Who Holds Influence over WTO Jurisprudence?
Mark Daku* and Krzysztof J. Pelc**

ABSTRACT
What shapes jurisprudence in international law? States dedicate considerable effort trying to influence not only the outcome, but also the content, of legal rulings. The stakes are high, as these legal opinions can redefine the meaning of the rules. Looking at the World Trade Organization, we ask whether some countries hold more influence over jurisprudence than others, and what such influence depends on. Using text analyses of every country submission in every ruling in the WTO era, we test a number of theoretical expectations. We find that some countries do appear to hold greater sway over the content of rulings than others: a country’s wealth, but especially its legal experience, account for much of this variation. Secondly, countries’ influence over the content of the verdict varies according to how novel the legal issue being ruled on is: states have more influence over the content of the ruling, the less precedent judges have to rely on in terms of prior legal decisions. The salience of the case and judges’ legal experience also follow expectations, as both are shown to take away from countries’ influence. Overall, the degree to which countries’ submissions influence the content of rulings appears to vary systematically. Legal capacity affects not only countries’ ability to file disputes, but also their ability to affect the shape of the resulting jurisprudence.

I. INTRODUCTION
It is a sign of the immense success of global governance that states today can afford to express power only in muted ways, and at the margins. International rules, on their face, aim for a hard won egalitarianism. They elevate sovereignty, and temper power differentials between members. In many institutions, such as the one we focus on in this article, the World Trade Organization (WTO), the formal rules appear prima facie skewed towards poorer nations, through the institutionalization of ‘special and differential treatment’. In such settings, do the strong seek alternative means of exercising power?

A handful of studies address this question in the specific context of the WTO. They find, for instance, that powerful countries may be largely setting the agenda of

* Institute for Health and Social Policy, McGill University (mark.daku@mcgill.ca).
** Department of Political Science, McGill University (kj.pelc@mcgill.ca). We thank Christina Davis, Joost Pauwelyn, participants at the APSA 2015 meeting in San Francisco, the LSA 2016 meeting in New Orleans, and three anonymous referees for valuable comments. All remaining errors are our own.

1 The WTO touts these provisions prominently. See https://www.wto.org/english/tratop_e/devel_e/dev_special_differential_provisions_e.htm (visited 22 May 2017).
WTO negotiations, such that choices through consensus that appear open on their face are likely to yield outcomes that suit the powerful. Large countries may exploit the informality of the rules when the stakes suddenly run high. In general, we have learned that any ambiguity in the rules favors the reinsertion of power. It may also be that in practice, existing legal avenues are not available to all: the last resort in WTO dispute settlement is trade retaliation, a threat poor countries find hard to deliver credibly. A minimum level of legal capacity may also be required for governments to defend their rights. If this is so, then powerful countries hold a source of deterrence that small countries lack.

In short, the exercise of power in international organizations is most likely to occur at the margins, rather than through open threats or coercion. This makes identifying the sites of power within international institutions analytically challenging. In this respect, empirical studies lag behind sociological and reflexive theories that emphasize the ways in which actors shape the institutions they are in turn shaped by. The biggest gap between theory and empirical work in this respect may be nowhere greater than in the study of the exercise of power through language. While this has arguably been the central concern of social theory since the linguistic turn of the 1960s, empirical work has thus far fallen short of delivering on the fundamental insight that influence over the terms of the debate amounts to influence over its outcome. This is our focus in this article, where we examine how countries push for their preferred terms in the rulings that make up the jurisprudence that the rest of the membership subsequently refers to. In sum, we ask, who shapes jurisprudence?

The question of interpretation is among the most controversial in international law. Governments jealously guard their sovereignty; they are wary of unelected judges changing the meaning of negotiated obligations through an innovative legal interpretation. To this end, treaty texts themselves attempt to circumscribe adjudicatory powers. The WTO’s Dispute Settlement Understanding (DSU) specifies that its goal is to ‘preserve the rights and obligations of Members under the covered agreements’. Insofar as it does ‘clarify the existing provisions of those agreements’, it must do so ‘in accordance with customary rules of interpretation of public international law’. Most pointedly, Article 3.2 of the DSU also warns that ‘rulings of the
DSB cannot add to or diminish the rights and obligations provided in the covered agreements'.

Despite these textual safeguards, it is generally agreed that by applying treaty law to specific cases, judges inevitably do engage in some form of rule-making. And because members value coherence in jurisprudence, these rulings have a way of becoming ‘definitive interpretations’ of the agreements.9 In some instances, in fact, governments may tacitly allow for such ‘interstitial legislating’. When negotiating parties struggle to reach agreement over some provision, they may leave it intentionally vague, to be clarified through adjudication at a later date.

What is less often remarked on is how such delegation to adjudicators, in turn, reopens the process to influence by states. It shifts the semantic struggle from a legislative process to a judicial one, but the stakes remain just as high. As Jan Klabbers puts it, ‘whoever controls this process [of interpretation] controls the meaning of the treaty, and therewith controls whether or not the obligations resting upon him are bearable or onerous’.10

There are several clues that governments themselves are deeply aware of this fact; they recognize the importance of legal interpretation in international law, over and above the direction of rulings. The very selection of disputes by governments seems geared not only towards resolving commercial disagreements, but also towards the impact on jurisprudence of the resulting legal reasoning.11 Litigants appeal findings not only to reverse them, but also to modify their formulation.12 And countries spend considerable resources pushing the court to act in ways that have no effect on the direction of the ruling, such as when pleading for the court to exercise judicial economy, or, conversely, to ‘complete the analysis’ and render rulings on subordinate claims that do not bear on the court’s final recommendations.13 Does this strategic behavior pay off? Are countries able to influence jurisprudence, and are some better at it than others?

We expect that powerful countries may be no more likely to win disputes, but that they may be better at inserting their language into jurisprudence. Stated from the point of view of judges, we expect that panelists will be more likely to defer to countries in the formulation of their verdicts than in the direction of their rulings. Following existing work, we posit that the best proxy for power in the context of dispute settlement may not be wealth, so much as legal capacity, measured through

---

9 This, in spite of the indication to the contrary contained in Article IX:2 of the WTO Agreement, which reads ‘The Ministerial Conference and the General Council shall have the exclusive authority to adopt interpretations of this Agreement and of the Multilateral Trade Agreements.’
litigation experience. We also hypothesize that this influence should be greater in some contexts than in others. Specifically, when a dispute concerns an unprecedented legal issue, countries should have more room to target their arguments, and the court should become more likely to adopt the language contained in countries’ submissions. This might explain why countries rush litigation to influence the interpretation of novel legal issues, before others have a chance to do so.

