The European Journal of International Law Vol. 28 no. 3

The David Effect and ISDS

Sergio Puig* and Anton Strezhnev**

Abstract

The legitimacy of international investment law is fiercely contested. Chiefly, scholars argue that investor–state dispute settlement empowers corporations from rich nations over governments of poor ones. Some also assert that poor nations facing investment claims have limited ability to improve their standing in this setting of adjudication. Based on a first-of-its-kind experiment conducted on 257 international arbitrators, this article argues that one avenue to improve standing is for developing countries to exploit their 'underdog' status. We presented arbitrators with a vignette describing an investor–state dispute and randomly assigned different features to test their effect. Our results suggest arbitrators are prone to a particular type of bias – surveyed professionals were more likely to grant poor respondent states reimbursement of their legal costs compared to wealthy states when the respondent won the dispute. Based on this 'David effect', we argue for re-conceptualizing investor–state arbitration as a tool for partially mitigating power imbalances.

> David reached into his shepherd's bag, took out a stone, hurled it from his sling … The stone sank into Goliath's forehead.
>
> – 1 Samuel 17:49

1 Introduction

Investor–state dispute settlement (ISDS) – the controversial method of investment dispute resolution – is among international law's most debated fields.[1] Some scholars argue that ISDS empowers investors from wealthy states over governments of

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

** PhD Candidate, Department of Government, Harvard University, Cambridge, MA, USA. Email: astrezhnev@fas.harvard.edu. The authors would like to thank Christopher T. Robertson, Susan D. Franck, Josh Kertzer and Beth Simmons for helpful comments or suggestions.

[1] Kaushal, 'How the Past Matters for the Present Backlash against the Foreign Investment Regime', 50 *Harvard International Law Journal* (*HILJ*) (2009) 491 (explaining the mounting critique against the regime and its internal contradictions); Kingsbury and Schill, 'Public Law Concepts to Balance Investors' Rights

developing host nations. In response, others assert that without it investors would suffer opportunistic actions like expropriations in countries with weak rule of law and institutions.

This debate about ISDS' role and purpose within a global governance system can be framed as a tale of two types of underdogs: relatively weak governments fighting corporate power or defenceless private actors fighting arbitrariness. Of course, which account is more accurate depends on one's perspective.[2] Consider two recent cases. When Russian authorities arrested Yukos Chairman Mikhail Khodorkovsky in 2003, he was forced to sell his company. Through tactics that suggest retaliation on the part of President Vladimir Putin for Khodorkovsky's political activism, he was slowly deprived of the company's assets (through charges of tax fraud, among others). Today, former Yukos assets are an important component of Rosneft, one of the largest publicly traded oil producers. Claiming that Moscow expropriated Yukos' assets for political reasons, Yukos shareholders unsuccessfully attempted to recover damages from Russia for many years in various domestic courts around the world. In 2014, however, the company found support from three investment arbitrators who rendered an award of more than US $50 billion in favour of the shareholders of the defunct company – the largest amount ever granted by an international tribunal.[3]

Contrast this case with that of tobacco giant Phillip Morris and its recent challenge to Uruguayan efforts to limit the marketing of tobacco, alleging that the measures constituted violations of, among others, its intellectual property rights.[4] The positions taken by Phillip Morris, a company that proudly admits its litigious corporate culture, could be seen as a constraint on the ability of states to regulate marketing and set health and safety standards, which are well-established areas of sovereign prerogative. And, yet, three investment arbitrators decided against the corporate tobacco giant and in favour of the regulating host state, providing a decisive victory and comfort to other

---

with State Regulatory Actions in the Public Interest: The Concept of Proportionality', in S. Schill (ed.), *International Investment Law and Comparative Public Law* (2010) 75, at 75 (for the authors, investment arbitration promotes 'democratic accountability and participation … good and orderly state administration and the protection of rights and other deserving interests'); see also Gordon and Pohl, 'Investment Treaties over Time: Treaty Practice and Interpretation in a Changing World', OECD Working Papers on International Investment 2015/02 (2015).

2 See, e.g., Reed and Martinez, 'The Convenient Myth of David and Goliath in Treaty Arbitration', 3 *World Arbitration and Mediation Review* (2009) 443.

3 S. Reed, '$50 Billion Awarded in Breakup of Yukos', *New York Times* (28 July 2014). Recently, the arbitral decision was overturned by the Hague Commercial Court. See Hague Commercial Court, *Russian Federation v. Veteran Petroleum Limited, Yukos Universal Limited and Hulley Enterprises Limited*, 20 April 2016, Case no. C/09/477160/HA ZA 15-1, 15-2 and 15112.

4 UNCITRAL, *Philip Morris Asia Limited v. Commonwealth of Australia*, Notice of Arbitration, 12 November 2012, PCA Case no. 2012-12, paras 7.15–7.17. A case under the Switzerland–Uruguay bilateral investment treaty (BIT) was filed by Phillip Morris before the International Centre for Settlement of Investment Disputes (ICSID) in relation to legislation requiring all cigarette manufacturers to adopt a single presentation requirement. See ICSID, *Philip Morris Brands Sàrl, Philip Morris Products S.A. and Abal Hermanos S.A. v. Oriental Republic of Uruguay*, ICSID case no. ARB/10/7.

nations who wish to require cigarette packages that warn consumers of the dangers of smoking.[5]

The stylized versions of these recent, complex legal battles help to explain how ISDS is often perceived in radically different ways. Each version assumes a completely different function of ISDS as part of our global economic system – one that may either attenuate or exacerbate material inequities between already unequal disputing parties. However, this debate hinges on a fundamental empirical question that has puzzled international law scholars for some years: is ISDS a pro-claimant or pro-defendant litigation setting? A more concrete answer could better inform the debates around a contested field and the reform direction to improve the workings of international investment law. Much of the debate surrounding these questions has focused on the legal and structural elements of ISDS – for example, the extent to which commitments embodied in bilateral investment treaties (BITs) unduly constrain states' authority to regulate and enact social policy. But whether ISDS rectifies or exacerbates power imbalances is not only a matter of treaty text, it is also a matter of practice. The behaviour and attitudes of the arbitrators deciding these cases play an important role in shaping the overall trajectory of ISDS, particularly when arbitral jurisprudence remains highly unsettled.[6]

In this article, we explore one important element of this broader question: how do investor–state arbitrators themselves incorporate differences in resource endowments between the parties into their decisions? Independently of the underlying legal issue, what extra-legal heuristics do arbitrators use in the adjudication process when they have discretion? To better understand arbitrators' decision-making processes, we carried out a survey experiment on a population of international arbitration professionals. In the experiment, we presented participants with a hypothetical choice task related to an investor–state arbitration proceeding. Participants were randomly assigned different characteristics of the respondent host state as well as the characteristics of the state of origin of the claimant investor. Based on 257 responses from arbitrators around the world, we found that arbitrators gave more favourable judgments to claimants from middle-income states compared to those from high-income states. Likewise, low-income respondent states received additional compensation compared to middle-income respondent states.

Our core findings show that, even in the absence of formal guidance, investment arbitrators pay attention to the capabilities and potential resource constraints of the parties and behave in a way that is consistent with a preference for rectifying inequalities in litigation resources. However, this preference is not simply a blank check to uniformly favour the weaker party. We found that the positive effect of claimant/respondent weakness on compensation is largely conditional on that party also winning the dispute at hand, a result we term the 'David effect'. This pattern, we argue, is consistent with a combination of two types of cognitive predispositions that have been demonstrated in many other contexts – an aversion to unequal resource distributions

[5] UNCITRAL, *Philip Morris Asia v. Commonwealth of Australia*, Award on Jurisdiction and Admissibility, 17 December 2015, PCA Case no. 2012-12.

[6] Franck, 'The Legitimacy Crisis in Investment Treaty Arbitration: Privatizing Public International Law through Inconsistent Decisions', 73 *Fordham Law Review* (2005) 1521.

and a preference to allocate resources only to 'deserving' individuals. Overall, our results suggest that, for many stakeholders, the legitimacy of legal regimes such as international investment law may require a minimum expectation of fairness between the parties with unequal resource endowment.

To be sure, our argument is not that investment arbitrators necessarily support all of the underdogs in complex legal disputes. Yet, the results of the experiment suggest that, because of a cognitive predisposition to help the party with fewer resources, or because of the contemporary contested standing of ISDS, or to ensure buy-in on the part of litigants (and secure 'customers' for this arbitration system), arbitrators tend to 'compensate' perceived economically weaker parties who are successful in a proceeding when exercising discretion. This pattern is highly relevant to debates regarding the future of ISDS, the design of international investment treaties and the legitimacy of this controversial legal field.

The article proceeds in four parts. The next section summarizes how the current legitimacy debate is broadly framed around two different perspectives of ISDS: the 'shadow of obsolescence' and the 'over-empowered investor' perspectives. The third part describes the experimental design and reports on the main results of our study. The fourth part discusses the most obvious challenges and limitations of our research approach and the implication of our findings.

2 Investment Arbitration: A Tale of Underdogs

A *Background*

Globalization and economic interdependence have led to an increase in cross-national business interactions and a corresponding rise in transnational litigation and arbitration.[7] Among the most controversial expansions is the extension of arbitration to resolve disputes involving actions by sovereign states.[8] Through a vast network of BITs, many states have granted private foreign investors the right to pursue arbitration against host states before an ad hoc international tribunal. Under investor–state dispute settlement, foreign investors can sue states for alleged violations of treaty commitments, foreign investment promotion laws and, in some cases, investment contracts.[9] With limited exceptions (for example, restitution), tribunals can only obligate the state to pay damages to the investor, creating an inherent asymmetry to the

[7] Drahozal, 'Private Ordering and International Commercial Arbitration', 113 *Pennsylvania State Law Review* (2008) 1031. For a socio-legal explanation of arbitration, see Y. Dezalay and B.G. Garth, *Dealing in Virtue: International Commercial Arbitration and the Construction of a Transnational Legal Order* (1996).

[8] Broches, 'Development of International Law by the International Bank for Reconstruction and Development', 59 *American Society of International Law Proceedings* (1965) 33, at 34–38, reprinted in A. Broches, *Aron Broches, Selected Essays: World Bank, ICSID, and Other Subjects of Public and Private International Law* (1995) 79, at 80–84; for a criticism, see G. Van Harten *et al.*, *Public Statement on the International Investment Regime*, 31 August 2010, available at www.osgoode.yorku.ca/public-statement-international-investment-regime-31-august-2010/.

