March 15, 2014

Conflating Politics and Development? Examining Investment Treaty Arbitration Outcomes

Susan D. Franck, *American University Washington College of Law*

| ARTICLE |

CONFLATING POLITICS AND DEVELOPMENT? EXAMINING INVESTMENT TREATY ARBITRATION OUTCOMES

SUSAN D. FRANCK[*]

    International dispute settlement is an area of ongoing evaluation and tension within the international political economy. As states continue their negotiations for the Trans-Pacific Partnership (TPP) and the Transatlantic Trade and Investment Partnership (TTIP), the efficacy of international arbitration as a method of dispute settlement remains controversial. Whereas some sing its praises as a method of protecting private property interests against improper government interference, others decry investment treaty arbitration (ITA) as biased against states. The literature has thus far not disentangled how politics and development contribute to investment dispute outcomes. In an effort to control for the effect of internal state politics, this Article offers the first analysis of ITA outcomes, focusing on respondent states' development status while simultaneously controlling for states' democracy levels. Using a dataset of 159 final ITA awards from prior to January 2012, the Article conducts quantitative analyses of outcomes as a function of raw wins and losses, amounts awarded, and relative investor success. Initially, when evaluating outcomes based on a respondent state's membership in the Organisation for Economic Cooperation and Development (OECD) or a state's score on the UN Development Programme (UNDP) Human Development Index, it was not possible to identify a reliable link to outcomes. Only defining a respondent's development status using a World Bank classification generated reliable differences for Upper-Middle income states, and only for two measures of outcome — namely raw wins and amounts awarded. Using the World Bank measure, there was no statistically significant relationship with relative investor

success. None of these analyses, however, controlled for the level of internal state democracy to identify how democracy levels, which can reflect good governance infrastructure, might contribute to outcomes. After controlling for the effect of a state's internal democracy levels, twelve analyses were unable to identify a reliable link with ITA outcomes and development status irrespective of how development status was defined. While the Article cannot conclusively exclude the possibility of systemic bias in ITA against the developing world, it provides additional evidence suggesting the potential absence of such bias or the importance of alternative explanatory variables. The results also suggest that focusing on development status alone may be unwarranted, and future research should explore internal levels of democracy or other indicators of good governance, which could be associated with the decreased risk of a state loss. The Article concludes that normative choices focused solely on respondent state development status miss an opportunity to craft normative solutions tailored to redress tangible problems. By focusing on variables that demonstrably contribute to variance in ITA outcomes, stakeholders could construct more appropriate international dispute settlement processes in a time of international economic transition.

INTRODUCTION

As global economic activity continues to expand during a time of worldwide economic shifts, international investment law has emerged as an area of mounting interest. In 2012, the global total of foreign-direct-investment stock was approximately US$23 trillion.[1] Meanwhile, the number of international investment treaties, which offer substantive investment rights and contain *ex ante* agreements permitting investors to require states to arbitrate substantive violations, continues to increase. With trillions of dollars protected by at least one international investment agreement (IIA) and theoretically subject to treaty-based arbitration,[2] changes in political and economic circumstances generate conflict, and investors have exercised their new treaty rights. Since the first award was rendered in 1990,[3] the United Nations Conference on Trade and Development (UNCTAD) estimates there have been more than 500 known disputes.[4]

With increased political, economic, and legal scrutiny on international investment and dispute resolution, investor-state dispute settlement has become a global hot button issue, as demonstrated by policy debates and negotiations surrounding the Trans-Pacific Partnership (TPP) and

---

1. U.N. CONFERENCE ON TRADE & DEV. [UNCTAD], WORLD INVESTMENT REPORT 2013: GLOBAL VALUE CHAINS: INVESTMENT AND TRADE FOR DEVELOPMENT, at xv, 217 Annex tbl. 2, U.N. Sales No. E.13.II.D.5 (2013) [hereinafter UNCTAD, WIR 2013].

2. *See id.* at 217 Annex tbl. 2; UNCTAD, WORLD INVESTMENT REPORT 2011: NON-EQUITY MODES OF INTERNATIONAL PRODUCTION AND DEVELOPMENT, at 102–03, 191, U.N. Sales No. E.11.II.D.2 (2012).

