Debiasing International Economic Law

Sergio Puig*

Abstract

A flourishing number of bodies evaluate the conduct of government officials against broad standards, decide complex questions of scientific probity and calculate the present value of past decisions. The effects of implicit biases (systematic patterns of deviation from rationality in judgment) impact the assessment of these issues, which are central to international economic law. Such effects are well understood by psychologists and increasingly confirmed by experiments involving legal actors, including judges. In this article, I provide three concrete examples of implicit biases affecting international tax, trade and investment adjudication, and I call for the incorporation of mechanisms to overcome such biases as well as their strategic exploitation by litigants. At a conceptual level, I propose a typology to think of 'debiasing tools' for international adjudication – mechanisms that can act as a centrepiece of coordination of information rather than mere inoculants of the habits of mind on adjudicators. At a normative level, I pose that biases may impact confidence in dispute settlement systems and that both concerns for sovereignty and a predilection for negotiated solutions make international economic law ripe for testing these interventions.

1 Introduction

International law has experienced a slow but steady process of transformation over recent decades. Instead of resolving disputes through institutionalized bargaining, states have delegated the task to multiple international courts and tribunals (ICs). Judges, arbitrators and other adjudicatory decision-makers are now responsible for

^{*} Professor of Law, James E. Rogers College of Law, University of Arizona, Tucson, AZ, USA. Email: spuig@ email.arizona.edu.

Shany, 'Assessing the Effectiveness of International Courts: A Goal-Based Approach', 106 American Journal of International Law (AJIL) (2012) 225.

contextualizing general, often broad legal standards of government conduct; assessing complex questions of scientific probity; and calculating the present value of past financial decisions. Yet, for all the benefits that this delegation of dispute settlement has afforded to states and other actors, support for these bodies seems to be fading, in part due to valid critiques of the failures of such bodies.²

Among other important critiques about this shift in international law is the intensification of the effects of implicit biases – systematic patterns of deviation from rationality in judgment by the individuals making decisions.³ Such effects are well understood by psychologists and increasingly confirmed by empirical work involving international legal actors.⁴ One reason that implicit biases can be problematic in adjudicatory settings is that litigants may strategically exploit their effects to 'nudge' decision-makers in a non-coercive fashion. With growing concerns over the role of ICs and increasing attempts to de-legitimize the work of these bodies, it is important to consider ways to improve the quality and to increase the trust of adjudicatory decision-making by 'debiasing'.⁵

In this article, I use three concrete areas of international economic law — tax, trade and investment — to explain how implicit biases may undermine adjudicatory decision-making. I also discuss how in each of these three settings debiasing could improve the work of these bodies. First, I examine how in international economic disputes adjudicators may be especially susceptible to 'anchoring bias' — the influence of irrelevant information when making decisions that involve numerical choices. Here, I describe how a 'final offer' in tax arbitration — wherein each party makes one offer, and the arbitrator's task is simply to pick one of the two proposals — may help to moderate the impact of anchors. Using data from a survey experiment, I show the proclivity of adjudicators to reject ambitious positions and the arbitration format's role in mitigating the effect of irrelevant anchors.

Second, I discuss how decisions in the context of risk and uncertainty are often affected by 'hindsight bias' – the inclination, after an event has occurred, to see the event as having been predictable. This effect is likely to impact domestic determinations evaluated by ICs, including a decision to deny or invalidate a patent for not satisfying

- See infra Section 4.
- ³ Broude, 'Behavioral International Law', 163 University of Pennsylvania Law Review (2015) 1099; Van Aaken, 'Behavioral International Law and Economics', 55 Harvard International Law Journal (HILJ) (2014) 421; Jolls, Sunstein and Thaler, 'A Behavioral Approach to Law and Economics', 50 Stanford Law Review (1998) 1471.
- ⁴ Hafner-Burton, Victor and Lupu, 'Political Science Research on International Law: The State of the Field', 106 AJIL (2012) 47; Hafner-Burton et al., 'The Behavioral Revolution and International Relations', 71 International Organization (2017) S1, at S21–22; Shereshevsky and Noah, 'Does Exposure to Preparatory Work Affect Treaty Interpretation? An Experimental Study on International Law Students and Experts', 28 European Journal of International Law (EJIL) (2017) 1287.
- 5 R.H. Thaler and C.R. Sunstein, Nudge: Improving Decisions about Health, Wealth, and Happiness (2009). For criticism of this approach, see A. Timms, 'The Sameness of Cass Sunstein', New Republic (20 June 2019), available at https://newrepublic.com/article/154236/sameness-cass-sunstein (arguing that Sunstein's approach represents 'the refusal of ... technocratic managerialism to engage with new circumstances, even as many in his own liberal camp reevaluate their priors').

'non-obviousness', a requirement enshrined in the World Trade Organization's (WTO) Agreement on Trade-Related Aspects of Intellectual Property (TRIPS Agreement). Here, the format of presenting information and giving advice by experts involved in the case could help to ameliorate the effect, calling for safe harbour provisions that reward governments by creating a presumption of validity when governments apply procedures aimed at neutralizing implicit biases.

Third, I describe how party-appointed arbitrators – a feature not uncommon in international economic law adjudication – may be especially affected by 'affiliation bias', a predisposition to favour the appointing party in the arbitration. In investor-state dispute settlement (ISDS), blinding appointments – so that nominees do not know who appointed them – could improve investment law by relieving arbitrators from the effect.

Based on these three concrete examples, the article continues by explaining how to think about biases in relation to distorting effects of the information incentivized by adversarial settings. Institutional designers can push against the failures of international adjudication, rather than encourage the habits of mind that produced them, through different methods. The universe of plausible debiasing tools is vast. Nevertheless, one can think of these procedural or substantive reforms as tools addressing the biases of decision-makers or addressing the effects resulting from incentives generated in adjudicatory settings. By considering mechanisms that address biases both directly and indirectly, reformers may find more possibilities to improve the quality of the decisions of ICs.

Finally, I argue that not only do biases undermine the confidence in international dispute settlement, but they also may contribute to a process of de-legitimization. In particular, rising concerns for state 'sovereignty' and a preference for negotiated outcomes make international economic adjudication not only susceptible to this process but also especially ripe for addressing biases with debiasing tools. This is not to dismiss the complex political considerations involving international relations, the challenges currently faced by international law or the distinct environment of international economic disputes. But well-crafted debiasing mechanisms can help to temper some of the most pernicious effects of biases by stimulating settlements among disputing parties, limiting extreme positions of lawyers, discouraging polarization among adjudicators and encouraging narrow decisions by ICs. While the implementation of such mechanisms is not without costs and controversy, and need to be explored more thoroughly in context, governments and institutions concerned with the future of international law at a challenging time for ICs should experiment with well-tested strategies. To encourage experimentation, I suggest some institutional conditions that might make the use of debiasing tools more successful.