[9] A. Parra, *The History of ICSID* (2012).

arbitration process. Corporations have used investor–state arbitration to challenge not only cases of direct expropriation, nationalization or confiscation of property but also domestic laws and regulatory decisions seen as being unfavourable to their business activities.[10]

The legitimacy of ISDS is fiercely contested.[11] A central critique of this mechanism is that it acts as a tool to empower investors from developed states over governments of developing host states. Since arbitrators are drawn often from the field of international business law, they are often criticized for not being able to adequately adjudicate investors' challenges to public laws in a way that is consistent with a state's expectation of sovereignty in domestic policy. Concerns regarding the delegation of such crucial policymaking authority have led some politicians and scholars to advocate for the removal of ISDS from future international commercial treaties.[12] Notably, the European Union (EU) has proposed an alternative permanent court system and the elimination of ISDS in recent treaty negotiations with Vietnam, Canada and the USA.[13]

Despite the debate, many arbitration professionals and some policymakers defend the role of ISDS – at least in a narrow context.[14] The nature of many foreign direct investments makes the credible enforcement of well-defined property or contractual rights essential. Indeed, one of the seminal insights into the challenges of foreigners investing abroad is Raymond Vernon's idea that capital-intensive investments are under a constant shadow of obsolescence.[15] Hardware bolted to the ground is ripe for expropriation and one of the central functions of legal protection backed by independent adjudication is to temper the host country's incentive to act opportunistically and expropriate and harm, or renege on, a contract after a costly investment has been made.[16] As we now explain, the arguments for or against ISDS often use a common set of legal standards, principles and procedures and construct them as two different tales

[10] E.g., TransCanada filed notice over a decision to halt construction of the Keystone XL pipeline on environmental grounds. The case was recently suspended. See I. Austen, 'TransCanada Seeks $15 Billion from U.S. over Keystone XL Pipeline', *New York Times* (6 January 2016).

[11] Yackee, '*Pacta Sunt Servanda* and State Promises to Foreign Investors before Bilateral Investment Treaties: Myth and Reality', 32 *Fordham International Law Journal* (2008–2009) 1550, at 1611; see generally M. Waibel *et al.* (eds), *The Backlash against Investment Arbitration* (2010); see also Kahale, 'Is Investor–State Arbitration Broken?', 7 *Transnational Dispute Management Journal* (2012).

[12] See, e.g., United States Trade Representative, Letter from Senator Warren to Ambassador Michael Froman, 17 December 2014 (criticizing investor–state dispute settlement (ISDS) as a mechanism to challenge US government policies before a panel of private attorneys that sits outside of any domestic legal system).

[13] *EU Finalises Proposal for Investment Protection and Court System for TTIP*, Commission Press Release IP/15/6059, 12 November 2015, available at http://europa.eu/rapid/press-release_IP-15-6059_en.htm.

[14] Paulsson, 'Arbitration without Privity', 10 *ICSID Review: Foreign Investment Law Journal* (*IRFILJ*) (1995) 232; Roberts, 'Clash of Paradigms: Actors and Analogies Shaping the Investment Treaty System', 107 *American Journal of International Law* (*AJIL*) (2013) 45 (dissecting changes in the field by reference to a clash of paradigms of actors with different backgrounds).

[15] R. Vernon, *Sovereignty at Bay: The Multinational Spread of US Enterprises* (1971); Woodhouse, 'The Obsolescing Bargain Redux? Foreign Investment in the Electric Power Sector in Developing Countries', 38 *New York University Journal of International Law and Politics* (2006) 121.

[16] Reisman, 'International Investment Arbitration and ADR: Married but Best Living Apart', 24 *IRFILJ* (2009) 185, at 185–192.

of underdogs. How one understands the relative bargaining strengths of the parties will colour the extent to which investor–state arbitration is considered a legitimate form for resolving legal disputes.

B *The Shadow of Obsolescence: The Investor as the Underdog*

One frame for investor–state arbitration is the tale of the investor as an underdog relative to the host state. For some, the allocation of sovereignty to states by international law justifies access to a special dispute settlement process outside the purview of the host state to rectify the imbalance. In other words, conflicts involving foreign investors and states are distinctive as governments may have a direct stake in taking or undermining the assets of an investor. ISDS can therefore be seen as a way to ensure effective justice when domestic remedies are inadequate.[17] Under this perception, ISDS responds to concerns over access to procedural justice when a state abuses its power to disrupt productive investments in opportunistic ways.[18]

To that end, ISDS establishes a process based on formal rules and a methodology that ensures that a host state cannot block proceedings by refusing to cooperate with them.[19] In theory, ISDS serves as a neutral dispute settlement process to enforce basic obligations and to compensate investors in the event of a finding of expropriation, unfair treatment or discrimination based on nationality. The *ex post* deterrent effect of ISDS may result in states with difficult legal or political environments becoming more attractive for investors *ex ante*. Michael Reisman summarizes a version of this argument as follows:

> A common feature of foreign direct investment is that the investor has sunk substantial capital in the host [s]tate, and cannot withdraw it or simply suspend delivery and write off a small loss as might a trader in a long-term trading relationship. … So rather than having an equality of bargaining power in an exclusively negotiation-based regime, parity will cease and things will tilt heavily in favor of the respondent [s]tate. *Unless*, that is, both sides appreciate that if negotiations fail, compulsory arbitration will follow.[20]

Empirical research suggests that this logic is especially likely to apply to disputes in resource intensive sectors (a significant percentage of arbitrations – close to 40 per cent of the cases – involve these sectors).[21] Moreover, the historical context of the

[17] M. Borchard, *The Diplomatic Protection of Citizens Abroad* (1915), at 29: '[I]t is clear that by international law there is no legal duty incumbent upon the state to extend diplomatic protection. Whether such a duty exists towards the citizen is a matter of municipal law of his own country, the general rule being that even under municipal law the state is under no legal duty to extend diplomatic protection.'

[18] Puig, 'Emergence and Dynamism in International Organization: ICSID, Investor–State Arbitration and International Investment Law', 44 *Georgetown Journal of International Law* (2013) 531; on procedural fairness, see Lind *et al.*, 'Individual and Corporate Dispute Resolution: Using Procedural Fairness as a Decision Heuristic', 38 *Administrative Science Quarterly* (1993) 224, at 225.

[19] Convention for the Settlement of Investment Disputes between States and Nationals of Other States (ICSID Convention) 1966, 575 UNTS 159.

[20] Reisman, *supra* note 16, at 190–191 (emphasis in original).

[21] Kobrin, 'Testing the Bargaining Hypothesis in the Manufacturing Sector in Developing Countries', 41(4) *International Organization* (*IO*) (1987) 609. To be sure, research suggests that as the bargaining hypothesis has limited application, it cannot be treated as a universal phenomenon. Nevertheless, the bargaining hypothesis continues to be part of the discourse around the need for ISDS. Puig, *supra* note 18, at 564–566.

development of ISDS suggests that it was, in part, driven by fears of such bargaining imbalances. After World War II, decolonization brought calls for a system of protection that could limit the repossession of the property of former colonial powers and their nationals, including oil, gas, mineral and other concessions. Attempts at creating a multilateral treaty or other global instruments were met with what eventually came to be known as the New International Economic Order. This geopolitical movement defended, among other positions, the application of a limited view of customary international law and championed domestic courts as the locale for disputes involving foreign investors.[22]

Within the World Bank, industrialized countries endorsed the negotiation of the Convention on the Settlement of Investment Disputes between States and Nationals of Other States (ICSID Convention).[23] Instead of addressing the specific rights and obligations afforded to foreigners investing abroad, which was then almost as contested a matter as it is today, the Convention simply enabled a system of 'private' enforcement of commitments via investor–state arbitration.[24] The International Centre for Settlement of Investment Disputes (ICSID), formed by the ICSID Convention, initially promoted the inclusion of investment arbitration clauses in contracts and legislation, but the surge in BITs that occurred in the 1980s and 1990s cemented ISDS.[25] While arbitration clauses in contracts typically permit either party to seek arbitration in the case of a contractual breach by the other party, BITs grant what is often referred to as 'arbitration without privity'.[26] This asymmetry in investment arbitration – claims are brought mostly by firms, while the respondents are almost always states or state-owned entities – is one of the many reasons why ISDS is often seen through a different lens, one where the state, rather than the investor, is in the position of the underdog.

C *The Over-Empowered Investors: The State as the Underdog*

Backlash against ISDS has intensified dramatically over the last decade.[27] The challenges to the legitimacy of the mechanism are often fuelled by a perception that the system serves as 'a private, global super court that empowers corporations to bend

[22] GA Res. 3201 (S-VI), 1 May 1974; see also Weston, 'The Charter of Economic Rights and Duties of States and the Deprivation of Foreign-Owned Wealth', 75 *AJIL* (1981) 437, at 439. The New International Economic Order also argued for broad government discretion to regulate and expropriate foreign capital to serve national interests. See Schwebel, 'The Story of the UN's Declaration on Permanent Sovereignty over Natural Resources', *49 American Bar Association Journal* (1963) 463, at 464.

[23] ICSID Convention, *supra* note 19.

[24] . Lowenfeld, 'ICSID Convention: Origins and Transformation', 38 *Georgia Journal of International and Comparative Law* (2009) 47, at 51–52; Broches, *supra* note 8.

[25] Parra, *supra* note 9.

[26] Paulsson, *supra* note 14 at 232: '[O]ne where the claimant need not to have a contractual relationship with the defendant and where the tables could not be turned: the defendant could not have initiated the arbitration, nor it is certain of being able even to bring a counterclaim.'

[27] See generally Gordon and Pohl, *supra* note 1; Waibel *et al.*, *supra* note 11; Puig *supra* note 18; see also Van Harten and Loughlin, 'Investment Treaty Arbitration as a Species of Global Administrative Law', 17 *European Journal of International Law* (*EJIL*) (2006) 121, at 147–148; G. Van Harten, *Investment Treaty Arbitration and Public Law* (2007), at 17.

[poor] countries to their will'.[28] Specifically, as tribunals interpret investment protections to encompass a wide array of state actions, from general legislation, to executive orders, to final court decisions, the threat of arbitration serves to limit the ability of governments to engage in public policy.[29] Thus, ISDS provisions, instead of incentivizing economic development through greater foreign direct investment (FDI) flows, may actually hinder the promotion of economic and institutional development within states.[30] As democratically elected governments refrain from enacting policies to avoid becoming embroiled in an ISDS suit, the preferences of domestic constituencies are overlooked and states' regulatory space becomes diminished.[31]

Scholars have argued that investor–state arbitration amplifies pre-existing power differentials. As Gus Van Harten explains, the imbalance is rooted in the origins of the system and the 'fact that early international arbitrations of investment disputes sometimes followed in the wake of foreign invasion and occupation'.[32] Thomas Schultz and Cédric Dupont expand on this idea and suggest that 'by maximizing investor protection, investment arbitration maximizes the protection of the economic interests of 'neo-colonial' powers, that is, of countries with a high level of economic development'.[33]

This 'state as the underdog' narrative has risen to prominence in both journalistic and academic accounts of ISDS. After years of being of interest to just a few insiders, the first ISDS cases under the North American Free Trade Agreement's Chapter 11 contributed to a collective awakening regarding the potential use of investor–state arbitration.[34] Such cases ripened at a critical juncture, when the globalization awareness movement was gaining steam in the late 1990s. Contemporary press and scholarly works from the early years of this century draw attention to the role of ISDS and its potential to force poor nations to compensate rich Western companies.[35] Hence,

[28] C. Hamby, 'The Court That Rules the World', *Buzzfeed News* (28 August 2016), available at www.buzzfeed.com/chrishamby/super-court.