Conversely, we expect that when jurisprudential stakes are high, that is, when a dispute is likely to have important systemic consequences in the future, the average influence of litigants should decrease, as the court asserts its independence, and becomes less likely to be swayed by the litigants’ reasoning. In this, we follow analogous beliefs advanced by scholars of the US Supreme Court. To provide further backing for our empirical approach, we also formulate expectations over (i) the relative influence of third parties and (ii) the impact of the legal experience of panelists.

We test these beliefs by extracting and assembling all the country submissions and rulings in every dispute in the WTO era. This yields a total of 1847 ‘voices’ that can be compared to one another. Using Jaccard similarity coefficients, we score the ‘proximity’ of each pair of texts through an evaluation of the words contained within each ruling. Specifically, we are interested in how close each country submission is to the associated panel and Appellate Body rulings. We then attempt to explain variation across these proximity measures within each dispute.

The key methodological challenge we face is to assess whether any proximity between submissions and rulings represents influence, or something more akin to anticipation. What we know about countries’ incentives pushes towards the former, but we derive two hypotheses that allow us to test these competing alternatives. In both cases, our findings suggest that textual proximity is indeed catching influence, rather than anticipation. Finally, we spend some time discussing the strengths and weaknesses of the text analysis methods we employ here: while powerful tools, their results are often difficult to convey in an intuitive fashion. We suggest one way to tackle this weakness for the corpus we examine.

The results suggest that language is one more means through which the powerful assert their influence at the margins. One central irony is that those same countries, like the US, that profess to be most wary of judicial rule-making are also best at leveraging the judicial process to define the meaning of the rules in a way that suits their interests. We show that by successfully inserting their preferred terms into the court’s final ruling, countries with greater legal capacity are able to affect the way in which issues are subsequently discussed. ‘Caesar is also lord over grammar’.

---

16 The phrase comes from Carl Schmitt, the legal theorist of the dark edges of the political, who added ‘whoever has true power is able to determine the content of concepts and words’. See Carl Schmitt, ‘Forms of Modern Imperialism in International Law’, in Stephen Legg (ed.), Spatiality, Sovereignty and Carl Schmitt: Geographies of the Nomos (London: Routledge (1933) 2011), at 44.
II. LEGAL INTERPRETATION IN THE WTO

As with all legal texts, WTO rules are not without ambivalence. They do not specify a precise course of action for every situation—if they did, fewer cases would call for dispute settlement in the first place. It is thus the task of the court to apply the rules to a given setting. But the implied level of delegation to the court is threatening to governments, who jealously guard their sovereignty. Insofar as the interpretation of courts may redefine the meaning of countries’ obligations, it treads on the function of rule-making, the traditional preserve of governments. States vehemently oppose ceding this function to courts. Consider the US’ most recent 2017 Trade Policy Agenda, which lists as one of its ‘key objectives’:

‘Resisting efforts by other countries—or Members of international bodies like the World Trade Organization (WTO)—to advance interpretations that would weaken the rights and benefits of, or increase the obligations under, the various trade agreements to which the United States is a party.’17

This statement is first, a nod to the aforementioned Article 3.2 of the DSU, which specifies that rulings ‘cannot add to or diminish’ the obligations of Member States. As for the mention of other countries having the same ability to push through such unfavorable interpretations, it serves as an apt motivation for this article’s question: are some countries better than others at influencing jurisprudence?

The question matters, since the implication of the inevitable ambivalence of the rules is that despite doctrinal warnings to the contrary, courts are continually redefining the meaning of the rules through their interpretations.18 In fact, the ambiguity built into the rules is partly by design. Warnings about judicial activism notwithstanding, country negotiators may prefer to delegate decisions over politically sensitive issues to subsequent deliberations by the court, if they think a political compromise is unlikely. Technological changes and exogenous shocks may also make it difficult to foresee, at the time of the rules’ design, future settings where rules will have to be applied. For all these reasons, states may tacitly allow international courts to ‘legislate interstitially’, filling in gaps in interpretation, clarifying lacunae, and applying the rules to unforeseen settings.19

Of course, courts do not do so do of their own accord. To begin with, international courts can only rule on those matters brought before them. In the case of the WTO, only governments can bring cases against one another. These governments act strategically, and will not challenge a foreign measure if the resulting jurisprudence is

likely to come back and haunt them. This is the common explanation for why rich countries do not challenge each other’s agricultural subsidies.

The court is similarly constrained once litigation begins, by discursive, constitutional and political constraints. It is bound by its terms of reference, spelt out in DSU Article 7, to adjudicating only those claims raised by the litigants. Before arguments are even presented, complainants choose under what claims to file. Panelists or Appellate Body members cannot go beyond the confines determined by the complainant’s initial claims, and the respondent’s defense, even if they believe that additional legal provisions may be relevant. For instance, if a defendant chooses not to justify a measure through an exception under Article XX’s General Exceptions, the court cannot go ahead and adjudicate the exception’s applicability, even if it is meritorious. In the past, some respondents have preferred to be found in formal violation than to have the invocation of an exception be adjudicated, from fear of the jurisprudential repercussions of the court’s ruling. These choices are made before any arguments are ever formulated. But governments afford the greatest care to crafting their legal arguments, which they submit to the panel and AB, and on the basis of which the court makes its ruling. For these reasons, countries have considerable latitude to push for their favored interpretation, seeking to insert their preferred terms into the final ruling, and in this way to pull the resulting jurisprudence towards their ideal point.

The rewards for successfully influencing jurisprudence are considerable. Precisely because these are often the issues that negotiators were unable to agree on during multilateral negotiations, the stakes may be especially high. Once an interpretation is set down in a ruling, others will face an uphill battle to try and dislodge it. Thus Klabbers can quip that ‘interpretation is the continuation of treaty negotiations by other means’. This is because even in public international law, where binding precedent does not formally apply, the interpretations of courts are quickly reified. In this way, early in the WTO regime, the European Union succeeded in considerably increasing the threshold for the use of safeguards, which it lacked a legal regime for.