⁶ Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS Agreement) 1994, 1869 UNTS 299.

See Jolls and Sunstein, 'Debiasing through Law', 35 Journal of Legal Studies (JLS) (2006) 2016 (discussing debiasing strategies to risk perception, including jury instructions or labelling requirements).

2 Biases and Debiasing in International Economic Law

Social scientists have long recognized that objectivity can be undermined by information and other extraneous knowledge (fully or partially) irrelevant to a particular analysis. Legal scholars have studied how these heuristics may affect decision-making by judges, experts and juries. Although the interest is growing, less attention has been given to this phenomenon among international law actors, despite the fact that ICs are growing in importance and may be acutely prone to different types of biases. In this part, I provide three examples to illustrate how biases interact in the adjudication of international tax, trade and investment matters and discuss interventions to minimize the particular effects identified.

A Anchoring and International Tax

Anchoring bias gives salience to pieces of partially or fully irrelevant information when making decisions, especially in numeric judgments. ¹¹ It is a tendency to be 'unduly influenced by [an] initial figure' when estimating a value. ¹² Anchoring has far-reaching implications in law because legal decisions often involve calculations. Essentially, a number presented in a case becomes stuck in the mind and influences a judgment or settlement. Parties can exploit this effect by advancing an anchor that influences adjudicators. ¹³ As I explain below, anchoring can interact with other heuristics such as extremeness aversion (or compromise effect), a tendency to choose intermediate rather than extreme positions. ¹⁴ Anchoring affects international economic adjudication. In investment arbitration, tribunals often decide questions of damages. At the WTO, adjudicators assess the 'reasonable period of time' given to a defendant to implement an adverse ruling. In international tax, arbitrators allocate taxation rights, as I now explain.

- ⁸ Kahneman and Tversky, 'Prospect Theory: An Analysis of Decision under Risk', 47 Econometrica (1979) 263; see also D. Kahneman, Thinking, Fast and Slow (2011).
- Teichman and Eyal, 'Judicial Decisionmaking: A Behavioral Perspective', in E. Zamir and D. Teichman (eds), Oxford Handbook of Behavioral Economics and the Law (2014) 664; Guthrie, Rachlinski and Wistrich, 'Inside the Judicial Mind', 86 Cornell Law Review (2001) 778; C. Sunstein (ed.), Behavioral Law and Economics (2000); Hafner-Burton, Victor and Lupu, 'Political Science Research on International Law: The State of the Field', 106 AJIL (2012) 47.
- Drahozal, 'A Behavioral Analysis of Private Judging', 67 Law and Contemporary Problems (LCP) (2004) 105. Previous studies suggest that international judges are biased in favour of co-nationals. Posner and de Figueredo, 'Is the International Court of Justice Biased?', 34 JLS (2005) 599. For insights into biases, see McDermott, 'New Directions for Experimental Work in International Relations', 55 International Studies Quarterly (2011) 503.
- Feldman, Schurr and Teichman, 'Anchoring Legal Standards', 13 Journal of Empirical Legal Studies (2016) 298; Tversky and Kahneman, 'The Framing of Decisions and the Psychology of Choice', 211 Science (1981) 453.
- Orr and Guthrie, 'Anchoring, Information, Expertise, and Negotiation: New Insights from Meta-Analysis', 21 Ohio State Journal on Dispute Resolution (2006) 597.
- ¹³ Burke, 'Prosecutorial Passion, Cognitive Bias, and Plea Bargaining', 91 Marquette Law Review (2007) 183.
- Simonson and Tversky, 'Choice in Context: Tradeoff Contrast and Extremeness Aversion', 29 Journal of Marketing Research (1992) 281.

1 Allocation of Resources

In cases involving the allocation of resources, the anchoring effect could create a perverse incentive for litigants to claim more extreme positions in order to yield a more advantageous 'middle' point. While the amount of claimed damages is partially relevant, it is independent from the actual figure of damages. The proposed anchor may affect the incentives of the parties and the objectivity and fairness of the analysis. In domestic courts, the focus is usually on punitive and pain and suffering damages since these damages entail unclear criteria (which may increase discretion).

There is evidence of the strategic manipulation of anchoring in different international contexts. Arbitrators, rather than using actual data to calculate the amount permitted as trade retaliation in WTO cases, are said to 'split the difference'. ¹⁵ In ISDS, claimants tend to exaggerate their losses to obtain more favourable awards. ¹⁶ Even the International Court of Justice (ICJ) has been accused of issuing 'Salomonic' judgments in boundary disputes. ¹⁷ In other words, in determinations of resource allocation, ICs have a tendency to operate with limited data points, apply uncertain criteria and compromise (including in the presence of extreme positions). While often unintentional, these choices tend to amplify the distance between the positions of the disputing parties (that is, tend to polarize the process).

Experimental evidence confirms that, like many legal actors, international adjudicators are affected by anchors. ¹⁸ Debiasing from anchors, however, is notoriously difficult; most 'inoculants can create alternative anchors or facilitate over-correction'. ¹⁹ Yet minimizing the effect is essential for fair decisions involving resource allocation. One option is 'final offer' (or 'baseball') arbitration. This format is used in international tax determinations (and baseball disputes) to offset the incentives created by anchoring and take advantage of the compromise effect by shifting the latter onto the litigants.

2 Production of Information: Baseball Arbitration

The Organisation for Economic Co-operation and Development's Multilateral Instrument establishes a default process used for interstate disputes involving taxation rights over a source of income or profit. The relevant provision states that the 'panel shall select as its decision one of the proposed resolutions for the case submitted

For examples, see Pauwelyn, 'Baseball Arbitration to Resolve International Law Disputes: Hit or Miss?', 22 Florida Tax Review (2018) 40.

Kahale III, 'The Inaugural Brooklyn Lecture on International Business Law: "ISDS: The Wild, Wild West of International Practice", 44 Brooklyn Journal of International Law (2018) 10.

 $^{^{17}\,\,}$ 'Peru-Chile Deal Will Strengthen Economies', World Affairs (4 February 2014).