3. Asian Agric. Prods. Ltd. (AAPL) v. Republic of Sri Lanka, ICSID Case No. ARB/87/3, Final Award (June 27, 1990), *available at* http://www.italaw.com/sites/default/files/case-documents/ita1034.pdf.

4. UNCTAD, WIR 2013, *supra* note 1, at 110–11 (through the end of 2012); UNCTAD, WORLD INVESTMENT REPORT 2012: TOWARDS A NEW GENERATION OF INVESTMENT POLICIES, at 86, U.N. Sales No. E.12.II.D.3 (2012) [hereinafter UNCTAD, WIR 2012]; *see also* UNCTAD, *IIA Issues Note No. 1: Recent Developments in Investor-State Dispute Settlement (ISDS)*, at 9–10, UNCTAD Doc. UNCTAD/WEB/DIAE/PCB/2014/3 (2014) [hereinafter UNCTAD, *IIA Issues Note 2014*], *available at* http://unctad.org/en/PublicationsLibrary/webdiaepcb2014d3_en.pdf (suggesting the number of concluded cases is more than 270, but failing to explain the definition for "concluded" cases to permit comparison to existing datasets).

Transatlantic Trade and Investment Partnership (TTIP). Part of these concerns derive from claims that investment treaty arbitration (ITA) is illegitimate[5] and that ITA detrimentally impacts the developing world to the benefit of only developed states.[6] As a result, there has been a deep re-examination of how investment treaty conflict should be resolved.[7] Some have proposed alternative or supplementary methods of managing investment treaty conflict, including mediation.[8] Others recommend exiting the system entirely.[9] In response, some states have voted with their feet to exit the international investment law regime, yet other states exhibit a preference to stay the course and rely on preexisting paradigms for dispute settlement.

While politics and legal disputes may be inextricably intertwined, the debate thus far has not disentangled internal state politics from the

---

5. *See, e.g.,* Susan D. Franck, *The Legitimacy Crisis in Investment Treaty Arbitration: Privatizing Public International Law Through Inconsistent Decisions,* 73 FORDHAM L. REV. 1521 (2004) [hereinafter Franck, *Legitimacy Crisis*]. *Compare* PIA EBERHARDT & CECILIA OLIVET, CORPORATE EUR. OBSERVATORY & TRANSNATIONAL INST., PROFITING FROM INJUSTICE: HOW LAW FIRMS, ARBITRATORS AND FINANCIERS ARE FUELLING AN INVESTMENT ARBITRATION BOOM (2012), *available at* http://corporateeurope.org/sites/default/files/publications/profiting-from-injustice.pdf (articulating concerns with ITA), *with* Charles N. Brower & Sadie Blanchard, *From "Dealing in Virtue" to "Profiting from Injustice": The Case Against "Re-Statification" of Investment Dispute Settlement,* 55 HARV. INT'L   L.J.   ONLINE   45   (Jan.   2014),   http://www.harvardilj.org/wp-content/uploads/2014/01/Brower_Blanchard_to_Publish.pdf (refuting claims about the concerns with ITA).

6. *See, e.g., ICSID—International Centre for Settlement of Investment Disputes,* BRETTON WOODS PROJECT (last modified Aug. 19, 2009), http://old.brettonwoodsproject.org/item.shtml?x=537853 ("Reasons for the vocal and mounting critiques against ICSID peg around its governance [and] its biasness [sic] in favour of rich countries . . . ."); Susan D. Franck, *Development and Outcomes of Investment Treaty Arbitration,* 50 HARV. INT'L L.J. 435, 437, 445–48 (2009) [hereinafter Franck, *Development and Outcomes*] (gathering sources identifying claims explaining that developing countries are more disadvantaged by ITA); Kevin Gallagher & Elen Shrestha, *Investment Treaty Arbitration and Developing Countries: A Re-Appraisal,* 12 J. WORLD INVESTMENT & TRADE 919 (2011) (arguing that developing countries experience larger risks than developed countries); Douglas Thompson, *Kahale Calls for Overhaul   of   BIT   System,*   GLOBAL   ARB.   REV.,   Apr.   11,   2014, http://globalarbitrationreview.com/news/article/32567/kahale-calls-overhaul-bit-system/ (referring to George Kahale's calling investment treaties "weapons of legal destruction" and stating there is "a systemic bias against states" (internal quotation marks omitted)).