[29] M. Sornarajah, *The International Law on Foreign Investment* (2nd edn, 2004); Center for International Environmental Law, *Letter from Kate Horner et al., Friends of the Earth US, to Wesley Scholz, US Department of State*, 31 July 2009, available at www.ciel.org/Publications/BIT_Comments_Aug09.pdf.

[30] See generally K. Sauvant and L. Sachs (eds), *The Effect of Treaties on Foreign Direct Investment: Bilateral Investment Treaties, Double Taxation Treaties and Investment Flows* (2009); Kingsbury and Schill, *supra* note 1, at 75.

[31] For a recent book trying to identify and address some of the systemic concerns, such as limitations on domestic policy space, see Waibel *et al.*, *supra* note 11.

[32] Van Harten, *supra* note 27, at 17.

[33] Shultz and Dupont, 'Investment Arbitration: Promoting the Rule of Law or Over-Empowering Investors? A Quantitative Empirical Study', 25 *EJIL* (2014) 1147, at 1153.

[34] See J. Thomas, *The Battle in Seattle: The Story behind and beyond the WTO Demonstrations* (2000); see also Parra, 'Applicable Substantive Law in ICSID Arbitrations Initiated under Investment Treaties', 16 *IRFILJ* (2001) 20 (referring to concerns of 'entrusting [the application of treaty standards] to investor-to-state arbitration' by non-governmental organizations [NGOs]). North American Free Trade Agreement 1992, 32 ILM 289, 605 (1993).

[35] International Institute for Sustainable Development, *Private Rights, Public Problems: A Guide to NAFTA's Controversial Chapter on Investor Rights* (2001), available at www.iisd.org/pdf/trade_citizensguide.pdf.

non-governmental organizations (NGOs) and many scholars have argued against ISDS as an unfair and imbalanced system.[36] This frame is also often adopted in media coverage that emphasizes ISDS' undemocratic and highly clandestine nature.[37]

Governments have reacted to the perceived bias in arbitration. In recent years, states have either withdrawn from the ICSID system (for example, Bolivia, Ecuador, Venezuela), or threatened to leave (for example, El Salvador, Nicaragua, Argentina), because of their exposure to large international claims.[38] Moreover, problems such as the limited pool of arbitrators, the lack of diversity and gender parity in the arbitrators' pool, the revolving door between counsel and arbitrators and a general dissatisfaction with investor–state arbitration have generally given ISDS and international investment law the image of an illegitimate legal regime.[39]

D *Empirical Evidence and Its Limitations*

The two (somewhat stylized) perspectives on ISDS – a mechanism against opportunistic actions by capricious states or a mechanism that over-empowers rich investors – are almost binary opposites. Determining which model is the more accurate depiction requires testing their implications against empirical evidence. Much of the existing literature evaluating these two narratives has looked for evidence of institutional or structural biases by examining dispute outcomes. For example, Shultz and Dupont note that most investment arbitrations are filed by investors from developed states, consistent with the hypothesis that the main function of investor–state arbitration is indeed to allow developed states to exploit developing countries. In 88 per cent of cases, the investor is a national of a country that is ranked as a high-income state by the World Bank. In 9 per cent of cases, the investor is from an upper middle-income country. In only 3 per cent of cases is the investor from a middle-income, lower middle-income or low-income country.[40] Moreover, they find that host states with a higher development status stand a greater chance of successfully fending off claims from weaker parties with a lower development status. Thus, economic power disparities seem to be a very relevant factor of success.[41] Susan Franck, on the other hand, looks

[36] Alliance for Justice, *Letter to US Congressional Officials and US Trade Representative*, 11 March 2015, at 1, available at http://bit.ly/1GKLy5Q. Other critiques address review by national courts, consistency in outcome and diversity of adjudicators. Franck, *supra* note 6.

[37] A. Depalma, 'NAFTA's Powerful Little Secret: Obscure Tribunals Settle Disputes, But Go Too Far Critics Say', *New York Times* (11 March 2001), at 1; B. Moyers, 'Reports: Trading Democracy', television broadcast, *Public Broadcasting Service* (5 February 2002), available at www.gwu.edu/~nsarchiv/NSAEBB/NSAEBB65/transcript.html (reporting that Massachusetts senator and former presidential candidate John Kerry said in 2002, 'not a single word was uttered in discussing Chapter 11. Why? Because we didn't know how this provision would play out. No one really knew just how high the stakes would get.')

[38] Vincentelli, 'The Uncertain Future of ICSID in Latin America', *Social Sciences Research Network* (20 February 2009), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1348016.

[39] Puig, *supra* note 18, at 564–566.

[40] Shultz and Dupont, *supra* note 33, at 1154.

[41] *Ibid.*

at outcomes as a function of respondent state development and democracy levels. She concludes that, with some minor exceptions, there is no statistically significant relationship between economic development and relative investor success.[42]

These studies seek to evaluate the overall bias in the ISDS system in that they do not distinguish between legal/institutional factors that may advantage one side (such as broad versus narrow interpretations of treaty provisions) and behavioural factors (such as the preferences or implicit biases of the arbitrators). While this is a much broader question, there are several methodological issues that limit any inferences about bias in the system from observational data on investment disputes alone, a fact that most of the cited scholars admit. One of the authors of this article has referred to numerous such issues in prior work. In short, ISDS takes place in a fluid and contextual environment where investors can choose which cases to bring, which ones to settle, and which dispute settlement options to pursue. Parties file cases anticipating any underlying structural 'biases' in the system. These strategic dynamics exacerbate sample selection issues and stretch the validity of 'selection on observable' assumptions required for causal inference in typical observational studies. For example, the likelihood that a firm will choose to bring an investment dispute is probably associated with both the underlying quality of the case and the quality of legal representation – two variables that are difficult to observe and control for.[43] On this point, Schultz and Dupont argue that 'determining the rightful winner of a case requires a detailed knowledge of the particulars of the case, including its precise facts, which may differ from the facts as described by the tribunal in its decision'.[44] Hence, selection bias issues resulting from decisions to file and decisions to settle, as well as litigants' strategic choices, like forum shopping, make determining whether ISDS is a pro-claimant or pro-defendant litigation setting from observed outcomes nearly impossible without a model of the underlying selection process. To estimate such a model would require not only full transparency and a disclosure of outcomes by governments for all of the cases brought against them – a contentious issue in the field – but also more knowledge of the universe of potential disputes that firms could have filed.

Considering the empirical challenges of assessing bias in ISDS as a whole, we instead focus on a specific element of arbitration decision making: are arbitrators themselves inherently biased? Given the power of arbitrators to define broad legal standards, the ad hoc nature of tribunals and the lack of an appeals process in ISDS, the behaviour of the arbitrators themselves is of great interest to the overall question of the legitimacy of investment law. Furthermore, we take a different approach to inferring evidence of bias. Rather than relying on macro-level data on observed dispute outcomes, we

[42] Franck, 'Development and Outcomes of Investment Treaty Arbitration', 50 *HILJ* (2009) 435; For the response, see Van Harten, 'The Use of Quantitative Methods to Examine Possible Bias in Investment Arbitration', 3 *Yearbook on International Investment Law and Policy* (2011) 859; see also Gallagher and Shrestha, Investment Treaty Arbitration and Development: A Re-Appraisal, Global Development and Environment Institute, Working Paper 11-01 (2011).

[43] Puig, 'Blinding International Justice', 56 *Virginia Journal of International Law* (2017) 647.

[44] Shultz and Dupont, *supra* note 33, at 1154.

instead examine the micro-level decision-making processes of individual arbitrators. Specifically, we used an innovative experiment conducted on arbitration professionals to understand how an arbitrator's own perceptions of the resources and capacity of the litigants affect how they will decide aspects of a dispute. To be sure, this approach will not settle all of the debates regarding bias in ISDS as an adjudicatory setting. However, it adds a unique and heretofore unexplored source of complementary micro-level data to a debate driven almost exclusively by analyses at the level of the investment dispute.

3 Bias in ISDS: An Experimental Approach

Social scientists have long recognized that objectivity can be undermined by irrelevant information and other extraneous knowledge immaterial to a particular analysis. Legal scholars have long studied how heuristics or cognitive illusions may affect decision making by judges, experts and juries.[45] Less attention, however, has been given to this phenomenon among arbitral and international law decision makers. Most of the limited work in this area tends to be aimed at comparing whether juries or arbitrators make 'better' decisions.[46] Experimental research on arbitration is just emerging, and, to our knowledge, no experiments around ISDS currently exist.[47] Moreover, few studies explore perceptions of 'inequity aversion' or 'deservingness' heuristics in the legal context.[48]

In this section, we develop a theoretical argument for why arbitrators might use information on the underlying resource endowments of the parties to the dispute, information that is irrelevant to the facts of the case, as a salient heuristic in their decision making. We predict that this will manifest as a form of bias in favour of the weaker party in the dispute. We then discuss the design of our survey experiment created to test this hypothesis with a sample of arbitrators and show that, even when the substance of the dispute is held constant, arbitrators in the experiment tend to render more favourable decisions to economically weaker parties, both as claimants and respondents.

A *The 'Underdog' Effect as a Heuristic*

Tales about successful underdogs are common across cultures and throughout history. An underdog is typically one who is disadvantaged in relation to another who is well

[45] C.R. Sunstein, *Behavioral Law and Economics* (2000); see, e.g., Murrie *et al.*, 'Are Forensic Experts Biased by the Side That Retained Them?', 24 *Psychological Science* (2013); see also D. Ariely, *Predictably Irrational: The Hidden Forces That Shape Our Decisions* (2008); D. Kahneman, *Thinking, Fast and Slow* (2013); Guthrie, Rachlinski and Wistrich, 'Inside the Judicial Mind', 86 *Cornell Law Review* (*CLR*) (2001) 777, at 778. For a more general application of psychology insights to international law, see Broude, 'Behavioral International Law', 163 *University of Pennsylvania Law Review* (2014) 1099.

[46] Drahozal, 'A Behavioral Analysis of Private Judging', 67 *Law and Contemporary Problems* (2004) 105.