Franck et al., 'Inside the Arbitrator's Mind', 66 Emory Law Journal (2017) 1115 (providing evidence that international arbitrators are influenced by anchors when awarding damages). For a discussion on biases of international actors, see Hyde, 'Experiments in International Relations: Lab, Survey, and Field', 18 Annual Review of Political Science (2015) 403, at 406–407.

¹⁹ Ibid.

by the [State] authorities ... and shall not include [justification]'.²⁰ In other words, 'final offer' arbitrators can pick between two offers (fixed monetary amounts) rather than having the discretion to make a decision based on a reasoned understanding of the case. This approach assumes that, at a certain point, the claimed amount can be perceived as suspect or outrageous and, hence, counterproductive to the arguing party's case.²¹

The 'final offer' format attempts to change the calculus of the litigating parties: facing the possibility of being perceived as unreasonable and, hence, having the arbitrators side with the party perceived as more reasonable, the binary mode tempers the litigation parties' inclinations to demand an extreme amount. The main hope is that this system will moderate the parties' requests (and expectations) and help arbitrators elect a choice closer to the actual amount in dispute (or a settlement). To my knowledge, there is no actual evidence on how this setting may affect the parties' strategic behaviour, but the operation is grounded in game theory. Anecdotally, parties to arbitration are said to adjust their positions based on the expectation of the arbitrators' likely decision and other factors.²²

In a recent survey experiment conducted on arbitration professionals (mostly lawyers involved in ISDS), Anton Strezhnev and I tested a 'baseball' arbitration scenario (in an investor-state, not an interstate context). Arbitrators were asked to select the amount that a hypothetical respondent should pay a hypothetical claimant in damages as a consequence of a treaty violation. In the vignette, we constrained our participants to choose exclusively between the Claimant's proposal (which was fixed at US \$50 million) and the respondent's proposal (which varied and randomly assigned in three ways: US \$25 million (N = 85), US \$12.5 million (N = 70) and US \$6.25 million (N = 93)) (see Appendix 1).²³ We did not require justification. This setting allowed us to test if and at what point arbitrators may consider an 'offer' somehow to be an 'unrealistic' expectation of the value of the hypothetical case.

Based on 248 responses, we found that arbitrators indeed sided with the claimant's proposal (fixed) at a higher rate when the amount requested by the respondent was further from the middle point – between the claimant's proposal and zero. However, the results suggest that the gap between the value of the proposals does not matter too much, at least given the parameters used in our hypothetical. In other words, there was very little difference between the most extreme proposal and the least extreme one. Based on our design, we identified an effect in the probability

Multilateral Convention to Implement Tax Treaty Related Measures to Prevent Base Erosion and Profit Shifting, 7 June 2017, available at www.oecd.org/tax/treaties/beps-mli-signatories-and-parties.pdf.

McAuliffc and Bornstein, 'All Anchors Are Not Created Equal: The Effects of Per Diem versus Lump Sum Requests on Pain and Suffering Awards', 34 Law and Human Behavior (LHB) (2010) 164. Tijmes, 'Who Wants What? – Final Offer Arbitration in the World Trade Organization', 26 EJIL (2015) 587.

²² Marburger, 'Exchangeable Arbitrator Behavior: A Closer Look', 43 Economics Letters (1993) 219.

²³ See Appendix 1 for vignette and results. The experiment is similar in design to a 'conjoint' multi-attribute choice experiment, an approach popular in social sciences. See Hainmueller, Hopkins and Yamamoto, 'Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments', 22 Political Analysis (2013) 1.

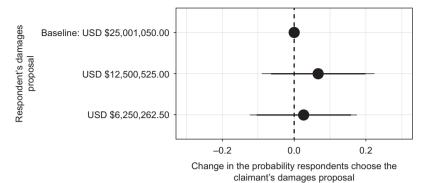


Figure 1: Representation of the Average Effect of an Arbitrator's Choice of Compensation in Baseball Arbitration

Note: The figure represents the change in probability of choosing the claimant's proposal given respondent's different scenarios. Lines denote 95% bootstrapped confidence intervals (N = 248).

of choice at a strong power (0.80 is typically the threshold used) in the order of six percentage points, although the external validity to tax arbitration professionals is unclear (see Figure 1).²⁴

B Hindsight and International Trade

Hindsight bias is the inclination, after an event has occurred, to see the event as having been predictable. ²⁵ This effect is particularly relevant in adjudicatory settings in which judges are required to assess decisions made in the context of risk or uncertainty. ²⁶ For instance, in negligence determinations, hindsight bias may result in juries' finding defendants liable more frequently than if cost benefit analysis were done indiscriminately. ²⁷ The *ex post* assessment of the facts results in the overestimation of the *ex ante* likelihood of the event; the consequence is that defendants are found liable more frequently than if the analysis were done without knowledge of the outcome. ²⁸

- In experiments, it is important to assess 'external validity' the extent to which the findings can be generalized to a population of interest. See Dunoff and Pollack, 'Experimenting with International Law', 28 EJIL (2017) 1317; see also Barabas and Jerit, 'Are Survey Experiments Externally Valid?', 104 American Political Science Review (2010) 226. On this particular experiment, see Puig and Strezhnev, 'The David Effect and ISDS', 28 EJIL (2017) 731.
- Hugh and Dekker, 'Hindsight Bias and Outcome Bias in the Social Construction of Medical Negligence: A Review', 16 Journal of Law and Medicine (2009) 846.
- ²⁶ Teichman and Eyal, supra note 9; Rachlinski, 'A Positive Psychological Theory of Judging in Hindsight', 65 University of Chicago Law Review (1998) 571.
- ²⁷ Kamin and Rachlinski, 'Ex-post ≠ Ex-ante: Determining Liability in Hindsight', 19 LHB (1995) 89; Hastie et al., 'Juror Judgments in Civil Cases: Hindsight Effects on Judgments of Liability for Punitive Damages', 23 LHB (1999) 597.
- Fischhoff, 'Hindsight=/Foresight: The Effect of Outcome Knowledge on Judgment under Uncertainty', 1 Journal of Experimental Psychology (1975) 288. In legal settings, hindsight bias can be exacerbated by complex mathematical concepts, statistical analysis, uncertain criteria and availability bias. See Tversky and Kahneman, 'Availability: A Heuristic for Judging Frequency and Probability', 5 Cognitive Psychology (1973) 207.