In an earlier classic example, the threshold for environmental exceptions was similarly rewritten through interpretation. In 1986, the USA challenged Canada’s export ban on herring and salmon. The case turned on whether Canada could justify its policy through an exception under Article XX(g), which reads that nothing in the agreement can prevent a country from adopting a measure “relating to the conservation of exhaustible natural resources”. The phrase “relating to” presented considerable ambiguity, and Canada pushed for a loose interpretation, while the US pushed for a

22 This was notably the case of the US choosing not to justify its 1980s embargo on Nicaragua through the security exception of Article XXI, even as this would have by all accounts resulted in no finding of violation. See above n 4 for a full discussion.
23 See above n 10 at 406.
24 See above n 21 and above n 11.
narrower interpretation.\textsuperscript{25} Canada lost the semantic struggle. Following the US’ cue, the panel ruled that “relating to”, in this context, meant that the measure had to be “primarily aimed at” conservation. Canada lost the case on this basis. But more importantly, the threshold for defenses alleging environmental conservation objectives was increased for the trade regime as a whole.\textsuperscript{26}

In fact, in the second case before the WTO, US – Gasoline, where Article XX(g) was invoked, neither litigant, nor any of the third parties, even tried to dispute the notion that the term ‘related to’ was now equivalent to ‘primarily aimed at’. The 1986 interpretation had been reified, and was guiding the interpretation of the rules a decade later. On appeal, the AB found itself compelled to note that ‘the phrase “primarily aimed at” is not itself treaty language’. Yet it went ahead and ruled as if it had been,\textsuperscript{27} finding that the US measure was indeed ‘primarily aimed at’ conservation, and thus fell within the scope of the Article XX(g) exception.

Just as countries fight for their interests during multilateral negotiations, they do the same during the clashes that occur in litigation. If a country can make the institution adopt its preferred terms of reference, its interests are considerably advanced. The question we ask is, do countries vary in their ability to achieve this outcome?

A. The incentives of the court

A large multidisciplinary literature has charted the particular challenge of international courts that must at once apply the letter and the spirit of the law, and contend with the interests of states.\textsuperscript{28} If courts fail to do the latter, states can flout rulings or rewrite the rules themselves, two outcomes that courts desperately seek to avoid. This risk is only increased when dealing with the foundational members of an institution—Germany, France, and the UK in the case of the European Union; the US and the EU in the case of the WTO. Yet observably deferring to political power would be just as costly, as the authority of courts flows from their being perceived as autonomous. A consistent bias in the verdicts of courts towards the powerful would

\textsuperscript{25} Canada stated ‘the issue was not whether these measures were conservation measures per se or even whether they were “essential” or “necessary” to the conservation regime’. The USA countered that ‘the primary motivation and effect must be conservation’. (GATT Panel Report on Canada – Herring and Salmon. 1986.)

\textsuperscript{26} See the WTO Analytical Index: GATT 1994, para 937. See also Ingo Venzke, ‘Making General Exceptions: The Spell of Precedents in Developing Article XX GATT into Standards for Domestic Regulatory Policy’, 12 German Law Journal, 1111 (2011) and above n 18 for an in-depth discussion.

\textsuperscript{27} As the AB ruled, ‘All the participants and the third participants in this appeal accept the propriety and applicability of the view of the Herring and Salmon report and the Panel Report that a measure must be “primarily aimed at” the conservation of exhaustible natural resources in order to fall within the scope of Article XX(g). Accordingly, we see no need to examine this point further . . . ’ United States – Standards for Reformulated and Conventional Gasoline. WT/DS2/AB/R, 19.

be quickly denounced—not least by the powerful countries themselves, to whom the
court’s credibility is arguably most valuable.

By comparison, deferring to power in the content of the ruling is less observable, and
thus less prone to censure, than any bias in the direction of verdicts. This be-
lief, that power is more likely to manifest itself at the margins, is the basis for our ar-
argument. Our premise is that the content of rulings is more likely to reflect power
than final verdicts. Judges should be more likely to defer to countries in the formulation
of their rulings, which are less observable, than in the direction of their verdicts.

Of course, influence of countries’ submissions is not reducible to the court’s will-
ningness to please litigants. The court may thus choose to borrow from the countries’
submissions because they are high-quality analyses of the law applicable to the meas-
ures. Or judges may be lazy, and rely on countries’ arguments to save time. The exact
reason why panelists or AB members pick up on countries’ wording in their rulings
matters little for the implications of this article. Whether judges are truly persuaded
by the submissions, or pick up on them mostly to economize judicial resources, what
matters to us is whether countries are able to inscribe their preferred language into
jurisprudence, and whether this varies systematically. Are some countries able to af-
flect the trade regime’s terms of reference more than others?

In some settings, the court may have an incentive to reassert its own autonomy. If
a dispute is particularly salient, and risks having considerable implications for coming
cases, the court may seek to shape jurisprudence according to its own interpretation
of the law, and purposefully rely to a lesser degree on countries’ interpretations.
Scholars of the USSC hold analogous expectations. As Pamela Corley argues, in ‘sali-
ient cases, the justices may spend more time and energy drafting the opinion, realiz-
ing that the words and phrases used will be analyzed and dissected by academic
scholars’. We test this expectation in our analysis: all else equal, do countries exert
relatively less influence over jurisprudence in those cases that are likely to have more
systemic impact?

III. MEASURING PROXIMITY TO CAPTURE INFLUENCE
We are interested in who shapes jurisprudence. To address this question, we extract
and assemble all the country submissions pertaining to every dispute in the WTO,
and the associated ruling, both at the panel and the appellate stage. This entails
manually gathering every country’s briefings and submissions in each case. We then
extract the panel and AB’s rulings, retaining only their reasoning and conclusion, and
leaving aside the traditional reiteration of the parties’ arguments.

First, we remove common stop words (e.g. ‘the’, ‘and’, as well as corpus-specific
words). We then calculate a co-occurrence matrix using the Wordstat software pack-
ge. This matrix reports Jaccard’s coefficient, a simple, yet revealing, measure of set

29 See Erik Voeten, ‘The Impartiality of International Judges: Evidence from The European Court of
Human Rights’, 102 American Political Science Review (2008), 417–33. See also Marc L. Busch and
University, 2017, on how judges calibrate the content of rulings to handle political controversy.
30 Pamela C. Corley, ‘The Supreme Court and Opinion Content the Influence of Parties’ Briefs’, 61(3)
Political Research Quarterly (2008), 468–78, at 471.
similarity, for each pair of cases. We keep only those observations that show the proximity of a country submission to its associated ruling, either at the panel or AB stage. The result is a dataset of 2771 observations at the dispute-stage-country level. Each observation corresponds to a country’s submission as a complainant, a defendant, or a third party in a given dispute. In some of our estimations below, we are interested in a narrower sample, looking for instance only at complainant submissions. The proximity measure is a score bound between 0 and 1, where scores closer to 1 designate more proximate pairs of texts.