[47] Cf. Franck *et al.*, 'Inside the Arbitrator's Mind', 66 *Emory Law Journal* (2017) 1115; see also Puig and Strezhnev, 'Affiliation Bias in Arbitration: An Experimental Approach', *Journal of Legal Studies* (2017) (forthcoming).

[48] For a review of the 'underdog effect' more generally, see Vandello, Goldschmied and Richards, 'The Appeal of the Underdog', 33 *Personality and Social Psychology Bulletin* (2007) 1603.

endowed in some capacity. Underdogs compete with fewer privileges and resources or face more obstacles and challenges than opponents. Inevitably – almost as an instinctive reaction – many people feel a sympathetic desire for underdogs to win.[49] Stories of underdog victories are recurrent and are often recounted in almost mythological fashion.

Why might arbitrators be more likely to render favourable outcomes for parties perceived as underdogs? Simply put, it is because they are humans, susceptible to the same cognitive biases as any person faced with a complex decision-making task. We characterize the underdog effect as the product of a combination of heuristics or short cuts for information processing. One such heuristic is inequity aversion or individuals' implicit preference for equitable outcomes and the distribution of resources.[50] That is, individuals have an underlying, inherent disutility for inequity and are willing to forgo some benefit if they believe that doing so will lead to a fairer outcome. Within the field of behavioural economics, the existence of inequity aversion helps explain otherwise puzzling results of lab experiments where participants do not always behave in a purely self-interested manner as predicted by a game-theoretic model.[51] Moreover, experimental results within psychology suggest that this general preference for 'fair' outcomes is an inherent trait within humans, finding effects consistent with inequity aversion in children as young as eight years old and even in non-human primates.[52] While the exact nature of this preference for fairness remains an open question within the social sciences, the existing research suggests that it is very plausible that adjudicators will be motivated, in part, by concerns of distributional justice, even in the absence of a formal rule of equity. We therefore hypothesize that, all else equal, arbitrators will tend to award greater compensation to parties that are perceived to have comparatively fewer economic resources.

However, 'inequity aversion' is not the only relevant heuristic for the underdog effect. Another common decision-making shortcut documented within the psychology literature is what is termed the 'deservingness' heuristic. In short, when individuals decide on an allocation of resources to some third party, they will make their assessment in part on whether they believe the recipient is deserving of the benefit.[53] Researchers have found this heuristic to be particularly salient under

[49] Fehr and Schmidt, 'A Theory of Fairness, Competition, and Cooperation', 114 *Quarterly Journal of Economics* (1999) 817. For the economic implications, see Kahneman, Knetsch and Thaler, 'Fairness as a Constraint on Profit Seeking: Entitlements in the Market', 76 *American Economic Review* (1986) 728.

[50] Fehr and Schmidt, *supra* note 49, at 819.

[51] *Ibid.*

[52] Shaw, Alex and Olson, 'Children Discard a Resource to Avoid Inequity', 141 *Journal of Experimental Psychology: General* (2012) 382; Brosnan and de Waal, 'Monkeys Reject Unequal Pay', 425 *Nature* (2003) 297.

[53] M. Gilens, *Why Americans Hate Welfare: Race, Media and the Politics of Antipoverty Policy* (1999); Van Oorschot, 'Who Should Get What, and Why? On Deservingness Criteria and the Conditionality of Solidarity among the Public', 28 *Policy and Politics* (2000) 33; Feather, Volkmer and McKee, 'Attitudes towards High Achievers in Public Life: Attributions, Deservingness, Personality, and Affect', 43 *Australian Journal of Psychology* (1991) 85; Kay and Jost, 'Complementary Justice: Effects of "Poor but Happy" and "Poor but Honest" Stereotype Exemplars on System Justification and Implicit Activation of the Justice Motive', 85 *Journal of Personality and Social Psychology* (2003) 823; Kay, Jost and Young, 'Victim Derogation and Victim Enhancement as Alternate Routes to System Justification', 16 *Psychological Science* (2005) 240.

conditions of scarcity or the uneven allocation of resources – for example, in the context of public attitudes towards social policies like health care, unemployment benefits and welfare.[54] What constitutes deservingness is in part a function of the recipient's underlying resource endowments, with less endowed recipients being seen as more deserving of a scarce resource. However, decision makers also assess deservingness in terms of whether the recipient is at fault for their condition, with individuals being perceived as being more deserving when their lack of resources is perceived as due to bad luck rather than to individual failings. Low resource endowments by themselves do not engender support if the benefit is not seen as deserved. Therefore, while the logic of 'deservingness' predicts a similar outcome to the logic of 'inequity aversion' – arbitrators render decisions favourable to underdogs – it is caveated by the additional requirement of sympathy or compassion for the party being compensated.

Why might a party be perceived as deserving or undeserving in an arbitration context? One factor that likely affects perceptions of deservingness independent of resource endowments is the substantive outcome of the case. Our particular experiment considers how arbitrators choose to assign the legal costs of arbitration. The logic behind 'cost shifting' (rather than simply letting each party be responsible for its own legal fees) is that it serves to compensate for investments in legal capacity that would have been unnecessary in the absence of the dispute. The legal costs of winning parties are, in effect, undeserved penalties necessitated by the losing party's actions. Successful claimants would not have needed to initiate a costly dispute had the respondent state acted in a manner consistent with its legal obligations. Conversely, victorious respondent states would not have had to spend effort and resources defending themselves had the claimant chosen not to pursue an unmeritorious claim. In essence, the legal burden of the winner is the fault of a choice made by the loser. Since determinations of 'deservingness' are, in part, based on this assignment of responsibility – victims of fate are seen as more deserving than victims of choice – we hypothesize that the effect of a party's resource endowments on compensation is conditional on whether that party also wins the dispute.

Given the centrality of asymmetry in litigants' resources to debates over the legitimacy of ISDS in global governance, we aim to test whether arbitrators seek to partially rectify imbalances and render decisions to compensate *ex post* for a perceived *ex ante* difference in material resources between the claimant and the respondent. That is, does being perceived as an underdog in some ways help a litigant and under what conditions?

[54] Jensen and Petersen, 'The Deservingness Heuristic and the Politics of Health Care', 61 *American Journal of Political Science* (2016) 68; Petersen *et al.*, 'Deservingness versus Values in Public Opinion on Welfare: The Automaticity of the Deserving-ness Heuristic', 50 *European Journal of Political Research* (2011) 24; Petersen, 'Evolutionary Political Psychology: On the Origins and Structure of Heuristics and Biases in Politics', 36 *Political Psychology* (2015) 45; Skitka and Tetlock, 'Providing Public Assistance: Cognitive and Motivational Processes Underlying Liberal and Conservative Policy Preferences', 65 *Journal of Personality and Social Psychology* (1993) 1205.

B *Experiment*

1 Context

Before discussing the experimental design, we first describe the general structure of investor–state arbitration or ISDS. Almost all ISDS cases involve investors seeking compensatory damages when affected by a host state's 'excessive' intervention. Hence, states are almost always the respondent, defending governmental measures taken by authorities. Arbitration tribunals are typically composed of three members – one arbitrator appointed by each party in the litigation and a third arbitrator appointed by a method that attempts to ensure neutrality. All arbitrators are supposed to be independent and impartial, including the two party-appointed arbitrators, and to have no prior involvement in the dispute.[55] When accepting an appointment, arbitrators often sign a declaration either confirming independence and impartiality or disclosing circumstances that might compromise independent judgment.[56] However, to our knowledge, arbitrators do not receive training or information on how implicit biases may affect decision making. The applicable rules often provide that arbitrators cannot share the nationality of either party in the dispute.[57]

2 Design

The ideal natural experiment for assessing the role of perceived resources would be to compare two essentially interchangeable arbitrators deciding two substantively identical cases by parties with opposite profiles: a well-resourced investor versus an underdog state or an underdog investor versus a well-resourced state. Not surprisingly, such perfectly paired counterfactuals are nearly impossible to find among the limited set of observed disputes. We therefore turn to an experimental design that would allow us to explicitly control all extraneous factors to be independent of the treatment of interest – parties' resource endowments.

We designed a survey experiment that presented respondents with a brief vignette describing a hypothetical arbitration scenario containing randomly manipulated elements. The advantage of a survey experiment is that researchers can directly control the assignment of each exposure independent of the underlying attributes of the respondents. Randomized experiments break the association between the treatment

[55] M. Waibel and Y. Wu, *Are Arbitrators Political? Evidence from International Investment Arbitration* (2017), available at www-bcf.usc.edu/~yanhuiwu/arbitrator.pdf.

[56] See, e.g., ICSID Rules of Procedure for Arbitration Proceedings (ICSID Arbitration Rules) (2006), Rule 6.

[57] This is a traditional practice in arbitration. See International Chamber of Commerce Arbitration Rules (1998), Art. 9(5); London Court of International Arbitration Rules (2014), Art. 6(1); Cf. American Arbitration Association, Consumer Arbitration Rules (2014), Art. 6(4); Paulsson, 'Moral Hazard in International Dispute Resolution', 25 *IRFILJ* (2010) 339 (arguing that rules that preclude the appointing of an arbitrator who shares the nationality are 'a step in the right direction'). Previous studies in other international legal contexts suggest that international judges may tend to favour co-nationals. Posner and de Figueredo, 'Is the International Court of Justice Biased?', 34 *Journal of Legal Studies* (2005) 599; but cf. Voeten, 'The Impartiality of International Judges: Evidence from the European Court of Human Rights', 102 *American Political Science Review* 102 (2008) 417 (showing that similar geopolitical biases are not apparent in European Court of Human Rights rulings).

and all observed or unobserved factors that predict the outcome. Therefore, any outcome differences we observe between respondents in different treatment categories are a consequence of the treatment itself rather than any spurious differences in the respondents. Figure 1 provides a sample version of the vignette showing the manipulation at issue. Each participant could be told that the claimant in the arbitration was a firm headquartered in either a high- or middle-income economy, while the respondent country was either a middle- or low-income economy (see Figure 1).[58]

Additional elements of the vignette were also randomized across respondents taking the survey. Notably, participants could be told that they were appointed by the respondent, by the claimant, by the parties (that is, by agreement of the litigating parties) or simply that they were appointed to the tribunal without mention of any appointing method (as shown in the sample vignette below). To assess how treatment effects are moderated by the substantive decision of the tribunal, the vignette also manipulated the dispute outcome. For instance, the tribunal's ruling could take on one of four possible conditions reflecting different ways in which a case litigated under the applicable procedural rules in the vignette (ICSID 2006 Arbitration Rules) could be resolved: (i) the respondent expropriated the claimant's property (claimant wins); (ii) the respondent did not expropriate the claimant's property (respondent wins); (iii) the dispute is outside of the tribunal's jurisdiction (respondent wins on 'procedural' grounds) or (iv) the claimant's claims are manifestly without legal merit (case dismissed summarily; that is, respondent decisively wins).[59] Each of the conditions had equal probability of being assigned to each individual, and the treatments were each randomized independent of one another.