7. *See generally* THE BACKLASH AGAINST INVESTMENT ARBITRATION: PERCEPTIONS AND REALITY (Michael Waibel et al. eds., 2010) [hereinafter THE BACKLASH] (discussing various concerns about ITA).

8. *See, e.g.,* Susan D. Franck, *Integrating Investment Treaty Conflict and Dispute Systems Design,* 92 MINN. L. REV. 161 (2007) [hereinafter Franck, *Dispute Systems Design*] (discussing the application of dispute systems design principles to managing investment treaty conflict); Susan D. Franck, *Using Investor-State Mediation Rules to Promote Conflict Management: An Introductory Guide,* 29 ICSID REV. 66, 66–68, 88 (2014) [hereinafter Franck, *Conflict Management*] (describing the potential application of mediation to investment treaty conflicts); Nancy A. Welsh & Andrea Kupfer Schneider, *The Thoughtful Integration of Mediation into Bilateral Investment Treaty Arbitration,* 18 HARV. NEGOT. L. REV. 71 *passim* (2013) (discussing how to integrate mediation into investment treaty dispute settlement).

9. *See, e.g., infra* notes 47–61 and accompanying text (discussing calls for reform related to eliminating ITA or placing disputes into the hands of national court judges).

resolution of investment treaty dispute settlement. Yet, attributing accountability of outcomes to development-related factors is unwarranted where unsupported by data or where other factors are more directly linked to outcome. Domestic politics — particularly where a state's political acts form the basis of the underlying dispute — potentially contribute to outcomes, and this may be similar to, or different from, development-related variables. It is therefore proper to examine the role of politics, particularly democracy levels within host states, and identify how democracy might — or might not — impact outcomes of investment treaty disputes. Disentangling the variance attributable to development and politics has the potential to generate a better understanding of variables contributing to outcomes. By isolating variables reliably linked to ITA outcomes, it becomes possible to differentiate between illegitimate bases for differential outcomes (e.g., development backgrounds that should be independent of a case's merits) and those potentially legitimate variables (e.g., a state's political environment that may be linked to the government activity that forms the basis of the cause of action).

This Article, therefore, will explore links among ITA outcomes, respondent state democracy levels, and state development indicators. It will use quantitative models to explore ITA outcomes and will re-examine outcomes as a function of the respondent state's level of democracy. While governmental structures and political activity may contribute to generating ITA conflict, the data used in this study were unable to provide evidence that a state's development background reliably affected ITA outcomes when controlling for a host state's democracy levels. While there is inevitably a risk of error, and recognizing that viewing limited development variables in isolation might create an impression of systemic bias, this Article will argue that ITA outcomes are a complicated phenomenon that is not attributable to a simple explanatory variable. While focusing on an isolated variable like development may be intuitively appealing, it nevertheless generates a risk of misperception.

In an effort to test intuition against data and address the gap within the literature, the Article will seek to offer information that stakeholders might use to craft normative reforms. It will do so by first exploring the history of investment dispute resolution and its doctrinal foundation. Next, the Article will highlight current aspects of the debate over ITA. Third, it will identify the research methodologies and hypotheses related to development, democracy, and outcomes of investment disputes. Fourth, the Article will demonstrate that although there was a reliable link between development and outcomes *only* when development was defined using a World Bank index, rather than two alternative measures of development, but after controlling for democracy levels, none of the models explored could identify a reliable link between outcome and a respondent's

development status. Fifth, the Article will interpret the results, explore the limitations, and recognize that the results require replication before definitively concluding there is no link between the respondents' development levels and ITA outcomes. In the interim, this Article will posit that a respondent's development status may be a facially enticing, but perhaps misleading, focal point. The Article will argue that ITA outcomes require a complicated, multivariate explanation, and focusing on alternative explanatory variables may ultimately prove more productive.