International adjudicators may review the assessment of questions of uncertainty, risk and scientific probity. One example, among many decisions of domestic bodies predicated in international law, is the assessment of 'non-obviousness', the requirement that prevents the patenting of trivial inventions, as I now explain.

1 Decision-Making under Uncertainty

The TRIPS Agreement regulates the conditions under which the WTO members should recognize patents. Patents confer the exclusive right to make, use or sell an invention, generally for at least 20 years. On the basis of international patent law, an invention needs to meet the domestic law requirements of novelty, industrial applicability (usefulness) and inventive step (non-obviousness).²⁹ Concerned with the effect of such practice on access to medicines, a growing list of countries have started to constrain evergreening, which is refiling patents with minor changes to obtain longer periods of monopolistic protection. Among other grounds, domestic authorities have determined that evergreening is impermissible because it does not satisfy the inventive step requirement.³⁰

One problem in this context is that hindsight bias can affect the assessment, outside factual grounds, of that determination. In hindsight, a small modification may be perceived as trivial; an obvious advancement over what was already known. Theoretically, a patent denial or invalidation on such basis may be challenged by a member of the WTO 'as applied'. More likely, however, a general practice could be challenged 'as such' if a national authority's policy discriminates between different fields of technology. The question presented could be whether the standard applied by a national authority to deny or invalidate a patent for advancing an obvious extension of the invention is inconsistent with the TRIPS Agreement. The assessment of non-obviousness in an unbiased way may require that experts suppress hindsight bias. To suppress hindsight bias, experts could rely on debiasing tools.

2 Format of Advice Giving: Safe Harbour Provisions

Studies show that the order and way in which information is framed can influence how people think about probability. In fact, the framing effect is also considered an

²⁹ TRIPS Agreement, *supra* note 6, Art. 27 (and footnote).

For discussion, see Du, 'Novartis Ag v. Union of India: "Evergreening", Trips, and "Enhanced Efficacy" under Section 3(d)', 21 Journal of Intellectual Property Law (2014) 223. Some countries have also required patent applicants to disclose a sound prediction of utility (or 'usefulness') of a patent at the time of application (and if a particular utility claim fails to deliver on it, then the patent can be invalidated). See UNCITRAL, Eli Lilly and Company v. Government of Canada, 16 March 2017, ICSID Case no. UNCT/14/2.

Mandel, 'Patently Non-Obvious: Empirical Demonstration That the Hindsight Bias Renders Patent Decisions Irrational', 67 Ohio State Law Journal (2006) 1391; Seymore, 'Foresight Bias in Patent Law', 90 Notre Dame Law Review (2015) 1105.

For 'as applied', see WTO, EU – Seizure of Generic Drugs in Transit – Requests of Consultations, 18 June 2010, WT/DS408/8. For 'as such', see WTO, Canada – Patent Protection of Pharmaceutical Products, 17 March 2000, WT/DS114/R.

implicit bias – most notably, when information is presented as a loss or as a gain.³³ Empirical evidence suggests that the form or order in which information is presented to experts may affect the advice they give.

In the patent example at issue, an innovation may not seem obvious if the assessing expert is unaware of key steps adopted or if the actual innovation is not fully disclosed prior to making an initial determination. Hence, to suppress hindsight bias that affects factual analysis, domestic courts have required that experts assess if a patent is advancing an obvious extension of the prior art, assuming the lack of knowledge of the actual innovation. Other authorities have gone further, demanding that, prior to making such assessment, experts list all inventions considered 'obvious'. The rationale is that, if it is in fact obvious, a reasonable person should be able to come up with that inventive step herself.

Recent experiments confirm that rules on framing and advice giving may temper hindsight bias in this context. When experts were presented with a problem faced by inventors and asked what sorts of obvious solutions they foresaw before learning of the actual solution, the experts were less likely to name a solution that was considered obvious in hindsight. A patent holder/seeker still can claim the innovation, but the expert does not learn of the specific innovation until after listing obvious solutions to the problem in question.

In the WTO, a safe harbour provision could encourage a determination by local authorities relying on properly implemented framing and advice-giving rules for experts. In other words, WTO members could agree that a patent denial or termination for lack of satisfaction of non-obviousness has a rebuttable presumption of validity when they follow a particular process – similar to other areas of WTO law where scientific probity is relevant. ³⁶ Such rules may also mitigate hindsight bias in other areas where domestic courts or administrative agencies assess *ex post* the *ex ante* likelihood of an event. For instance, in 'safeguard' determinations, the investigating authority has to assess if, as a result of the effect of the obligations incurred and given a development unforeseen at the time when it incurred its obligation, there has been an increase of imports that causes serious injury to a domestic industry. Such an *ex post* assessment of foreseeability, like others in the international trade context, may be also affected by hindsight bias.

C Affiliation and International Investment

Affiliation bias is the tendency to form allegiances, to the detriment of impartiality, and to favour a position based on an expert's relationship with a particular group

³³ Kahneman et al., 'Experimental Tests of the Endowment Effect and the Coase Theorem', 98 Journal of Political Economy (1990) 1325.

³⁴ Pozzoli SPA v BDMO C.S., [2007] EWCA Civ 588.

³⁵ For discussion, see 'How Do You Protect Patents from Judicial and Expert Hindsight?', *The IPKat*, available at http://ipkitten.blogspot.in/2018/03/how-do-you-protect-patents-from.html?m=1.

Agreement on Sanitary and Phytosanitary Measures 1994, 1867 UNTS 493, Art. 3.2; EC – Measures Concerning Meat and Meat Products (Hormones) – Report of the Panel, 18 August 1997, WT/DS26/R, para. 4.87.

or interest.³⁷ There are several instances in which affiliation bias may impact international economic adjudication. For one, parties to proceedings often rely on experts whose impartiality may be affected by the relationship with a party. The clearest example is perhaps the appointment of adjudicators, as I briefly explain in the context of ISDS.