There is an important time aspect to these data: all country submissions, both by litigants and by third parties, occur prior to the ruling. In a typical dispute, the litigants begin by offering their ‘requests for findings’, a brief list of the claims made by each party. Then, each litigant, starting with the complainant, offers its arguments, where it makes claims about the legality of the measure by interpreting the legal texts, often in view of existing jurisprudence, if prior cases have dealt with a similar legal issue, and rules of legal interpretation. If any third parties have joined the dispute, they then present their submissions, either in oral or in written form. The panel or AB then presents its findings, by first reviewing the arguments of the parties, and then presenting its own legal reasoning. Each report concludes with a series of findings and recommendations, where the legality of each claim is ruled on.

The different categories of country submissions afford us a useful plausibility probe of our proximity measure. We can start by verifying that the litigants bear a stronger relation to the final ruling than third parties, on average. Indeed, while third parties often take sides in the dispute (usually siding with the complainant), they also sometimes pursue a narrow agenda of their own. At other times, they do the opposite, and take a broad stance, warning the court of the delicate nature of the issue at hand, and its dangerous implications for jurisprudence. Thus, we should expect that third party submissions would bear a weaker relationship with the final ruling than the litigants’ own submissions.

Figure 1 shows the average proximity between a country’s submission and the ruling in that dispute by the type of voice. As expected, third party submissions are significantly less proximate to the ruling than either litigant. The proximity of third parties to the ruling is also more widely dispersed, in accordance with the variety of their intentions. Of equal interest is the fact that the submissions of complainants and defendants are about equidistant from the reasoning in the ruling, on average. If

---

31 Jaccard’s coefficient is computed as \( \frac{a}{a + b + c} \), where \( a \) represents words that occur in both cases, and \( b \) and \( c \) represent instances where a word appears in one case but not the other. Our objective is similar to that of Manfred Elsig, Todd Allee, and Andrew Lugg. The Ties Between the World Trade Organization and Preferential Trade Agreements: A Textual Analysis. Journal of International Economic Law, 2017, in this issue, since they too are interested in comparing texts—in their case, those of individual PTAs with the WTO texts. Since they are interested in evidence of copying and pasting between texts they rely on strict matches of six identical words in the same sequence. We are interested in evidence of influence, something that we expect will be expressed through the use of similar, though not necessarily identical phrases. Thus, our use of Jaccard’s coefficient does not rely on identical texts, but instead looks for co-occurrence of words at the document level.

32 In this way, panelists and AB judges will often refer to the Vienna Convention on the Law of Treaties.

33 See above n 13.
anything, defendants’ submissions are slightly closer, but this difference is not statistically significant across the corpus.

A. Influence versus anticipation

The key premise of our analysis is that the proximity between country submissions on rulings proxies for the influence of the first on the second. This assumption is consistent with the sequence of the different voices. It is also consistent with what we know of countries’ incentives. Countries stand to gain from influencing the court’s legal reasoning, and anecdotal evidence suggests that they try hard to do so. Their incentive is to offer the court discursive resources that will be picked up on in the final ruling. Finally, analogous studies in the case of the US Supreme Court, which have analyzed the influence of legal briefs on the rhetoric of the majority opinion, have routinely relied on the premise that similarity serves as a proxy for influence without problematizing it.34

The alternative we nonetheless want to guard against is that any proximity between countries’ submissions and the ruling is due not to influence on, but to anticipation of, the legal opinion. States may derive some benefit simply from having the final ruling ‘agree’ with their submission. Such agreement might be sold as a success to domestic audiences, for instance. The theoretical distinction between influence and anticipation is plain, but it relies on a counterfactual which we cannot evaluate directly. When we talk of influence, we mean that a ruling would be different, had this submission not been made. When we talk of anticipation, we mean that the submission sought to approximate the expected content of the ruling, rather than to affect it in a given direction. In that case, any proximity between the two would not be affected by whether the court ever read the submission. Anticipation is possible because

34 See Corley, above n 30.
countries’ submissions and the ruling are referring to the same legal issue within the same legal framework.

It is considerably harder to make this distinction empirically. To do so, the mechanisms of influence and anticipation must have divergent testable implications. We generate two sets of expectations that achieve this. First, we test the impact of how established the legal issue examined in every dispute is. Second, we formulate expectations over the impact of panelists’ experience on proximity with the average submission. These act both as substantive hypotheses in their own right, and as tests of the proximity measure’s validity.

B. Legal novelty and panelist experience

Countries strategically choose which measures to challenge. Existing work shows that some cases are filed more for their precedential value than for the underlying commercial concern.\(^{35}\) Anecdotal evidence suggests that countries look for good ‘test cases’ to push their favored interpretation of a provision that has not yet been ruled on. The European Union thus filed a number of challenges of safeguards in the WTO’s early days, and the US did the same on intellectual property as soon as the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPs) was implemented. All countries hope to have influence over the ruling’s content in such seminal cases. This leads to a testable expectation. If proximity between submissions and rulings is capturing influence, then we would expect that less established cases, over legal measures that have never been challenged, would be associated with greater proximity. This belief is also in keeping with the findings of the literature on the European Court of Justice (ECJ), which shows that judges assert their independence more forcefully when they can draw on existing precedent.\(^{36}\)

Conversely, if what we are measuring is anticipation, more legally unprecedented cases should lead to lower proximity. Litigants would have a harder time anticipating the likely ruling, even if this was their primary aim. The court has more discretion in the absence of relevant case law, since it is not bound by prior rulings, and its reasoning should thus be less predictable. In sum, more legally established claims should be associated with lower proximity between submissions and the ruling if we are dealing with influence, and higher proximity if, instead, we are dealing with anticipation. We code how legally established each case is by counting the number of times a given legal claim (e.g. TBT Article 2.2) has been brought up in a dispute prior to the case under observation, using the ruling date as our ordering variable.