After being presented with the vignette, survey respondents were asked how they thought the parties' expenses in the dispute, including the cost of legal representation, should be apportioned. Survey respondents could choose to have one party reimburse the other party for either all or some of their expenses or have each party bear their own expenses. Participants were then asked to briefly discuss the reasoning behind their decision in an open-ended question. We chose to ask arbitrators to decide on the allocation of costs rather than on the actual merits of a case mainly for two reasons. First, summarizing the dispute such that participants would have enough information to render an educated decision on the merits of a full case would require an impossibly long vignette. We were conscious of the fact that participants would likely not have the time or interest to spend hours on our survey. Additionally, a long factual pattern could open the case to complex questions of interpretation. Instead, prompting survey respondents to render a clear decision on costs given the result of the case allowed us to use a vignette that participants could easily read and analyse in a practical amount of time.

Second, we needed an outcome that would exhibit meaningful variation in responses. To achieve that, we selected a question on which there is little pre-existing

[58] To avoid having too many treatments relative to the sample size, thus causing low power, we limit our treatments to two categories only for each disputant. Because ISDS claimants are rarely from low-income countries and respondent states are typically not high-income countries, we excluded low-income claimants and high-income respondents. This allows the vignette to capture more 'typical' investment arbitration scenarios.

[59] ICSID Arbitration Rules, *supra* note 56.

Imagine an investor–state dispute being conducted under the 2006 Arbitration Rules of the International Centre for Settlement of Investment Disputes (ICSID). The claimant is a firm headquartered in a [high-income/middle-income economy]. The respondent is a country classified by the World Bank as a [middle-income/low-income economy].

The claimant alleged that the respondent violated the provisions of a bilateral investment treaty to which the respondent is a party. Among other arguments, the claimant argued that the investor and its investments had been treated unfairly and that ultimately the respondent expropriated the claimant's investment located within the respondent's territory. The underlying dispute concerns an infrastructure project undertaken by the claimant under a concession contract with a governmental agency. The respondent argued in response that the claimant had violated provisions of the contract and that the investors received all compensation to which they were entitled.

You were appointed to the tribunal. After careful consideration of the facts of the case, the tribunal unanimously decided that the respondent unfairly treated and wrongfully expropriated the claimant's investment and that the claimant is entitled to compensation.

In their submissions on costs, both parties have requested that the other party bear the costs of the proceedings in full, including legal fees and expenses. The counsels for both parties behaved professionally and ethically during the proceedings.

Figure 1: *Sample Vignette of an Investor–State Dispute*

legal guidance (as a matter of legal rule or precedent), and arbitrators have a range of discretion among reasonable choices – otherwise, participants would likely all provide very similar, if not identical, answers.[60] Cost allocation was an ideal question in this

[60] Franck, 'Rationalizing Costs in Investment Treaty Arbitration', 88 *Washington University Law Review* (2011) 769; but see Gotanda, 'Awarding Costs and Attorneys' Fees in International Commercial Arbitration', 21 *Michigan Journal of International Law* (1999) 1 (discussing international arbitral rules generally);

respect because, in the context of the procedural rules applicable in our vignette, little formal or informal precedent exists. The arbitration rules grant discretion to tribunals to decide how the parties should pay the costs, and such discretion could plausibly be swayed by extra-legal considerations.

We conducted the survey in the autumn of 2015 and recruited participants by collecting publicly available email addresses of arbitrators from the lists of arbitral institutions.[61] In addition to the vignette responses, we surveyed the respondents with basic information regarding their professional activities and background. Nearly 16 per cent of surveyed professionals declared to have served as arbitrators in investor–state cases. All others confirmed having legal expertise in the field of international arbitration, and a large majority had previously served as arbitrators in commercial arbitrations. To increase our response rate, we sent several reminders to participants but promised only an advanced circulation of any articles summarizing the research. We obtained approval from the institutional review boards of both of the authors' academic institutions.

3 Analysis and Results

Prior to conducting the analysis, we pre-processed the 628 responses to our survey request. However, not all respondents that began the survey completed it. Only 257 respondents fully completed the survey and responded to the outcome question. To our knowledge, this is still the largest experimental survey conducted for this specialized population. The outcome observed for each participant is a five-category decision on costs: losing party pays some, losing party pays all, split costs (each party pays their own), winning party pays some, and winning party pays all. For the purposes of our study, we were interested only in the first three choices as these were the only outcomes of practical interest. In an arbitration context, having the winner of a dispute reimburse the loser would make little sense. Thirteen of our survey respondents, however, selected one of these two outcomes. It is most likely that these choices are a result of measurement error and respondents responding incorrectly on the survey. Since this outcome question involves post-treatment quantity, dropping observations where respondents chose to have the winner reimbursed may induce bias.[62] As a result, we retained these responses in the analysis (coding them as choosing none of the three outcomes of interest). Removing these observations from the analysis, however, does not substantively change any of the results.

Rubins, 'The Allocation of Costs and Attorney's Fees in Investor–State Arbitrations', 18 *IRFILJ* (2003) 109, at 109–110.

[61] These institutions include both investor–state arbitral institutions such as ICSID, along with international private arbitral institutions like the International Chamber of Commerce. In total, 28,832 recruitment emails were sent, though many of these were to addresses that were no longer active. Arbitrators were not paid but could opt in to receive updates regarding the results of the experiment.

[62] Montgomery, Nyhan and Torres, *How Conditioning on Post-Treatment Variables Can Ruin Your Experiment and What to Do About It*, Working Paper (2016), available at www.dartmouth.edu/~nyhan/post-treatment-bias.pdf.

Our primary independent variables of interest are the wealth level of the winning party and the wealth level of the losing party. Since we did not have enough observations to conduct entirely separate analyses of wealth level conditional on claimant/respondent victory with sufficient power, we pooled together claimant and respondent wealth when constructing the winning and losing party resource variables. The wealth of the winning party is coded as 'high' if the claimant wins the dispute and the claimant is from a high-income economy or the respondent wins the dispute and the respondent is a middle-income economy. Conversely, wealth is coded as low if the claimant wins and is from a middle-income economy or if the respondent wins and is a low-income economy.[63] Losing party wealth is coded analogously.

The measure of interest is the average effect of the party wealth treatment on the probability that a respondent would pick one of the three outcome choices of interest. Because each attribute of the treatment was randomized independently, we can estimate these effects without bias. Moreover, these effects can be estimated without any additional modelling assumptions using simply the difference in the share of respondents that picked each category under control and under treatment.[64] Figure 2 plots the estimated average effect of the winning party having a lower resource endowment, as proxied by the wealth of the respondent's or claimant's home country. Within our sample, arbitrators had a statistically significant higher chance of choosing to have the losing party compensate some or all of the winner's litigation costs when the winner was a middle-income claimant or a low-income respondent and were less likely to choose to split the costs ($p < 0.05$). On average, lower-resourced winners were about 12 per cent less likely to receive a decision that split the costs.

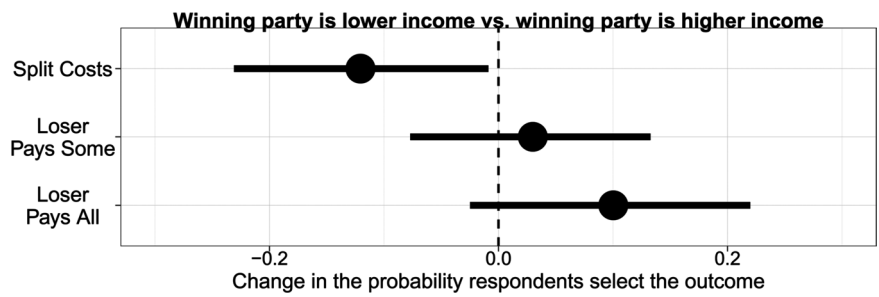

Figure 2: *Estimated Average Effects of Winning Party Wealth in an Investor–State Dispute*[65]

[63] When we interact our winner and loser wealth variables with whether the winner was the claimant or respondent, we do not find a statistically significant difference between the two effects on any of the three outcomes ($p > 0.05$).

[64] Our experiment is similar in design to a 'conjoint' multi-attribute choice experiment, an approach that has become increasingly popular within the social sciences. See Hainmueller, Hopkins and Yamamoto, 'Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments', 22 *Political Analysis* (2013) 1.

[65] Lines denote 95 per cent bootstrapped confidence intervals. Confidence intervals that do not cross zero imply that the difference between resource endowments is statistically distinguishable from 0 at ($p < 0.05$). Number of observations = 257.

Figure 3, conversely, plots the estimated average effects of the losing party having a lower resource endowment. The effect that we observe for winners' wealth does not appear at all. These general patterns are consistent with a story of arbitrators assigning costs and compensation based on litigants' perceived need or ability to pay – a factor that is, by design, orthogonal to the substantive aspects of the dispute. But the ability to pay only factors into arbitrators' decisions when the party also happens to win the dispute. These effects amount to quite a substantial shift in parties' welfare, given that the median legal costs for a claimant in ISDS are around US $3.1 million, while the median award is only US $10.5 million.[66] Legal costs alone can eat up over a quarter of the total amount under dispute.

We replicated this result in a parallel vignette involving an international commercial arbitration that we also presented to the same group of respondents (see Figure 4).[68] In this dispute, participants were given a scenario where the claimant firm won and was seeking damages from another firm. Both the claimant and respondent firms could be either a Fortune 500 company (high resource endowment) or a 'modest but growing local company in an emerging economy country' (lower resource endowment). Figure 4 shows the text of the vignette and the manipulation of interests. We also varied (i) the appointing party, (ii) whether the decision was unanimous, and (iii) the presentation of the submission on damages (either by two party-selected experts or an independent expert). Survey respondents were again asked to decide on the legal costs issue.