1 Nomination of Experts and Adjudicators

The practice of each party unilaterally appointing one decision-maker is widespread in international arbitration. 'Party appointments' may bias decision-making in multiple ways but especially through the introduction of two effects. First, a reasonable litigant will select an arbitrator who has reliably shown a favourable approach towards the set of issues relevant to the litigant's goals. Separately from such selection effect, arbitrators may also find it difficult to maintain impartiality, making it difficult to neutralize preferences through introspective reflection. In fact, an arbitrator may even resist the influence of, and yet still be affected by, this appointer or affiliation effect. Combined, selection and affiliation may exacerbate the polarization among the perspectives of decision-makers, especially in areas that routinely confront values, interests and political ideas in legal disputes like ISDS. ³⁸

Affiliation bias has been documented in controlled experiments, including amongst international arbitrators.³⁹ Empirical evidence also confirms that the effect is stronger when adjudicators have more formal discretion (for example, the question of arbitration costs within the International Centre for Settlement of Investment Disputes [ICSID]). While constraining discretion may moderate the intensity of the effect, evidence confirms that bias remains when arbitrators decide questions of a less discretionary nature (for example, the question of damages under ICSID arbitration).

2 Non-Disclosure of Information: Blinding Mechanisms

Significant evidence confirms that knowing the course of the appointment creates unconscious pressures on arbitrators.⁴⁰ This also suggests that a 'blinding' mechanism could help by relieving the arbitrators from the information that triggers the effect.⁴¹ In fact, in the experiment mentioned above, Strezhnev and I found that, other things being equal, 'blinded' arbitrators were less likely to show the same affiliation bias and more likely to behave much like the arbitrators appointed by agreement.

³⁷ Harris, 'Testimony for Sale: The Law and Ethics of Snitches and Experts', 28 Pepperdine Law Review (2000)
1, at 3; Sperling, 'Expert Evidence: The Problem of Bias and Other Things', 4 Judicial Review (1999) 429.

³⁸ For a discussion, see Puig, 'Blinding International Justice', 56 Virginia Journal of International Law (2017) 647 (proposing a typology to understand bias in investor-state dispute settlement (ISDS) and explaining the complications to implement blinding as well as the limitations of this debiasing tool).

³⁹ Puig and Strezhnev, 'Affiliation Bias in Arbitration: An Experimental Approach', 46 JLS (2017) 371.

⁴⁰ Rojas-Elgueta, 'Understanding Discovery in International Commercial Arbitration through Behavioral Law and Economics: Journey Inside the Minds of Parties and Arbitrators', 16 Harvard Negotiation Law Review (2011) 165.

⁴¹ Robertson, 'Blind Expertise', 85 New York University Law Review (2010) 174.

Hence, blinding arbitrators by preventing them from knowing the party that appointed them to a tribunal would be a very sensible reform to reduce affiliation effects in international economic adjudication. While other reforms could work better for ISDS (think, for instance, of a random appointment of arbitrators) and blinding could be more effective in symmetrical arbitration settings (think, for instance, of commercial arbitration where consenting parties may end up as claimants or respondents), blinding is easy to implement. Moreover, the practice maintains the legitimacy and litigant-centric elements of party appointment systems and may correct the bias affiliation bias in arbitration, as I have explained at length elsewhere.

3 Information and Debiasing as Coordination Tools

As explained, various mechanisms have the potential to limit anchoring, hindsight or affiliation, among other biases, prominent in different areas of international economic adjudication. Before expanding on why this field is especially ripe for experimentation with such tools, I first propose a framework for thinking of debiasing mechanisms, which can be organized depending on the targeted actor (decision-makers or litigants) and the type of rules through which debiasing is implemented (procedural or substantive). Second, as some implicit biases are inevitable, biases should be accepted as a constant in adjudicatory decision-making, leveraging some of these biases when possible and increasing awareness about the problems in adjudication while noting the potential benefits of debiasing interventions for ICs. These tasks are facilitated by conceptualizing debiasing as a regulatory function of legal institutions, which involves shaping expectations to prevent, deter, eliminate or create checks against intuitive assessments by adjudicators or their strategic exploitation by litigants.

A 'Debiasing': A Conceptual Framework

Adjudicatory decision-making can be improved with mechanisms that directly suppress (or create immunity against) implicit biases. An obvious way is to target the decision-maker, who is often behaviourally responsible, via procedure. For instance, in the case of affiliation, blinding relieves the arbitrator of the allegiance in favour of the appointer. The incorporation of a simple process to limit the disclosure of relevant information insulates the decision-maker from the effect.

Yet, in some other instances, where a particular procedural mechanism is impractical or unavailable, a more fundamental change may be required to alleviate or at least check or control the effects of the bias in decision-making. One option is directly targeting the same decision-maker but with a substantive rule instead. According to Christine Jolls and Cass Sunstein, debiasing through substantive law (as opposed to procedure) is a distinctive and sometimes far preferable alternative to insulating legal outcomes from the effects of bounded rationality. Debiasing by adapting rules is less invasive and often cheaper.

One example of this second alternative is substantive rules (or precedent, binding interpretative guidance and so on) that constrain discretion. To be sure, there are

many positive aspects associated with judicial discretion (for example, distinguishing or narrowing situations; emphasizing a different perspective or applying other values), and discretion can be restrained through procedural means as well (for example, a functioning appellate system). But when discretion can aggravate the bias and procedural fixes are unavailable or costly, substantive reform may be the only solution. For instance, in the same affiliation bias example provided above, the observed effect is intensified when discretion is greater. Hence, one could imagine an alternative solution to blinding that shares similar goals. For one, a clear cost-shifting rule that constrains discretion in the allocation of costs could mitigate the affiliation effect (when it comes to costs decisions). While the mechanism still targets the adjudicator, it does so with a substantive provision that mandates how to evaluate information. To some extent, final offer arbitration adopts this approach by constraining the interpretative authority and the decision options of the tribunal.

Not all biases are susceptible to similar approaches either because it is practically impossible to change a rule or a process or because of the powerful influence that bias exerts on adjudicators. For example, in anchoring, most inoculants can facilitate alternative biases. Hindsight bias, as I have explained, is practically irreversible and is exacerbated by unclear criteria (for example, fair and equitable standard of treatment) or complex information like statistical or financial calculations.

Yet, even in such contexts, the effects of implicit biases could be moderated using existing tools that target bias indirectly. These circumvention strategies can adapt the legal environment that fosters intuitive (*mis*)judgment by using behavioural insights that target the incentives facing litigants. Hence, when targeting the adjudicator's intuition is practically impossible, procedural rules or substantive standards designed to change the litigants' behaviour may advance the same regulatory goal of a less biased decision-making process. Final offer arbitration uses an unorthodox process instead of targeting the effect of the anchor on arbitrators with an inoculant. It aims at forcing reliance on better information by targeting litigants indirectly via procedure. Moreover, instead of biasing the decision with a compromise effect, that same effect is put into work in favour of the process as both parties are incentivized to be perceived as the most reasonable.