In our sample, LEGAL HISTORY OF DISPUTE ranges from 0 to 81, and we use its log in our estimations.

The legal experience of panelists offers us another opportunity to distinguish these two competing effects. The experience of panelists has been shown to matter


\(^{36}\) See above n 28.
in the WTO: more experienced panel chairs (panels are composed of three individuals, of whom one chair) are associated with significantly lower odds of the AB overturning a ruling on appeal.\textsuperscript{37} All panelists prefer to have their rulings stand on appeal, but those with greater knowledge of the law and the institution’s procedures are more likely to achieve this. We expect that legal experience should make individual panelists less prone to borrowing directly from the reasoning of either litigant, and more likely to offer their own original reasoning instead. Conversely, the decisions of panelists with more legal experience should be easier to anticipate. That is also why their decisions are less likely to be overturned on appeal: experienced panelists will offer the reasoning that the AB, as the institution’s standing legal body, will find most apt. They also have a longer record of rulings for litigants to pour over, if anticipation is their main goal.

It follows that if countries are trying to anticipate rulings rather than influence them, the consistency that comes with judicial experience should have a positive effect on proximity between country submissions and the final ruling. We count the number of prior cases each panelist in each dispute has sat on, using the Horn and Mavroidis data hosted by the World Bank.\textsuperscript{38} We then sum this prior experience across all panelists in each dispute to obtain our variable PANELIST EXPERIENCE.

Both sets of expectations are assessed in Table 1. The first column of Table 1 considers our maximal sample, that is, it estimates the proximity of all country submissions to the associated ruling, whether at the panel stage or the AB stage. The results show that as expected, the more legally established a dispute is, the smaller each country’s influence over the content of the final ruling becomes. Put otherwise, more legally novel disputes afford countries greater influence over jurisprudence.

In the second and third columns of Table 1, we limit the sample to the panel ruling stage, since we are concerned with panelist experience.\textsuperscript{39} Column 2 shows the univariate regression results with PANELIST EXPERIENCE, and then Column 3 adds LEGAL HISTORY OF DISPUTE on this smaller sample. In both cases, panelist experience is negatively related to countries’ influence over the content of the ruling, suggesting that, as expected, panelists grow more autonomous as they acquire experience of WTO proceedings, and become less likely to rely on country submissions. Apart from demonstrating the effects of legal novelty and panelist experience on the average influence of countries over the content of the ruling, the results of Table 1 should offer greater confidence that, in keeping with analogous studies of legal briefs in the US Supreme Court, the observed proximity between countries’ submissions and the associated ruling does reflect influence, rather than some form of anticipation by countries. That alternative was associated with the expectation of both LEGAL HISTORY OF DISPUTE and PANELIST EXPERIENCE being positively related to proximity.


\textsuperscript{38} We are careful to count only distinct cases in this respect: if a panel renders two distinct panel reports to two complainants over the exact same legal issue, as in the recent COOL report, where the two co-complainants, Canada and Mexico, received technically separate panel reports, we count this as a single case for the purpose of judicial experience.

\textsuperscript{39} Experience is less meaningful at the AB stage, since it is a standing body. The AB stage can never feature three ‘first time’ judges, while the panel often does.
C. Who holds the most influence?

We argue that power is most likely to be exercised at the margins. In the case of the WTO, such power can take on two forms: it can be proxied by wealth, or by capacity. The belief we want to test in each case is that more powerful countries, either in terms of wealth or legal capacity, are better able to get the court to adopt the language of their submissions in its final rulings. We begin by visualizing the variation in influence across countries in the simplest way possible. Figure 2 shows a dot plot of the average proximity of countries’ submissions to the final ruling. We consider the 15 most frequent users of dispute settlement, to avoid having countries with a single submission skewing the results. This figure makes no distinction as to the role countries were playing in the dispute, and mixes in the panel and AB stage.\(^{40}\)

This quick snapshot of influence appears to match expectations: submissions by the USA and the EU appear as most influential. In fact, US submissions appear significantly more influential than any other country. In keeping with the research on legal capacity and learning,\(^{41}\) a number of far less economically powerful countries, but with considerable judicial experience in the trade regime, such as Argentina and Thailand, figure among the countries with the most influential submissions.

For a more precise test, we estimate an OLS model where the dependent variable is once again the proximity between a country submission and its associated ruling.\(^{42}\) Among our explanatory variables, we include a measure of countries’ legal experience. We code this variable, COUNTRY LEGAL EXPERIENCE, as the number of prior cases the country has offered a submission on, in any capacity. What we care about is experience in actually formulating and submitting a country’s position on the legal measures at issue,

\(^{40}\) That is, submissions at both the panel and AB level are included into the calculation of averages, but the proximity scores are based on the appropriate level: e.g. submissions at the panel level are only compared to the panel report.

\(^{41}\) See above n 6 and above n 14.

\(^{42}\) The dataset is a cross-section with no time dimension. As before, robust standard errors are clustered on the common dispute.
and thus we do not count e.g. a country’s third party participation if that country did not submit its views to the panel. Our second proxy for capacity is GDP per capita, coded as the log of GDP per capita in constant US dollars, from the World Development Indicators, for the year of the dispute’s initiation. We then add back the two variables from our prior analysis, Legal History of Dispute and Panelist Experience.

The results are shown in Table 2. The first three columns consider the influence of all country submissions on the final ruling. Yet most of the literature is concerned especially with complainants’ legal capacity, since they are the initiators of challenges. From the point of view of the trade regime, the concern is whether countries are able to avail themselves of the institution’s resources to challenge violations against them. Transposed to our analysis, this question becomes, are countries that challenge violations against them equally able to shape the jurisprudential outcome? To test this, in columns 3–6, we rerun the same estimations on a narrower sample containing only complainant submissions and their proximity to the associated ruling, whether at the panel or AB stage. Disputes with multiple complainants result in separate observations, which is one more reason why robust standard errors are clustered by dispute. Finally, to ensure consistency with our results above, in the last column we add back the two prior variables of interest, Legal History of Dispute and Panelist Experience. Here, the sample looks at all submissions from complainants, defendants, and third parties alike, but it is naturally limited to the panel stage, since we are interested in the effect of panelist experience.