This Article will conclude by arguing that normative choices focused solely on a respondent state's development status miss the opportunity to generate normative solutions tailored to address real problems. International investment law should focus instead on designing dispute resolution systems tailored to generate meaningful solutions to demonstrable problems. Stakeholders should focus on other areas, such as promoting and implementing good governance practices and preventing investment conflict from erupting in the first instance. The objective should be to provide states with incentives to comply with domestic and international law, to promote a favorable investment climate, and to limit exposure to unnecessary or unmeritorious claims. Similarly, although it must be done with caution, states may also wish to focus on improving the quality of their domestic regulatory activities, promoting good governance practices, and exploring how best to promote democratic institutions within the state.[10] By focusing on variables that demonstrably contribute to the variance of ITA outcomes, ITA will generate a more legitimate and appropriate dispute resolution process during a period of international economic transition.

I. THE HISTORY AND DOCTRINE OF INVESTMENT TREATY DISPUTE RESOLUTION

ITA is a hybrid *sui generis* creature that owes its doctrinal and policy basis to both public and private international law. ITA involves the rights of private individuals, their commercial activities, the vindication of rights through an adjudicative forum, and remedies involving damage awards. ITA's public international law elements involve the law of state responsibility, the regulatory activity of state actors, and states' capacity to craft international law instruments that voluntarily curtail their regulatory authority to incentivize private commercial activity.[11]

---

10. *See infra* notes 151–54 (indicating that, although two analyses were on the cusp of significance for the group of all awards and not significant for the subset of cases where states won, there was some facial evidence that lower levels of democracy were associated with adverse awards in the form of higher amounts awarded or a greater degree of an investor's relative success).

11. *See generally* JOSÉ E. ALVAREZ, THE PUBLIC INTERNATIONAL LAW REGIME GOVERNING

Historically, investors' options for addressing disputes arising from government activity included simply ignoring investment conflict and accepting it as a cost of doing business (and pricing commercial risk accordingly), obtaining political risk insurance coverage and passing the cost of conflict on to a secondary market, or asking the investor's home state to invoke its diplomatic rights on behalf of the investor.[12] Other solutions required a political assessment, including declarations of war, exercises of "gunboat diplomacy," lobbying the investor's home state to invoke formal diplomatic relief, or soliciting an investor's home state to espouse its claim at the International Court of Justice (ICJ).[13] Yet indirect dispute resolution under international law was largely unsatisfactory, as it prohibited direct and immediate access to neutral dispute resolution, politicized dispute resolution by placing it within the sphere of international relations, and generated a remedy for the state but not for the damaged investor.[14]

Against this backdrop, states and investors began using IIAs to permit international arbitration of international investment law obligations. ITA was an evolutionary step intended to provide investors with direct rights and a real forum for non-politicized rule of law adjudication. The objective was to permit states to avoid making the difficult — and ironically often political — decision of whether to espouse an investor's claim.[15] ITA granted investors the direct choice to decide whether, in light of the net costs and benefits, pursuing a claim was appropriate. ITA also offered a tried and tested process — namely international arbitration —

INTERNATIONAL INVESTMENT (2011) (discussing investment treaty dispute settlement); KRISTA NADAKAVUKAREN SCHEFER, INTERNATIONAL INVESTMENT LAW: TEXT, CASES AND MATERIALS (2013) (same); Anthea Roberts, *Clash of Paradigms: Actors and Analogies Shaping the Investment Treaty System*, 107 AM. J. INT'L L. 45 (2013) (discussing private and public law aspects of state liability under international investment law).

12. *Cf.* STEPHAN W. SCHILL, THE MULTILATERALIZATION OF INTERNATIONAL INVESTMENT LAW 1–22 (2009) (tracing the evolution of international investment law); NATHAN M. JENSEN ET AL., POLITICS AND FOREIGN DIRECT INVESTMENT 115–45 (2012) (discussing how firms can lobby related to foreign investment disputes); Franck, *Legitimacy Crisis*, *supra* note 5, at 1521–35.

13. *See* RUDOLF DOLZER & CHRISTOPH SCHREUER, PRINCIPLES OF INTERNATIONAL INVESTMENT LAW (2012) (describing the history of international investment law); *see also* Anthea Roberts, *State-to-State Investment Treaty Arbitration: A Hybrid Theory of Interdependent Rights and Shared Interpretive Authority*, 55 HARV. INT'L L.J. 1, 15–17 (2014) (discussing gunboat diplomacy and the history of ITA); Jason Webb Yackee, *Controlling the International Investment Law Agency*, 53 HARV. INT'L L.J. 391 (2012).