Finally, debiasing adjudication from hindsight bias *ex post* is often impossible as one cannot travel back in time. Yet substantive rules can encourage litigants not only to produce particular information to substantiate a decision (an expert opinion in a patent case or science-based analysis in a case involving sanitary and phytosanitary measures) but also to require that the assessment is conducted based on information provided in a particular fashion – in this case, using the framing effect to yield a more accurate assessment.

The four different types of proposed approaches for debiasing are summarized in Table 1. In short, debiasing tools can play an important role in ICs. The core function of these tools is to act as a centrepiece of coordination rather than a mere inoculant of the minds of judges. Debiasing tools must address information that exacerbates bias and should be designed as choice preserving aids of decision-making, as I now explain.

	Adjudicators (Direct)	Litigants (Indirect)
Procedural	Non-disclosure of information	Production of information
	E.g., blinding appointments	E.g., baseball arbitration
Substantive	Rule constraining discretion to assess information	Rule on format of information and advice-giving E.g., non-disclosure of patented solutions
	E.g., cost-shifting rules	

Table 1: Typology of Debiasing Tools for International Economic Law

B Debiasing by Shaping Expectations

Institutions, including international law institutions, can be structured to presume the existence of implicit biases and, when appropriate, to correct for them. How does the topology of debiasing tools help to think more generally about institutional design? One key insight in Jolls and Sunstein's classic piece, *Debiasing through Law*, is that many strategies 'respond to information failures by providing additional facts. Indeed, many forms of debiasing through law may be seen as a distinctive kind of informational regulation'. ⁴² To a large extent, this is insightful and correct. In many instances, debiasing efforts work by generating additional facts. However, there are several instances where the problems generated by implicit biases may not be resolved with additional facts and may require strategies that work in different directions. In other words, sometimes more (biased) facts may be the problem.

For one, the topology and examples presented in the prior part of this article require strategies that are distinct to one demanding more facts. While they are all in the domain of informational regulation, in some instances the requirement can be to demand less information (blinding mechanism), while in others there is a mandate to discard information (format of information and advice giving). Debiasing strategies may also aim at generating accurate information about the parties' case (special procedures) or aim at constraining the paths to take with certain information (limiting discretion).

My main point is simple: debiasing adjudication is a strategy to generate more accurate decision-making by providing focal points (actions or legal concepts that pull attention) that shape expectations or help produce relevant information. This framing expands the number and type of plausible strategies for debiasing. In some instances, debiasing can be accomplished by insulating the effect and, in others, by circumventing it. One could view insulation and circumvention as interdependent (versus dominant) strategies that seek to produce the correct assessments that bounded rationality impedes: one by addressing the cognitive aspects of the decision-making; the other by generating the incentives to shield the process from those effects while preserving choice. In both instances, the expectations can be shaped behaviourally to prevent, deter, eliminate or create checks against implicit biases.⁴³

⁴² Jolls and Sunstein, *supra* note 7. The literature on debiasing is in its early days. For a review, see Lilienfeld *et al.*, 'Giving Debiasing Away: Can Psychological Research on Correcting Cognitive Errors Promote Human Welfare?', 4 *Perspectives on Psychological Science* (2009) 390.

⁴³ Jolls, Sunstein and Thaler, *supra* note 4.

4 Why Debiasing International Economic Law?

There is a backlash against international economic adjudication. At the WTO, the USA blocked to paralysis the appointment of members of the Appellate Body;44 Zimbabwe did the same with the judges of the Tribunal for the South African Development Community, and this court is now suspended; ⁴⁵ China refused to participate in the South China Sea arbitration; 46 and, since the Great Recession, some countries have denounced ISDS as being illegitimate. 47 The anxieties illustrated by these actions reflect both the successes and inherent tensions of international adjudication as much as they signal politically thorny times for international law. In part, states are reacting against ICs because their decisions matter; generally, they are binding and affect sovereign states and their citizens. Hence, decision-making affected by biases may cause confidence loss in international law and runs afoul of rule-of-law values. In this final part, I explain two additional reasons to debias adjudication as it can help increase support for international economic adjudication by: (i) encouraging agreement and negotiated solutions in dispute settlement and (ii) alleviating concerns over sovereignty. While debiasing alone will not solve the (growing) crisis affecting ICs, I also offer brief remarks about the plausible conditions that might make the use of debiasing tools more legitimate and successful as well as some limitations that may be unique to international economic adjudication.

A Encouraging Negotiated Solutions

One important purpose of international law is to help participants settle their disputes peacefully. While a pro-settlement leaning would be more controversial in certain areas of domestic law, ⁴⁸ states tend to prefer negotiated solutions in international economic law, some of which often happen without much transparency and in the shadow of, or after, legal disputes. Recognizing this, adjudicatory venues like the WTO profess a preference for transparent settlements.

Debiasing tools may encourage better information and, with that, perhaps also more negotiated solutions. Admittedly, under some conditions involving parties with unequal power, bargaining may favour the more powerful actor, which may motivate options other than adjudication. However, in many cases, settlement is more likely if the parties can estimate accurately the expected outcome of adjudication, making

- ⁴⁴ G. Shaffer, M. Elsig and M Pollack, 'Trump Is Fighting an Open War on Trade: His Stealth War on Trade May Be Even More Important', Washington Post (27 September 2017).
- ⁴⁵ Alter et al., 'Backlash against International Courts in West, East and Southern Africa: Causes and Consequences', 27 EJIL (2016) 293; M. Swart, 'A House of Justice for Africa: Resurrecting the SADC Tribunal', Africa in Focus (2 April 2018).
- ⁴⁶ PCA, South China Sea Arbitration (Philippines v. China), PCA Case no. 2013–19, Award, 12 July 2016.
- ⁴⁷ Kaushal, 'How the Past Matters for the Present Backlash against the Foreign Investment Regime', 50 HILJ (2009) 491; Van Aaken, 'Perils of Success? The Case of International Investment Protection', 9 European Business Organization Law Review (2008) 1.
- ⁴⁸ Fiss, 'Against Settlement', 93 Yale Law Journal (1984) 1073 (explaining how adjudication has value beyond the perspective of the adjudicating parties).

it easier to find mutually acceptable terms.⁴⁹ Debiasing tools may help in that regard by generating better information, helping adjudicators increase 'navigability' and, by doing so, find common ground prior or after a decision has been issued. At the very least, better information might actually result in less inaccurate decisions. By eliminating anchors based on irrelevant information or by relieving adjudicators of information that is unrelated to the analysis, debiasing tools can encourage more settlements.