In total, 245 respondents also completed the investor–investor dispute vignette. Figure 5 plots the average effect of the claimant resource treatment. Among the respondents in the sample, claimants that were small local companies were estimated to be around 18 per cent more likely to receive an award of some or all of their costs

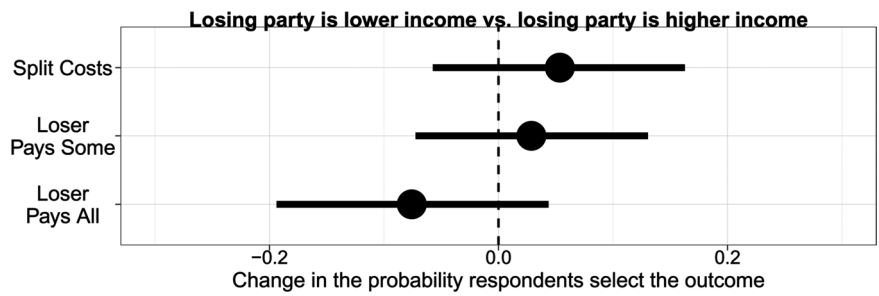

Figure 3: *Estimated Average Effects of Losing Party Wealth in an Investor–State Dispute*[67]

[66] Hodgson, 'Costs in Investment Treaty Arbitration: The Case for Reform', in J.E. Kalicki and A. Joubin-Bret (eds), *Reshaping the Investor-State Dispute Settlement System* (2015) 748.

[67] Lines denote 95 per cent bootstrapped confidence intervals. Confidence intervals that do not cross zero imply that the difference between resource endowments is statistically distinguishable from 0 at ($p < 0.05$). Number of observations = 257.

[68] The vignette presented to the respondent first was randomly chosen to minimize ordering effects. Some respondents completed the investor–state vignette first, while others completed the investor–investor vignette first. To avoid revealing to respondents that the appointing party was a treatment of interest to the researchers, the appointing party treatment was randomized across respondents but kept constant for each respondent between vignettes.

Imagine a commercial arbitration dispute being conducted under the arbitration rules of the International Chamber of Commerce, Paris.

The claimant is a [Fortune 500 company/modest but growing local company in an emerging economy country]. The respondent is a [Fortune 500 company/modest but growing local company in an emerging economy country].

The parties to the case agreed to bifurcate the proceedings in two phases, merits and damages. The primary dispute concerned whether the respondent complied with several contractual provisions over the course of several years. You were appointed to the tribunal. After careful consideration of the facts of the case, the tribunal decided unanimously that the respondent breached the contract and that the claimant is entitled to damages.

In their submissions on damages the parties agreed on an independent expert randomly selected from a pool of qualified experts who calculated damages for US $2,034,040. The counsels for both parties behaved professionally and ethically during the proceedings.

Figure 4: *Sample Vignette: International Commercial Arbitration*

compared to claimants that were Fortune 500 companies. Interestingly, it appears as though arbitrators make a two-stage decision when allocating costs, first choosing whether costs will be awarded at all and then deciding on the amount. In our investor–investor vignette, resource endowment seems to factor primarily into the first step. Arbitrators moved by the treatment to award costs instead of letting them fall where they may chose fairly evenly between awarding 'some' or 'all' of the costs. As before, Figure 6 shows no statistically significant difference between the outcome distributions for losing party wealth. We summarize the results of all four sets of comparisons in Table 1.

Overall, the experiment provides evidence that arbitrators respond meaningfully to both the economic development levels of the parties' home countries – proxies for the relative resource endowments of the two parties – when assigning costs. However, income levels affect a party's compensation only when that party also wins the dispute, which is consistent with when the arbitrators are averse to unequal cost

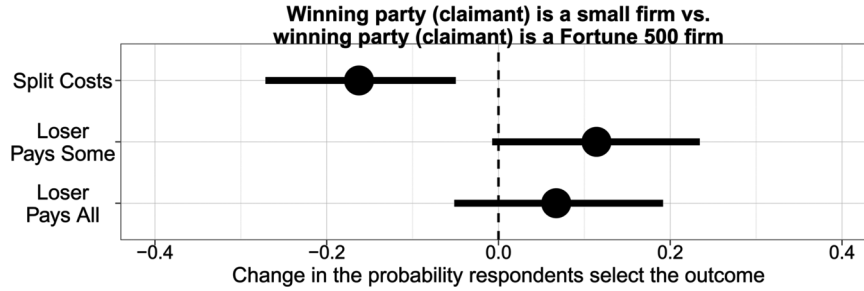

Figure 5: *Estimated Average Effects of Winning Party Wealth in an Investor–Investor Dispute*[69]

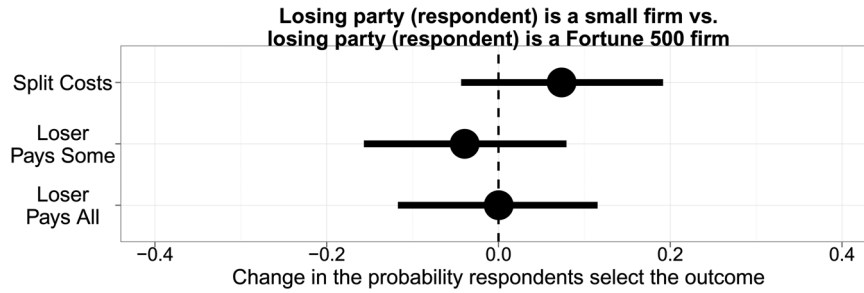

Figure 6: *Estimated Average Effects of Losing Party Wealth in an Investor–Investor Dispute*[70]

burdens post-dispute (the 'inequity aversion' heuristic) as well as when the circumstances necessitate those legal expenditures (the 'deservingness' heuristic). Hence, we describe the result as a 'David effect' – as in the biblical story of David versus Goliath. When the winning party has fewer resources than its opponent, yet it nevertheless obtains an improbable victory, arbitrators respond by increasing the magnitude of the winner's compensation.

4 External Validity and Sample Composition

Since the treatments were randomly assigned, our estimates of their effect have high 'internal validity'. We can be confident that, by design, the observed associations between treatment and outcome reflect true causal effects and are not simply due to an omitted common cause. However, it is also important to assess the 'external validity' of our estimates – the extent to which our findings can be generalized to the population of interest.

[69] Lines denote 95 per cent bootstrapped confidence intervals. Confidence intervals that do not cross zero imply that the difference between resource endowments is statistically distinguishable from 0 at ($p < 0.05$). Number of observations = 245.

[70] Lines denote 95 percent bootstrapped confidence intervals. Confidence intervals that do not cross zero imply that the difference between resource endowments is statistically distinguishable from 0 at ($p < 0.05$). Number of observations = 245.

Table 1: Summary of Treatment Effects for Investor–State and Investor–Investor Disputes

	Investment Arbitration Vignette		**Commercial Arbitration Vignette**	
	Winning party has lower income (versus higher)	**Losing party has lower income (versus higher)**	**Claimant (winning party) is a small firm (versus Fortune 500)**	**Respondent (losing party) is a small firm (versus Fortune 500)**
Split costs	–0.121*	0.054	–0.162**	0.074
	(0.057)	(0.056)	(0.056)	(0.059)
Losing party pays some costs	0.030	0.029	0.114+	–0.039
	(0.052)	(0.052)	(0.061)	(0.059)
Losing party pays all costs	0.100	–0.076	0.067	0.0003
	(0.062)	(0.061)	(0.061)	(0.059)
Losing party pays some or all costs	0.130*	–0.047	0.181**	–0.039
	(0.059)	(0.058)	(0.060)	(0.062)

Notes: Point estimates denote differences in the share of respondents choosing each outcome. Bootstrapped standard errors in parentheses.
+ = $p < 0.10$; * = $p < 0.05$; ** = $p < 0.01$

One potential critique of our experiment is that our response rates limit generalizability since both the share of individuals contacted who replied to our survey request and the share of actual respondents that completed the survey are small. However, response rates alone do not indicate that our sample estimates are non-representative. Indeed, low response rates are ubiquitous in both academic and commercial survey research.[71] It is very challenging to motivate busy individuals to dedicate scarce time to answering survey questions. But, for low response rates to generate bias, the set of individuals who chose to respond must be qualitatively different from those who did not respond with respect to the parameter of interest – in this case, our treatment effects. It is not enough to state that non-respondents might have different characteristics; those characteristics would also have to modify the treatment effect. Otherwise, non-response can be treated essentially as the random selection of respondents from the target population.[72]

While it is possible that our sample composition may differ in meaningful ways from that of the target population, prior research on non-response in survey research suggests that representativeness problems induced by sample selection are likely to be

[71] E.g., a 2012 report on non-response in Pew Research Center surveys found the response rate of a standard Pew telephone survey is just 9 per cent compared to 36 per cent in 1997. See Pew Research Center, *Assessing the Representativeness of Public Opinion Surveys* (2012), available at www.people-press.org/2012/05/15/assessing-the-representativeness-of-public-opinion-surveys/.

[72] Holbrook, Krosnick and Pfent, 'The Causes and Consequences of Response Rates in Surveys by the News Media and Government Contractor Survey Research Firms', in J.M. Lepkowski *et al.* (eds), *Advances in Telephone Survey Methodology* (2007) 500.

small. In an analysis of response rates across 14 major survey research firms, Allyson Holbrook, Jon Krosnick and Alison Pfent found that, while surveys with higher response rates tend to be somewhat more representative of the target population on demographic covariates, the association is relatively weak. They conclude: '[D]evoting substantial effort and material resources to increasing response rates may have no measurable effect on the demographic accuracy of a survey sample.'[73] Still, it may be possible that our sample may meaningfully differ from the target population of ISDS arbitrators. While, ideally, we would have surveyed every single individual who has served as an arbitrator in an investor–state dispute, privacy as well as cost and time constraints made such a study infeasible. However, while not all of our sampled respondents have served as investor–state arbitrators, we found that the sample is remarkably similar to the actual composition of ISDS arbitrators on key covariates.

We examined the distributions of three possible effect modifiers that we were able to measure for both a large set of our respondents and for all of the individuals in the target population: gender, the country's legal tradition and employment background. Of the 257 respondents who finished the survey and completed the investor–state vignette, 240 also provided information on all three of these covariates. We then gathered and compared data for a reasonable target population. Ideally, we would know the distributions for all ISDS arbitrators across all institutions throughout history. However, many ISDS disputes are confidential, making it hard to determine which exact individuals actually make up that population. Moreover, covariate information for some arbitrators may be difficult to obtain, particularly for older arbitrators without extensive presence on the Internet. We instead compare our sample to the closest feasible target population – all of the arbitrators who served on tribunals constituted under the auspices of ICSID between 2010 and 2015.[74]

We found 188 unique arbitrators that served on at least one ICSID tribunal constituted in that five-year period. For each arbitrator, we used publicly available data to code our three variables of interest. We then calculated the proportion of arbitrators in each coded category for each variable and compared these target proportions to the proportions observed within our sample. Figure 7 visualizes the differences between the sample and population covariate distributions. While small discrepancies are evident, we were pleased to find that our sampled arbitration experts are similar to the pool of actual arbitrators.

[73] *Ibid.*, at 527. Other scholars draw similar conclusions in an analysis of two sets of Pew Research Center surveys with comparable questions but varying non-response rates. They find that on over 90 per cent of questions, the difference between estimates from the low response rate and high response rate surveys was statistically indistinguishable from zero. See, e.g., Keeter *et al.*, 'Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey', 70 *Public Opinion Quarterly* (2006) 759.