The results are consistent across both our full sample and the complainants-only sample: legal capacity matters. And experience as a litigant matters more than wealth, in a more systematic fashion. Experience swamps wealth. This offers further

43 While both show significant effects in univariate estimations, when we run them alongside one another, the effect of wealth remains positive, but loses statistical significance, while the effect of experience remains highly significant.
support to studies that argue that countries can build legal capacity by being engaged in dispute settlement proceedings. The greater effect of experience remains when we control for LEGAL HISTORY OF DISPUTE and PANELIST EXPERIENCE, which also show consistent effects with the results in Table 1.

A question remains as to what type of experience matters, and to what degree our variables are correlated. To offer a clearer picture, we include a list of simple correlations in Table 3 between disaggregated experience variables and influence. What comes across is that experience as a complainant has the strongest association with complainants’ influence, but experience as a defendant is not far behind. Experience as a third party is also positively correlated with influence, though at a lesser degree. Finally, wealth shows an even smaller correlation, though direct comparisons are hard to draw in this respect, given the different units of measure.

The premise of this article is that countries hold more influence at the margins. Put in terms of the incentives of an international court, because the content of rulings is harder to scrutinize than ‘who wins’, politically savvy judges are more likely to defer to country interests in the content of the ruling than in its direction. For the sake of comparison, we thus show the effects of legal capacity on the direction of the verdict itself. To do this, we code the direction of every claim in every WTO dispute from 1995 to 2013. All told, panels have delivered 1429 findings on 820 individual claims. We first collapse these findings at the claim level, and then collapse claims at the dispute level, to obtain the number of claims won by the complainant. We divide

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experience as litigant</strong></td>
<td>0.031***</td>
<td>0.023***</td>
<td>0.038***</td>
<td>0.029***</td>
<td>0.034***</td>
<td>0.023***</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td><strong>GDP/cap. (log)</strong></td>
<td>0.022***</td>
<td>0.011</td>
<td>0.021*</td>
<td>0.009</td>
<td>–0.002</td>
<td>0.022***</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td><strong>Judicial experience of panel</strong></td>
<td>–0.085***</td>
<td>–0.025**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Legal novelty of dispute</strong></td>
<td>–0.025**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>0.560***</td>
<td>0.427***</td>
<td>0.477***</td>
<td>0.675***</td>
<td>0.565***</td>
<td>0.614***</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.04)</td>
<td>(0.12)</td>
<td>(0.11)</td>
<td>(0.07)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>2771</td>
<td>2633</td>
<td>2633</td>
<td>597</td>
<td>560</td>
<td>560</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.03</td>
<td>0.01</td>
<td>0.02</td>
<td>0.05</td>
<td>0.02</td>
<td>0.05</td>
</tr>
</tbody>
</table>

DV is textual proximity between a country submission and the associated ruling. Columns 1–3 consider all country submissions and both panel and AB rulings. Columns 4–6 consider only complainant submissions and both panel and AB rulings. Column 7 considers all country submissions, but only on panel reports. Robust standard errors clustered on dispute.

* p < 0.10, ** p < 0.05, *** p < 0.01.

44 See above n 6 and above n 14.
this number by the total number of claims ruled on, to obtain the percentage of claims won by the complainant, which varies from 0 to 1; its mean is 0.74 for panel rulings, and 0.69 for appellate rulings. The results are shown in Table 4, where we estimate the odds of a pro-complainant verdict, using the same explanatory variables as in our estimation of influence.

What Table 4 shows is that capacity, however measured, does not help countries ‘win cases’. Column 1 considers all verdicts, while Column 2 is restricted to panel rulings, and Column 3 to AB rulings. Neither wealth nor legal experience as a litigant appears to increase the odds of a pro-complainant verdict.\(^45\) Taken together, the contrasting findings of Table 2 and 4 support our central contention that power is more likely to be exercised at the margins. Judges are wary of deferring to political power in the direction of their verdicts, while they seem more likely to do so in the content of these verdicts.

Recall that we also held beliefs about what might make judges less likely to lean directly on the litigants’ arguments. If a case promises to have systemic repercussions going forward, the court may be especially careful to craft a legal opinion that takes future applications in consideration. In doing so, panelists and AB members may be less likely to borrow their language directly from the litigant’s submissions, and more likely to rely on their own interpretation of the legal texts.

To capture such salience, we exploit the procedure by which third parties join a dispute. All countries can request to join a dispute as a third party, which allows them to be present during otherwise private deliberations, and submit their views to the court. Crucially for our purpose, third parties can join either by claiming a ‘trade interest’, if they have a direct economic stake in the dispute (if the dispute concerns solar energy, then all other solar energy producers may want to join), or they can cite a ‘systemic interest’. They do the latter if they can demonstrate that the dispute

---

\(^{45}\) Ideally, the direction of the ruling would be estimated in a two-stage regression, where the first stage is the likelihood of a ruling in the first place. Indeed, there is reason to believe that countries with high capacity are better at settling cases before a panel is even formed, and are thus left with a different pool of cases for litigation. Although accounting for this falls out of the empirical scope of the present article, Leslie Johns and Krzysztof J. Pelc perform a similar two-stage estimation, and also find no effect for either complainant or defendant GDP on the direction of the ruling. See: Leslie Johns and Krzysztof J. Pelc, ‘Free-riding on Enforcement in the WTO’, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2782560, 2015 (visited 22 May 2017).
Table 4. Effect of legal capacity on direction of the verdict

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All rulings</td>
<td>Panel rulings</td>
<td>AB rulings</td>
</tr>
<tr>
<td>Experience as litigant</td>
<td>−0.016</td>
<td>−0.072***</td>
<td>−0.019</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>GDP per cap. (log)</td>
<td>−0.032</td>
<td>0.036</td>
<td>−0.032</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.038***</td>
<td>0.598***</td>
<td>1.058***</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.20)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>N</td>
<td>557</td>
<td>314</td>
<td>235</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.02</td>
<td>0.05</td>
<td>0.02</td>
</tr>
</tbody>
</table>

DV is proportion of claims ruled in favor of complainant. Robust standard errors clustered on dispute.

*p < 0.10, **p < 0.05, *** p < 0.01.

has a bearing on the membership as a whole. The resulting variable, Systemic, is a variable coded as 1 if any third party joined by citing a systemic interest, and 0 otherwise. 68% of the disputes that proceed to a ruling feature at least one such third party.46

In Table 5, we estimate countries’ influence in view of whether the dispute has systemic implications. The first column shows a parsimonious estimation, while the second column adds the variables from Table 2. As expected, disputes with systemic implications appear to reduce the influence of all country submissions. When a dispute holds important implications for jurisprudence, the court appears to follow its own interpretation of the law, and rely less on that of governments. This effect is robust to controls for country capacity. Equally interesting, even once we account for the systemic importance of disputes, the relationship identified above in Table 2 holds: while both litigant experience and country wealth are positively related to influence, only litigant experience is highly significant.