14. SUSAN D. FRANCK, INVESTMENT TREATY ARBITRATION: MYTHS, REALITIES AND COSTS (forthcoming 2015) [hereinafter FRANCK, MYTHS & REALITIES]; *see also* Susan D. Franck, *The ICSID Effect? Considering Potential Variations in Arbitration Awards*, 51 VA. J. INT'L L. 825, 833–34 (2011) [hereinafter Franck, *ICSID Effect*]; Sergio Puig, *Recasting ICSID's Legitimacy Debate: Towards a Goal-Based Empirical Agenda*, 36 FORDHAM INT'L L.J. 465, 471–75 (2013) (discussing the history of investor-state dispute settlement).

15. *See* Andreas F. Lowenfeld, *The ICSID Convention: Origins and Transformation*, 38 GA. J. INT'L & COMP. L. 47, 51–52 (2009).

whose historical pedigree offered an opportunity for adjudicative neutrality. Moreover, the existence of two preexisting treaties, namely the New York Convention[16] and the International Centre for Settlement of Investment Disputes (ICSID) Convention,[17] offered a reliable, established, and streamlined enforcement regime. In theory, ITA generated a more balanced playing field and fostered adjudication of investment disputes in a manner designed to minimize commercial risk, decrease political risk, and maximize rule of law.[18]

The theoretical framework, however, must be placed within a doctrinal context to understand the basis of the arbitration agreement and applicable law.[19] First, it is critical to observe that ITA derives from IIAs. The net objective of an IIA is to entice inbound foreign investment and to protect a state's own investors abroad while minimizing the risk of state liability. IIAs typically involve a pair or group of countries signing a treaty, such as the Dominican Republic-Central American Free Trade Agreement (CAFTA-DR),[20] that promises to protect international investment within the territory of the host state. IIAs grant reciprocal investment rights — of both a procedural and substantive nature — to the foreign investors of the signatory countries. In some, but not all, instances, IIAs secure the intended benefit of foreign investment.[21]

Procedurally, IIAs offer clear dispute resolution rights, including the right to arbitrate treaty disputes. After complying with certain prerequisites, such as notification of disputes and attempting amicable settlement, IIAs permit investors to directly bring a claim against a state to a tribunal in which both parties appoint adjudicators. ITA also has the virtue of preventing blanket claims of immunity from suit — whether by virtue of sovereign immunity or political question doctrines — that might

---

16. Convention on the Recognition and Enforcement of Foreign Arbitral Awards, June 10, 1958, 21 U.S.T. 2517, 330 U.N.T.S. 38.

17. Convention on the Settlement of Investment Disputes Between States and Nationals of Other States, International Centre for Settlement of Investment Disputes, *opened for signature* Mar. 18, 1965, 17 U.S.T. 1270, 575 U.N.T.S. 159 [hereinafter ICSID Convention].

18. FRANCK, MYTHS & REALITIES, *supra* note 14.

19. Much of the overview of ITA doctrinal context and arbitration mechanics in this Part derives from the author's earlier summaries. *See, e.g.*, Franck, *ICSID Effect*, *supra* note 14, at 835–37.

20. Dominican Republic-Central America Free Trade Agreement, U.S.-Costa Rica-Dom. Rep.-El Sal.-Guat.-Hond.-Nicar., Aug. 5, 2004, 43 I.L.M. 514 (2004).

21. *See* Jennifer L. Tobin & Susan Rose-Ackerman, *When BITs Have Some Bite: The Political-Economic Environment for Bilateral Investment Treaties*, 6 REV. INT'L ORGANIZATIONS 1 (2011) (reviewing conditions where IIAs are linked with increased investment); Todd Allee & Clint Peinhardt, *Contingent Credibility: The Impact of Investment Treaty Violations on Foreign Direct Investment*, 65 INT'L ORG. 401 (2011) (discussing the utility of IIAs and the capacity to secure investment flows, particularly in light of threatened or successful ITA disputes). *Compare* Todd Allee & Clint Peinhardt, *Delegating Differences: Bilateral Investment Treaties and Bargaining Power over Dispute Resolution Provisions*, 54 INT'L STUD. Q. 1 (2010), *with* Jason Webb Yackee, *Do Bilateral Investment Treaties Promote Foreign Investment? Some Hints from Alternative Evidence*, 51 VA. J. INT'L L. 397 (2010).