[74] Because ICSID disputes are registered publicly, we know the names of all arbitrators serving during this time period. Additionally, constraining the sample to recent years yields a more policy-relevant target population in debates over the future of ISDS and its legitimacy care specifically about how modern arbitrators are deciding cases.

First, we found that the distribution of ICSID arbitrators in recent cases is predominantly male, consistent with the general criticisms that women are under-represented in ISDS. Only 9.5 per cent of arbitrators who served on at least one ICSID tribunal in the 2010–2015 period were women. This skew is also evident in our sample, which is very close to the population distribution. However, our sample does contain a slightly larger share of women arbitration experts – 13.8 per cent – and the 95 per cent confidence interval for this proportion just barely fails to overlap the population truth. Nevertheless, a 4.5 per cent difference in gender share is unlikely to result in an extreme difference between sample and population effects, particularly when we have little reason to believe male and female arbitrators respond in significantly different ways to treatment.

Arbitrator nationality is the second variable we considered. Just as in the actual population, the majority of respondents in our sample stated that they were nationals of European or North American countries.[75] However, a fair number of our respondents also indicated that they were from Latin American, Asian or African countries, representing arbitrators from both rich and poor nations. We are unlikely to be simply capturing an effect for arbitrators from a single country. But nationality may be correlated with other salient factors that modify the treatment effect in our experiment. One possible mechanism considers an arbitrator's legal culture and legal training. Specifically, two dominant approaches with respect to the allocation of costs exist in domestic courts – the English and American rules. Under the English rule, the loser pays litigation costs, whereas under the American rule, each party pays its own costs.[76] Whether a respondent was trained in the US system may therefore affect their underlying propensity to shift costs and thereby responsiveness to treatment. Unfortunately, determining each arbitrator's legal training is challenging given that nationality is not a perfect proxy – many arbitrators attend foreign law schools to obtain a master's or doctorate degree abroad after an original law degree. However, a simple comparison of the number of US arbitrators versus non-US arbitrators in our sample and in the population is a reasonable first cut at this question.

In the population, we found that about 16 per cent of all arbitrators are US nationals (including those with dual nationality). Within our sample, however, that share jumps to 27 per cent. The slight over-representation of Americans makes sense given that our survey was administered in English (and not in French or Spanish, the two other official languages of ICSID). However, US respondents still make up a minority of our sample, which makes it hard to attribute the observed patterns exclusively to effects within a single nationality of respondents. Moreover, if effect heterogeneity by nationality does exist, it makes our finding even more interesting and policy relevant as it implies the parties can exert some control over the 'underdog effect' in practice through the arbitrator appointment process. While the investigation of the sub-group

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

[76] Riesenberg, 'Fee Shifting in Investor-State Arbitration: Doctrine and Policy Justifying Application of the English Rule', 60 *Duke Law Journal* (2010) 1000.

effects in our experiment is unreliable given how small our samples become once we segment our 257 respondents by covariates, subsequent studies should consider explicitly over-sampling nationalities of interest to obtain greater precision in estimating interaction effects.

Finally, we considered our arbitrators' career backgrounds. When administering our survey, we allowed respondents to indicate their current area of employment via four general categories: private law, academia, government or other. We chose these categories because they reflect the most common career backgrounds of international arbitrators.[77] Respondents who answered 'other' were recoded into one of the three remaining categories based on their open-ended response. We found that most respondents answering 'other' described themselves as independent arbitrators and were therefore recoded as being in the private sector. A few also noted employment in an international organization, which we chose to group with individuals selecting 'government' due to similarities in career trajectories. We applied the same coding scheme to our target population using each arbitrator's most recent area of employment based on data collected from publicly available information sources (for example, websites and curricula vitae).

As shown in Figure 7, the majority of respondents work in the private sector. While this remains the case among ICSID arbitrators, a larger share of the target population originates in academia or the public sector. About 92 per cent of our respondents indicated private sector employment compared to 69 per cent in the ICSID group. While this difference suggests a potential limitation of our sample, it is the case that arbitrators with backgrounds in the private sector comprise the largest sector of the overall international arbitration pool. Even in the unlikely case that private sector arbitrators respond to treatment in a manner entirely different from academic or public sector arbitrators, our experiment still has real-world utility, as it credibly identifies the treatment effect for that private sector sub-group – the modal career background for ICSID arbitrators. However, given that investment arbitrators often move between each of these three employment paths throughout their career – academics consult for the private sector, and many arbitrators working for firms also have experience in the public sector – we do not expect career path to modify the treatment effect to such an extreme.

In sum, while our surveyed arbitrators are most certainly not perfectly representative of the true population of arbitrators, it would be inaccurate to treat our sample as simply a convenience sample. We explicitly drew respondents from lists provided by arbitral institutions (and from those individuals mainly involved in ISDS). These are individuals who not only have legal training but who also have specific experience with arbitration proceedings. They are precisely the types of individuals who get selected into the exclusive pool of investor–state arbitrators. The fact that 40 of our participants (roughly 21 per cent of all of the arbitrators appointed to a tribunal between 2010 and 2015) had previously served in ISDS tribunals highlights the fact

[77] Costa, 'Comparing WTO Panelists and ICSID Arbitrators: The Creation of International Legal Fields', 1 *Oñati Socio-Legal Series* (2011) 1, at 14.

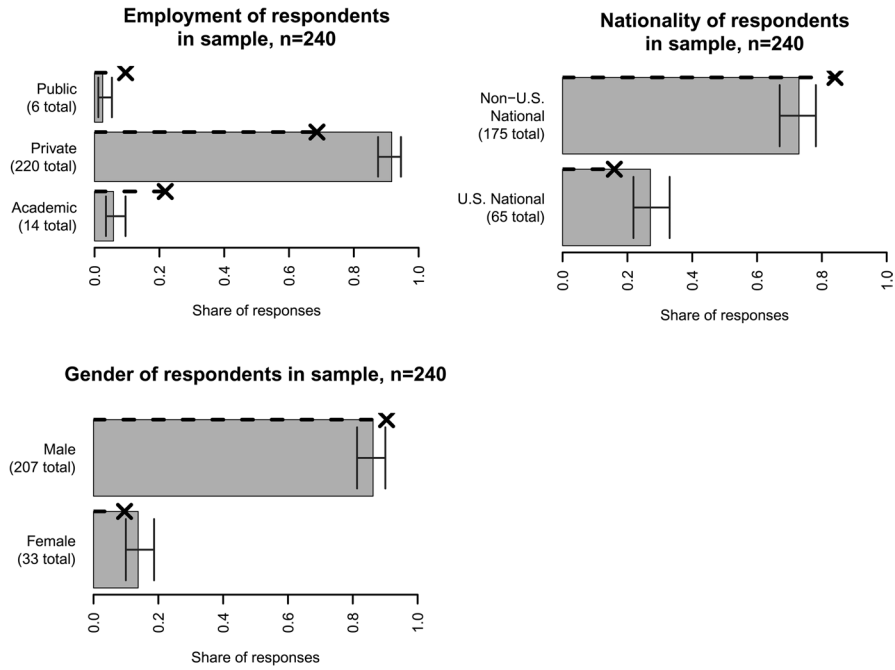

Figure 7: *Characteristics of Arbitrators in the Experimental Sample*[78]

that we are, in part, capturing the population of interest. Even if response rates are low, the weight of the evidence, both from past studies of non-response bias and from a comparison of covariate distributions between sample and target populations, suggests that we can make reasonable generalizations from our experiment about how actual arbitrators think and behave.

C *Limitations of Our Experimental Approach*

Our experimental approach has clear limitations. For one, it is unable to assess how exactly the observed effects, with respect to individual arbitrators in a control setting, would impact the final outcome of a collective body in a deliberative environment. This is certainly an important limitation not only of our survey experiment but also of most experiments involving individuals who engage in collective decision-making processes. This limits our ability to predict what the impact of the identified effect would be with respect to the decision of an arbitral panel, particularly if the process in question involves compromises or strategic actions for the professional advancement of those involved.[79]

Future experiments could add more insight on this question – taking the unit of analysis to be the arbitral tribunal rather than the individual arbitrator. Prior research

[78] Crosses denote proportions observed among ICSID arbitrators 2010–2015. Lines denote 95 per cent confidence intervals.

[79] Paulsson, *supra* note 57.

in social sciences provides some guidance as to what we might expect from this analysis. Specifically, researchers have found that in collective deliberations, biases tend to be reinforced and not corrected – an effect known as bias accentuation.[80] In other words, preferences and perceptions could become further entrenched throughout the arbitrators' deliberation process rather than revised. This is in contrast to the classic assumption that tribunals attenuate biases via the symmetrical structure of party appointments. While that may be true for some types of biases that happen to be correlated to the party of appointment, it may not be the case for other types of biases and heuristics such as the one we examine, which may be common to all members of the tribunal.

Other factors such as the specialization, scrutiny and incentives affecting ISDS limit the broad generalization of this analysis to other fields of arbitration. ISDS is unique as a form of dispute settlement, and the perception of the power differentials between claimants and respondents among scholars and arbitrators is, to some extent, reflective of the field's complex politics.[81]

4 ISDS, Fairness and the Support for Underdogs

How does our finding that arbitrators may be affected by heuristics that connect potentially irrelevant information about material short-handedness with a norm of 'equity' and 'deservingness' inform the current debate over, and practice of, ISDS? In this final section, we provide some preliminary thoughts and argue that there are at least two important implications of our core empirical result.

A *ISDS and the Current Debate over Its Inclusion in Economic Treaties*

Developing countries have primarily been the state parties to ISDS proceedings, although claims against developed states are not uncommon. Conversely, the investors have been mostly companies controlled by nationals or corporate groups based in Europe (Netherlands, United Kingdom, Germany, France and Italy) and North America (USA and Canada).[82] While some ISDS cases in fact address 'excessive' interventions or opportunistic actions on the part of the host state, plenty are 'fence' cases – claims where reasonable people could disagree about the reasonable outcome.

The general perception of ISDS (at least in the developing world, Europe and now the USA), however, does not reflect this distribution of fence versus excessive intervention cases. For example, very few actors – mostly transnational corporate interests –

[80] Fiedler and Kutzner, 'Information Sampling and Reasoning Biases: Implications for Research in Judgment and Decision Making', in G. Keren and G. Wu (eds), *The Wiley Blackwell Handbook of Judgment and Decision Making* (2015) 380.

[81] Roberts, *supra* note 14; Trakman, 'The ICSID under Siege', 45 *Cornell International Law Journal* (2012) 603.