The relationship between proximity and systemic disputes should be read as a positive sign: judges appear able to recognize those cases that have significant systemic implications, and they reassert their own authority in crafting the legal reasoning in these cases, knowing that it is likely to affect jurisprudence in the years to come.

D. Signs of learning? The case of China

The evidence above suggests that legal capacity has an observable effect on a country’s influence over the content of the ruling, and that this effect trumps the effect of sheer economic wealth. One implication is that it should be possible to trace a country’s rising influence as it gains in legal capacity. No country is brandished more often as an example of learning and growing legal capacity than China. Following its entry into the WTO in late 2001, China joined practically every WTO dispute as a third party. Most observers agreed that China’s main goal in doing so was to gain

46 We ensure that the results are robust to a count version of this variable, coded as the number of third parties in any given dispute that joined citing systemic interests.
experience, and that since its accession, ‘China has used its experience in the WTO to build up its capabilities, learn about how the organisation operates, influence the WTO’s procedures and defend its own economic interests’. These widespread anecdotal beliefs produce a further testable implication. Can we observe China’s influence over proceedings increasing over time? To answer this question, we construct a sample consisting of all of China’s submissions, and regress their influence across time.

The results are shown graphically in a scatterplot on Figure 3. There, we can see how across a single decade, China has effectively doubled its average level of influence over panel and AB rulings. Every dot in Figure 3 in the scatterplot indicates the proximity of China’s submission in a dispute that it participated in. The solid red line shows the trend across time. This sharp upwards trend is all the more striking given that, in keeping with the findings above concerning legal novelty, the influence of the average country on proceedings actually decreased over the same period. For comparison, we add a dashed gray line that represents the trend in proximity of all other countries’ submissions across time, not including China.

When we regress China’s influence on a time indicator, we obtain the same highly positive relationship, which is statistically significant at the 0.001 level. The negative trend for all other countries also proves statistically significant, but only at the

Table 5. Effect of systemic disputes

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systemic dispute</td>
<td>-0.080***</td>
<td>-0.085***</td>
<td>-0.068***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Experience as litigant (log)</td>
<td>0.025***</td>
<td>0.029***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>GDP per cap. (log)</td>
<td>0.010</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>Legal novelty of dispute</td>
<td></td>
<td></td>
<td>-0.023**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.685***</td>
<td>0.544***</td>
<td>0.601***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.07)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>N</td>
<td>2771</td>
<td>2633</td>
<td>2633</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.02</td>
<td>0.05</td>
<td>0.06</td>
</tr>
</tbody>
</table>

DV is textual proximity between a country submission and the associated ruling. Robust standard errors clustered on dispute.

*p < 0.10, **p < 0.05, *** p < 0.01.


Table not shown to save space, available in replication materials.
0.05 level. In sum, the same learning that we show taking place in Tables 4 and 5 can be observed within a given country. As China gained legal experience within the trade regime, its influence over legal rulings grew in tandem.

E. The limitations of text analysis, and tracing seminal phrases

In the analyses above, we used automated text analysis to provide reliable measures of the proximity between pairs of texts, in a way that a human coder could not accurately assess. But this also means that the measures produced by these methods hold little intuition: it is difficult to convey what the US’ average proximity score to rulings of 0.751 actually corresponds to. These measures are useful in their relative relationships, as when we compare the US’ proximity score to that of Taiwan, which is 0.427. The utility of text analysis measures and their lack of intuition are two sides of the same coin. It is because a human reader cannot accurately gauge the proximity between two texts that a Jaccard’s similarity coefficient is useful; but this also means that there is no aspect of these texts that can be easily pointed to as illustration for the resulting measures. In fact, such limitations are in part what leads Broude et al. to opt for manual coding of the similarity between state regulatory space provisions between the Trans-Pacific Partnership (TPP) and existing agreements.50

---

50 Tomer Broude, Yoram Z. Haftel, and Alexander Thompson, ‘The Trans-Pacific Partnership and Regulatory Space: A Comparison Of Treaty Texts’, Journal of International Economic Law (2017). As the authors recognize in comparing the two approaches, ‘The computational approach is especially useful for studying which countries had the most influence over the language of [the treaty]’. But we share Broude et al.’s view that computational text analysis methods should serve to amplify human efforts rather than to substitute for them. The two are complementary.
To get a better handle on what influence looks like, in this section we seek to trace the actual words that countries voice and judges pick up on. To do this, we generate a list of all the phrases of eight words or more spoken by any country, and then taken up by the Appellate Body in its ruling over the corresponding dispute. There are nearly 19,000 such phrases across the corpus, though many of these are exact duplicates, or subtle variants of the same phrases. In many cases, these are formulaic citations from the legal texts. But many phrases are of the very type that we have in mind in this article. These might be called ‘seminal phrases’: phrases that a country successfully coins, and that subsequently gain currency.

The first thing to note is that the number of phrases a country succeeds in introducing into the AB report is highly correlated with the similarity scores above, though they are not equivalent. Overall, the bivariate correlation between the number of distinct eight-word phrases introduced into the report and that country’s similarity score with the report is 0.41. That is, the similarity scores reflect the number of eight-word phrases countries successfully insert into panel and AB reports, but the two are by no means equivalent. While the similarity score is undoubtedly a better overall measure of what we have in mind by influence, the number of phrases is more intuitive. It is worth offering a sense of a couple of these phrases below, which represent two main means of affecting jurisprudence: defining vague terms, and reifying practice.

These phrases offer a more tangible sense for how influence operates. In a dispute against Canada’s dairy subsidies, the AB argued:

As the Panel observed, the dictionary meaning of the word ‘payment’ is not limited to payments made in monetary form. In support of this, the Panel cited the Oxford English Dictionary, which defines ‘payment’ as ‘the remuneration of a person with money or its equivalent’. . . . Thus, according to these meanings, a ‘payment’ could be made in a form, other than money, that confers value, such as by way of goods or services.