Second, debiasing tools can soften entrenched positions and, as a result, lead to more agreement between parties or adjudicators (that is, less polarization and dissent). Choice preserving, behaviourally informed mechanisms may help in instances where adjudicators face a 'self-control' problem (for example, affiliation) or where litigants can draw an advantage from the lack of self-control (for example, anchoring). In such instances, debiasing tools may exert a positive force to move beyond narrower interests or opportunistic advantage, which in turn could lead to greater cooperation. For instance, blinding can help to reduce affiliation effects and, hence, dis-entrench arbitrators' views. At the very least, by educating lawyers to fully anticipate biases, these heuristics may be less likely to create obstacles for cooperation.

Finally, debiasing tools can help to shape agreeable outcomes (preserving some choice) and, by doing so, decrease the cost of agreement or increase the cost of disagreement. Behaviourally informed procedures can help to align incentives of disputing parties or they can encourage litigation parties to take more predicable paths. Just as opt-out clauses in treaties can impose costs on governments deciding to safeguard an exception, debiasing tools can help adjudicators and litigants avoid costly issues—for instance, by discouraging adjudicatory decision-makers from making broad principled decisions as opposed to resolving narrow issues of quantum. To be sure, there is a great deal of idealization in these examples, and reality may prove more complicated due to practical and political challenges unexplored here due to space limitations. The larger point, however, still holds: behaviourally informed rules and procedures can act as a centrepiece of coordination to shape or encourage consensus in adjudication.

B Alleviating the 'Sovereignty Loss'

As international adjudication is taking on increasingly more contentious topics that affect areas of policy that were previously the exclusive domain of national prerogative – from tax to trade and from investment to intellectual property – firms and

This proposition is well founded in social science analysis. Cooter and Rubinfeld affirm that '[t]he decision to assert a claim is a decision under uncertainty to be solved recursively by computing the expected values of subsequent stages in the dispute'. Cooter and Rubinfeld, 'Economic Analysis of Legal Disputes and Their Resolution', 27 Journal of Economic Literature (1989) 1067. To be sure, a legal system that produces bad information and, therefore, uncertainty might also encourage private bargaining. In that case, states may exit what would be an unpredictable system. In other situations, better information might actually push some states to let an international court or tribunal decide because they no longer fear inaccurate decisions. In such cases, decisions would be more legitimate as both parties may want to have the predicted outcome rather than a settlement. Ibid., at 1070. On state exit of unpredictable international economic systems, see Pauwelyn, 'The Transformation of World Trade', 104 Michigan Law Review (2005) 1.

governments have expanded the use of ICs. In the effort to legitimize these bodies and their procedures, scholars and policy-makers have suggested a wide array of reforms aimed at promoting transparency, increasing predictability and generating decisions more responsive to social welfare concerns.⁵⁰

The response has been in part the result of well-founded critiques. While the expanded scope and authority of international law, especially in economic affairs, has come with many important benefits – such as rising levels of investment, trade and welfare – it has intrinsically also created suspicions and distrust. Today, a growing sentiment of economic nationalism is impacting ICs as many actors see them as being implicated in what is usually referred to as 'sovereignty loss'. Yet, states and adjudicatory institutions have neglected the use of debiasing tools to ameliorate these concerns – albeit modestly. While debiasing will not resolve the current crisis alone, three reasons merit considering this path as a complementary strategy.

First, debiasing can improve the quality of decision-making – therefore, making adjudication less controversial. Debiasing can encourage decision-making with better information, which may result in decisions that are less unpalatable for states. Relatedly, debiasing can help by narrowing the scope of issues to be decided by adjudicators, hopefully allowing for the adjudication of less controversial issues. By constraining but maintaining choices (baseball arbitration), by creating a presumption of consistency through processes (safe harbour provisions) or by creating rules (as opposed to standards) in contexts where implicit biases are pervasive, adjudicators in ICs can be freed to focus more on narrow issues – a concern that, for example, relates to contemporary complaints against the expansive role of the WTO's Appellate Body.⁵¹

Finally, debiasing tools can dis-entrench positions by removing features that create judges' allegiances, such as polarization or loyalties, that may trigger impulses to act more as adjudicatory lawmakers instead of assuming a narrower role as 'dispute settlement' bodies. Affiliation bias in ISDS is the most obvious example – in such cases, blinding can mitigate the strength of selection effects (and, therefore, partisanship) by curtailing arbitrator opportunities to cater to the appointing party to develop a partisan reputation. But other areas of litigation prone to normative or ideological polarization such as international trade could also benefit from tools that target bounded rationality and the lack of self-control to support a particular party or position.

C A Debiased Field? Legitimacy, Possibilities and Limits

International economic law is a good candidate to experiment with debiasing tools. Not only are these fields of international law particularly 'judicialized', but their structure features some suitable conditions for experimentation. For instance, it is in these adjudicatory settings where states have the ability to appoint adjudicators to hear cases as well as where states and private actors have more control over legal processes

⁵⁰ See Alter et al., 'How Context Shapes the Authority of International Courts', 79 LCP (2016) 1.

Statement by the United States at the Meeting of the WTO Dispute Settlement Body, 23 May 2016, at 1, available at https://geneva.usmission.gov/wp-content/uploads/2016/06/Jun22.DSB_.pdf.

and disputes, relative to other areas of international law. It is in this field where anxieties over adjudication are openly manifested more often – perhaps because of the number, complexity and types of questions involving risk, uncertainty and resource allocation that have to be assessed against broad standards of conduct. And it is in these areas where dispute settlement is less formal, allowing for the experimentation with unorthodox procedural features.