[82] Parra, *supra* note 9, at 136; see United Nations Conference on Trade and Development, *World Investment Report 2015*, available at http://unctad.org/en/PublicationsLibrary/wir2015_en.pdf.

explicitly support ISDS and its inclusion in treaties like the Transatlantic Trade and Investment Partnership (TTIP).[83] When the European Commission opened up consultations with the public over the inclusion of ISDS into the trade pact between the EU and the USA, the vast majority of individuals filing replies expressed strong opposition to ISDS.[84] Only 60 companies submitted comments, most of which were in support of ISDS. Among these firms, however, Phillip Morris' endorsement seems particularly damaging to the cause of ISDS, given that the recent cases brought by the tobacco giant are generally perceived as being audacious. Cecilia Malmström, Europe's trade commissioner, concludes that in Europe 'there is a huge scepticism against the ISDS'.[85] Among academics, as well, the merits of ISDS are hotly disputed. As explained by leading scholars in the field, '[i]t is easy to find respectable academics arguing that it is something close to an unmitigated good, and others, just as respectable, arguing the opposite'.[86]

This current dissatisfaction may be the result of a perception that ISDS is unfair and does little to improve developing host states (the traditional underdogs), while empowering rich multinational corporations. This perception is not without merit. For one, corporations that take part in the system economically dwarf smaller states. Chevron, a repeat player in ISDS, had revenues in 2014 of US $200 billion, putting it in a remarkable third place in the Fortune 500 list.[87] The only African nation that has a gross domestic product (GDP) above that number is Nigeria, the African powerhouse. Moreover, according to the World Bank's 2015 listing of countries by GDP, only 26 countries have a GDP larger than the annual revenues of Walmart, currently the world's largest corporation, despite corruption scandals and controversial labour practices.[88] Some 70 corporations on the Fortune 500 list had annual revenues of US $100 billion or more, while only 60 nations had a GDP in 2014 in excess of US $100 billion.[89]

This highlights an oddity of the central feature of the international investment law system: corporations that rely on ISDS can be affected by excessive government intervention, but they are generally not in the position of being underdogs in terms of

[83] Poulsen, Bonnitcha and Yackee, Transatlantic Investment Treaty Protection: Paper no. 3 in the CEPS-CTR Project on 'TTIP in the Balance', CEPS Special Report no. 102 (2015). Transatlantic Trade and Investment Partnership (draft dated 12 November 2015).

[84] Online Public Consultation on Investment Protection and Investor-to-State Dispute Settlement (ISDS) in the *Transatlantic Trade and Investment Partnership Agreement*, Commission Consultation, 13 July 2014, available at http://trade.ec.europa.eu/consultations/index.cfm?consul_id=179.

[85] Quoted in C. Oliver, 'Public Backlash Threatens EU Trade Deal with the US', *Financial Times* (13 January 2015).

[86] Poulsen, Bonnitcha and Yackee, *supra* note 83.

[87] See 'Fortune Globa 500', *Fortune* (2016), available at http://fortune.com/global500/list.

[88] *Ibid.*; see, e.g., M. Trottman and S. Banjo, 'Wal-Mart Accused of Violating Workers' Rights', *Wall Street Journal* (15 January 2014); D. Barstow and A. Xanic von Bertrab, 'The Bribery Aisle: How Wal-Mart Got Its Way in Mexico', *New York Times* (17 December 2012).

[89] World Bank, *GDP (Current US$)* (2015), available at http://data.worldbank.org/indicator/NY.GDP.MKTP.CD.

material resources – hence, the generalized perception of the overpowered investor.[90] But there is also a silver lining. Our experiment suggests that a bias towards wealthy firms is not a characteristic of the arbitrators themselves. Indeed, arbitrators meaningfully react to material short-handedness in a way that compensates the weaker party more heavily. The apparent advantage held by well-financed investors is likely more a function of the current institutional arrangements underpinning ISDS than an inherent aspect of arbitration. This may be a positive development for investment law as a whole since arbitrators are just as capable of sympathizing with both underdog investors and states and even attribute some value to rectifying the perceived inequality of legal 'arms.' In other words, within the political economy of international investment law, we find some minimum expectation of fairness between litigating parties – a preference for a kind of equality of legal opportunity.

Many prescriptions have been made to make BITs and ISDS fairer endeavours.[91] At the political level, however, policymakers need a better way to justify ISDS on the basis of the rule of law and fairness for host states, not only for wealthy investors, and to reflect such arguments in the core design features of BITs. At the rhetorical level, this exercise includes explaining ISDS as mitigating imbalances or improving the position of state underdogs by formally maintaining a restrictive policy towards the diplomatic espousal of investment claims by more powerful states.[92] In some contexts, ISDS may create a mutually beneficial setting for all of the parties involved, particularly when compared to the alternatives. Such benefits include compartmentalizing potentially daunting conflicts between states into individual disputes. This may help to 'depoliticize' internationally distressing conflicts, liberating energy towards building constructive relationships.[93] It can also have a therapeutic effect on affected investors of more modest means and help to tame the use of political power against underdogs – businesses and states. Without addressing the disconnect described above, the political sustainability of ISDS will continue to be in doubt.

To be sure, the challenge is not simply a political one. A framing of this nature requires treaty design features that truly sustain ISDS as a tool for balancing power to benefit the 'have-nots' as opposed to over-empowering the 'haves'. One could start by reinserting important flexibilities into BITs that were originally conceived in the system's earlier days and that have diminished, in part, as a result of the political efforts and bargaining power of some developed states.[94] These flexibilities include: (i) allowing states to require the exhaustion of local remedies prior to accessing ISDS;

[90] For a similar account on the power of a transnational corporation, see Martinez, 'New Territorialism and Old Territorialism', 99 *CLR* (2014) 1387.

[91] Kingsbury and Schill, *supra* note 1, at 75; Trakman, *supra* note 81.

[92] Shihata, 'Towards a Greater Depoliticization of Investment Disputes: The Role of ICSID and MIGA', 1 *IRFILJ* (1986).

[93] ICSID, Annual Report (1985); cf. Paparinskis, 'The Limits of Depoliticisation in Contemporary Investor-State Arbitration', 3 *Select Proceedings of the European Society of International Law* (2010) 271.

[94] Alschner, 'Americanization of the BIT Universe: The Influence of Friendship, Commerce and Navigation (FCN) Treaties on Modern Investment Treaty Law', 5 *Goettingen Journal of International Law* (2013) 455.

(ii) permitting countries to stipulate that their relationship with foreign investors is governed by domestic law; (iii) excepting especially sensitive policy areas from ISDS challenges (that is, tobacco control or access to medicine measures) or certain investments from the protection of BITs more generally; (iv) prohibiting an investor home state from giving diplomatic protection to nationals unless the host state demands it for justifiable reasons; and (v) limiting most-favourable-nation treatment to narrowly defined substantive matters.[95] These basic concepts could be easily operationalized if treaty parties decide that improving ISDS is less costly than completely replacing it with a new 'court-like' system.[96] Nevertheless, the flexibility within the ICSID Convention exists and may align ISDS with what we believe is the threshold expectation observed in our narrow experiment.

B *The David Effect and Litigating Difficult Cases*

A prominent lawyer, defending a claim involving the regulatory change of a poor Central American nation against a large European company, once explained his legal strategy of the case as follows: 'I'm trying to bring the moral character of this case into display.' The lawyer's client decisively won the case after showing, among other misjudgments, that the investor's chief executive officer arrogantly threatened government regulators of the poor developing country. Conversely, the melodramatic behaviour of Prime Minister Muhammad Mossadegh, who defended Iran's nationalization of the Anglo-Iranian Oil Company without the assistance of outside legal counsel in preliminary proceedings before the International Court of Justice, only served to make him the butt of jokes among lawyers and diplomats.[97] The ageing and frail Mossadegh overplayed his country's underdog card, neglecting to make a legal case (as opposed to a moral one) and arguably losing the decisive stage for Iran.

Our experiment has an important implication for the way that developing host states can turn a disadvantage into a slight advantage in actual litigation. States should always bring the best possible legal arguments, but they should not hesitate to highlight how they may have been affected by unfairness inherent in the system as well as by any endowment imbalances, emphasizing their position as underdogs. As we observe from our experiment, there is a significant probability that such shorthandedness can be translated into a heuristic with limited, but important, material consequences. This does not mean that arbitrators necessarily always favour underdogs. However, if a host state is perceived as an underdog and wins, it may have a sharp advantage at least when it comes to collecting costs. We can only presume that

[95] I. Shihata, 'ICSID and Latin America', *News from ICSID* (1984), at 2.

[96] *EU Finalizes Proposal for Investment Protection and Court System for TTIP*, Commission Press Release, 12 November 2015, available at http://trade.ec.europa.eu/doclib/press/index.cfm?id=1396; European Union TTIP Team, *We're Proposing a New Investment Court System*, 16 September 2015, available at https://twitter.com/EU_TTIP_team/status/644110990242639873.

[97] Rolin, 'The International Court of Justice and Domestic Jurisdiction: Notes on the Aglo-Iranian Case', 8 *IO* (1954) 36.

this tactical adaptation could have an impact when arbitrators exercise discretion but not in all aspects of investment disputes.

An alternative explanation to our main empirical finding suggests that arbitrators may be interested in ensuring buy-in on the part of state underdogs without deterring the use of the system by investors that could find themselves in an underdog position. Hence, litigants perceived as being economically weaker that win receive more compensation because arbitrators want to ensure that litigants are not deterred from using the system.

Finally, the fact that the legitimacy of ISDS is connected with a minimum expectation of fairness should not surprise anyone. However, what our experiment shows is how certain elements that may be unrelated to the actual merits of a case can act as shorthand for such a fairness assessment. In this vein, the 'deservingness' and 'inequity aversion' heuristics may have deep connections with the way arbitrators actually use discretion. Scholars should continue to clarify how this may affect the distribution of outcomes in ISDS.

5 Conclusion

Does investor–state arbitration favour wealthy investors over states or does it protect risk-taking firms from predatory governments? While the institutional arrangements of ISDS may make one underdog narrative more or less prevalent than the other among those disputes that are litigated, it is important to understand how the arbitrators themselves consider material inequalities when adjudicating a dispute. The difficulty of answering this question with observational evidence remains high, and disentangling the decision-making logics of arbitrators from the limited number of filed disputes is presently impossible. However, using an experimental approach allows us to isolate our analysis on the arbitrators in particular. The results of our survey experiment provide strong evidence that arbitrators may be prone to the 'David effect' – biased towards the perceived underdog or 'weaker' party when this party wins. While our results would benefit from confirmation in broader settings, they suggest that the legitimacy of legal regimes is connected to a minimum expectation of fairness among the parties involved. This threshold expectation highlights the importance of recasting ISDS as a tool to mitigate imbalances between different actors.