But counter to what the AB suggested, this was not the panel’s own observation; it was New Zealand’s argument during the panel proceedings. The panel then picked up on it, and this definition made its way into the AB report. The choice of what constitutes ‘ordinary meaning’ in this case is far from evident. Under Article 9.1(c), ‘payments’ are ‘financed by virtue of governmental action’, which could plausibly be taken to mean that if payments must be ‘financed’, they must, as Canada argued, amount to an actual monetary payment. Canada also cited the French language version of the legal text, which translates payment as ‘versement’, and which Canada argued ‘meant literally to remit money’. The US also provided competing dictionary definitions, which the panel chose not to pick up.

In sum, there were plausible arguments advanced on all sides about the ‘ordinary meaning’ of the term ‘payment’ as understood in the Agreement on Agriculture. Yet the one retained was an expansive meaning put forth by New Zealand, and it is the one that appears in our set of seminal phrases. No part of Canada’s argument about the meaning in the French version made it into the AB’s conclusion, and none is recorded as a seminal phrase in the corresponding dataset.
The distributional effects are straightforward. If the term ‘payment’ includes not only direct monetary transfers, but also non-monetary payments in kind, then the subsidy rules of the Agreement on Agriculture increase considerably in scope. They begin to cover not only direct monetary transfers, but also any advantages conferred.

Consider another instance, India challenged a US safeguard on wool shirts and blouses. Much of the US defense consisted of arguing that ‘the burden was on India in the first instance to make a prima facie case that the United States’ application of a transitional safeguard [was inconsistent with the rules].’ In making its point, the US drew on ‘GATT practice’ to shift the onus onto the complainant, in opposition to India’s position. India argued that since the US safeguard was the result of a domestic investigation, and since it was an exceptional measure, the onus rested on the US to defend the legality of its application, which included demonstrating ‘serious damage’ to a US industry or ‘actual threat thereof’.

Never mind that the US lost the case on the merits; it successfully introduced the concept of ‘prima facie case’ into jurisprudence, permanently modifying the understanding of complainants’ burden of proof. The following year, in EC–Hormones, the AB spoke of “The general rule in a dispute settlement proceeding requiring a complaining party to establish a prima facie case of inconsistency with a provision of the SPS Agreement.” The US made sure to cite this position in its submission as a third party in a dispute between Thailand and Poland, where the argument over the need for the complainant to make a prima facie case of inconsistency became Thailand’s main rejoinder to Poland’s challenge of its antidumping duties on iron and steel. Through these invocations, the prima facie case standard spread from an argument in the narrow case of a safeguard within the Agreement on Textiles and Clothing to the SPS Agreement and, most significantly, the Agreement on Antidumping. But these and other similar invocations all harken back to the US argument that GATT ‘practice’ demanded that the complainant bear the burden of proof of presenting a prima facie case that a domestic investigation had not the requirements put on it by the articles the complainant claims were being flouted.

The appeal of text analysis methods is that they are able to quickly aggregate hundreds of occurrences like the New Zealand defining ‘payment’, and the US using a phrase to redefine the burden of proof on complainants. This is what makes it possible to estimate whether some countries are better at inserting their preferred language into jurisprudence on average. The similarity measures we use in the analyses above reflect not only these eight-word phrases, but the extent to which the submissions are then reflected in legal opinions in a more general sense.

52 Ibid, emphasis added.
IV. CONCLUSION

The insight that actors exercise power by influencing the terms of the debate is a foregone conclusion for much of social theory. In international relations, it has been especially woven into aspects of constructivism. Yet it has been largely overlooked by empirical studies. This is all the more surprising since global norms have all but eliminated reliance on open threats and coercion, making the settings most often studied by empiricists especially well-suited to the beliefs of social theory: within international institutions, states are likely to exercise power at the margins, as when pushing for their favored interpretation of the institution’s own rules.

Looking specifically at WTO law, our premise is that state power is likely to influence the content of legal rulings more than their direction. Indeed, judges are wary of appearing to defer to state power in their verdicts, but are more likely to be swayed by countries’ arguments in drafting the content of their opinions. The expectation we want to test is whether some countries are better at inserting their favored interpretations into the trade regime’s jurisprudence.

The short answer is yes. We find that some countries are systematically more likely to have the wording of their submissions picked up by judges both at the panel and the AB ruling stage. Yet the implications are not as sinister as might appear at first glance. In accordance with the literature on legal capacity, we find that countries’ influence depends more on legal experience than on wealth. This is why small economies that are frequent users of dispute settlement, like Argentina, appear to hold as much average influence over jurisprudence as wealthy countries like Japan. And experience as a defendant is almost as valuable in this respect as experience as a complainant. Both types of experience appear to matter more than a country’s wealth. By comparison, neither wealth nor judicial experience appear to increase countries’ odds of ‘winning cases’. Capacity, however measured, does not actually affect the direction of the ruling. Power is felt in nuanced ways, and at the margins.

We also conduct tests that increase our confidence that our measure of textual proximity is picking up influence on, rather than anticipation of, the ruling. To get at this, we show that panelist experience, and the extent to which the underlying legal issue has been established in prior cases, are both negatively related to countries’ influence on the final ruling. As we argue, if our measure of proximity were picking up some form of anticipation by countries of the content of the ruling, we should expect both these variables to have negative effects. Instead, we find that panelists with more experience follow their own interpretations to a greater degree, and that the less established the legal claims in a dispute, the more sway countries appear to have over the resulting legal opinion.

We then show that judges do rely on their own interpretations when a dispute carries considerable jurisprudential weight. In such ‘systemic’ cases, all countries come to hold less average influence over the outcome. The implication is that the court is able to distance itself from countries’ submissions when it feels its interpretation is likely to loom large in the future.

Finally, we apply these insights to the case of China, which is often presented as a country that, since its WTO accession in 2001, has quickly gained legal capacity that allows it to use the rules to its advantage. Consistently with this view, we find that China’s influence over jurisprudence has risen in a significant fashion through time. This is all the more striking since other countries’ average influence has waned over the same period, likely in reflection of how the accumulation of the *acquis* has dictated panelists’ and AB members’ decisions.

As social scientists embrace text analysis tools and techniques, we can hope to see empirical work providing support for what social theorists have long suspected: power is likely to be exerted most forcefully in those settings where it is least visible. Legal capacity may not determine countries’ ability to win disputes, but it does help them win the semantic struggles therein.