There are questions of legitimacy around, as well as potentially unknown consequences of, implementing debiasing interventions. For one, heuristics have many legitimate uses in a world in which legal actors operate under time and information constraints. Bounded rationality can sometimes improve decision-making by allowing for quick, accurate assessments. More importantly, citizens worry about government manipulation even with widely accepted debiasing techniques. Finally, since we are still unable to determine how exactly implicit biases of an individual decision-maker affect outcomes of the collective body, it is also somehow speculative to conclude that we know the exact effects of biases on actual legal outcomes – hence, the call in this article for experimentation. ⁵²

These are very important questions that warrant proceeding with caution. It is naive, however, to draw sharp lines between acceptable and unacceptable informational regulation, especially as evidence builds around the role of biases and the potential of a large number of debiasing strategies. From a policy perspective, it is also important to ask why not adapt international legal institutions to the growing body of knowledge based on social science evidence. Many other practical challenges exist for the implementation of debiasing tools in each field – some of which I have addressed elsewhere. Fig. 14 if implemented successfully, debiasing interventions may not all be cost effective. There are many reasons why, but the main one is that other biases would emerge or persist. In many instances, litigants may become aware of the goal of the intervention and adapt to take advantage of these biases. Hence, future research should focus on identifying evidence-based, legitimate and cost-effective interventions that could improve the workings of ICs.

As a more general point, any solution to implicit biases in legal adjudication must account for context. For example, most areas of international legal adjudication take place in a close-knit community of repeat players who interact routinely and who respond to political considerations. Hence, while the thick social structure of the international law(yers) community may be valuable for experimentation with debiasing tools, it may also create additional hurdles by accentuating problems arising from repeat interactions or professional norms, just to mention some contextual factors.⁵⁴

Franck et al., supra note 18. Fiedler and Kutzner 'Reasoning Biases: Implications for Research in Judgment and Decision Making', 2 Wiley Blackwell Handbook of Judgment and Decision Making (2016) 380.

Puig, supra note 38.

Pauwelyn, 'The Rule of Law without the Rule of Lawyers? Why Investment Arbitrators Are from Mars, Trade Adjudicators from Venus', 109 AJIL (2015) 761.

5 Conclusions

A flourishing number of ICs decide questions of scientific probity, evaluate the conduct of government officials and calculate the present value of complex financial decisions of the past. The assessment of these questions in adjudicatory settings could be impacted by multiple implicit biases, such as affiliation, anchoring or hindsight. This article has described contexts in international economic adjudications where biases are pervasive as well as examples of debiasing tools that may help to overcome them. It has argued for experimentation with debiasing strategies to overcome implicit biases as well as their strategic exploitation by litigants. In a time of increasing anxieties about the role of ICs, the potential of debiasing to encourage settlement and to reduce concerns over sovereignty loss makes these pleas not only relevant but also timely. While questions over the legitimacy and the practical possibilities of a debiased law exist, institutional designers should embrace the application of behavioural approaches to avoid errors and improve – albeit modestly – the quality of international law.

Appendix 1: Methodology

Table A1: Summary of Number Observations Assigned to Each Condition Blinding and Anchoring Experiments (N = 248)

Variable		Conditions		
	Appointed by the respondent	Appointed by the claimant	Appointed by the parties	Blind appointment
Appointer effect	88 (90) US\$25,001,050.00	79 (79) US\$12,500,525.00	45 (45) US\$6,250,262.50	36 (38)
Respondent's proposed damages	85 (86)	70 (73)	93 (93)	

Notes: Counts denote the number of observations among survey respondents who answered the question on damages. Counts in parentheses denote the number of observations among respondents who answered the question on costs. Chi-squared tests for the marginal counts across all four conditions fail to reject the null that the counts are generated by the distributions we specified for randomization (p > 0.10).

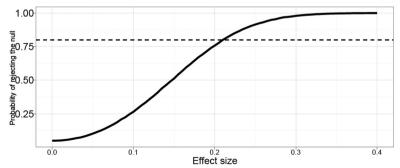


Figure A1: Probability Distribution of Compensation Proposal in 'Baseball' Format (Showing How Large an Effect Could Be Expected Given the Sample Size of the Experiment).

Imagine an investor-state dispute being conducted under the 2006 Arbitration Rules of the International Centre for Settlement of Investment Disputes (ICSID). The Claimant investor alleged that the Respondent state violated the provisions of a bilateral investment treaty to which the Respondent is a party. Among other arguments, the Claimant argued that the Respondent violated the treaty's fair and equitable treatment provisions and mistreated the Claimant's investment. You were appointed to the Tribunal [by the Respondent.]/[by the Claimant.]/[by the Parties.]/[.] After careful consideration of the facts of the case, the tribunal (you and your fellow arbitrators) unanimously decided that the Respondent unfairly treated the Claimant's investment in violation of the treaty and that the Claimant is entitled to compensation.

You are now asked to decide on the amount of damages owed to the Claimant by the Respondent. The parties have agreed that the tribunal's task is simply to pick one of the two positions of the parties' experts and decide how the expenses should be apportioned in this dispute. In its relevant part, the bilateral investment treaty provides as follows:

- 1. A Tribunal may award monetary damages and any applicable interest, only.
- 2. A Tribunal may also award costs in accordance with the applicable arbitration rules.
- 3. A Tribunal may not order a Party to pay punitive damages.

The Claimant has argued that they should be compensated for lost future profits that would have been realized had the measure not taken place, plus interest. The Claimant justifies this claim on the grounds that the enterprise operated profitably for a period of almost three years prior to the violation. The Claimant cites Metalclad v. Mexico ICSID Case No. ARB(AF)/97/1 which considered a minimum presence of at least two or three years necessary for an award of future profits. The Claimant's expert has calculated damages for US\$50,002,100.00 based on the discounted cash flow value of the expected returns from the Claimant firm's 10-year investment plan. The Respondent has argued that the enterprise had not operated for a sufficient period of time to establish itself as a "going concern" and that the ability of the enterprise to generate future earnings was uncertain and compromised. The Respondent cites Tecmed v. Mexico ICSID Case No. ARB(AF)/00/2, arguing that the tribunal in that dispute ruled that the Claimant's operating history of two and a half years was insufficient to establish enough objective data on profitability to apply a discounted cash flow analysis. Therefore, any estimate of future profits would be highly speculative. The Respondent instead proposes that damages should be based on the liquidation value of the firm and their expert has calculated damages for IUS\$25,001,050,00, US\$12,500,525,00, US\$6,250,262.50].

In ICSID disputes, the average award for Claimants who are awarded damages is about US\$45.6 million. The median award is US\$10.9 million (See: Franck, S. D., & Wylie, L. E. (2015). Predicting outcomes in investment treaty arbitration. Duke Law Journal, 65(3), 494-527.). Throughout the proceedings, both disputing parties were cooperative and the counsels for both parties behaved efficiently, professionally and ethically. The parties have not agreed on how and by whom the expenses shall be paid.

Figure A2: Sample Experimental Vignette Showing Key Manipulations