otherwise prevent investors from bringing their claims in national courts.[22] Despite intending to provide a neutral, legitimate forum for states to resolve disputes without allegations of bias, ITA has fallen prey to claims that it is politicized and biased — particularly against developing states.[23]

Substantively, IIAs involve state promises that foreign investors will receive certain minimal treatment, including the right to freedom from unlawful expropriation without proper compensation, the right to freedom from discrimination, and guarantees of fair and equitable treatment. These rights are analogous to using international law to provide an economic bill of rights for foreign investors.[24] Some legal disputes arising under IIAs involve public law elements, such as Zimbabwe's expropriation of land belonging to certain white farmers or the imposition of an environmental regulation that has a *de facto* disparate impact on foreign companies.[25] Other disputes have a more commercial flavor, such as the revocation of a banking license or breach of contract.[26] Certain disputes, although not necessarily representative of the larger whole, have become iconic symbols in ITA. For example, groups of investors have sued Argentina for its 2001 currency crisis that led to the devaluation of the Argentine peso and decreased the value of investments, particularly those denominated in U.S. Dollars.[27] Similarly, Phillip Morris sued Australia and Uruguay for their plain-packaging cigarette regulations as an unlawful expropriation of its intellectual property rights.[28]

---

22. *See, e.g.,* Stephen E. Blythe, *The Advantages of Investor-State Arbitration as a Dispute Resolution Mechanism in Bilateral Investment Treaties,* 47 INT'L LAW. 273, 274–79 (2013).

23. *See* Leon E. Trakman, *The ICSID Under Siege,* 45 CORNELL INT'L L.J. 603 (2012).

24. *See* Susan D. Franck, *The Nature and Enforcement of Investor Rights Under Investment Treaties: Do Investment Treaties Have a Bright Future,* 12 U.C. DAVIS J. INT'L L. & POL'Y 47, 48 (2005); David Schneiderman, *Investment Rules and the New Constitutionalism,* 25 LAW & SOC. INQUIRY 757, 767 (2000).

25. *See* Funnekotter v. Republic of Zim., ICSID Case No. ARB/05/6, Award (Apr. 22, 2009), *available at* http://italaw.com/documents/ZimbabweAward.pdf (deciding claim against Zimbabwe for repossession of land from white farmers); S.D. Myers, Inc. v. Canada, United Nations Commission on International Trade Law [UNCITRAL], Partial Award (Nov. 13, 2000), *available at* http://www.italaw.com/sites/default/files/case-documents/ita0747.pdf (adjudicating dispute involving claim that foreigner was disparately impacted by domestic environmental regulation of PCBs); *see also* Jason Webb Yackee, Pacta Sunt Servanda *and State Promises to Foreign Investors Before Bilateral Investment Treaties: Myth and Reality,* 32 FORDHAM INT'L L.J. 1550, 1555–57 (2009) (describing the South Africa and Argentina cases); Julie A. Maupin, *Differentiating Among International Investment Disputes, in* THE FOUNDATIONS OF INTERNATIONAL INVESTMENT LAW: BRINGING THEORY INTO PRACTICE 467 (Zachary Douglas et al. eds., 2014) (discussing multiple cases using socio-legal, territorial, and political dimensions).

26. *See, e.g.,* Ross P. Buckley & Paul Blyschak, *Guarding the Open Door: Non-Party Participation Before the International Centre for Settlement of Investment Disputes,* 22 BANK. & FIN. L. REV. 353, 366 (2007); *see also* Michael Waibel, *Opening Pandora's Box: Sovereign Bonds in International Arbitration,* 101 AM. J. INT'L L. 711, 712–13 (2007).

27. *See generally* ALVAREZ, *supra* note 11 (compiling cases).

28. Philip Morris Brands Sàrl v. Oriental Republic of Uruguay, ICSID Case No. ARB/10/7, Decision on Jurisdiction (July 2, 2013), *available at* http://www.italaw.com/sites/default/files/case-documents/italaw1531.pdf; Philip Morris Asia Ltd. v. Australia, UNCITRAL, Notice of Arbitration

As a matter of arbitration mechanics, if a foreign investor believes a host state has violated its substantive treaty rights, IIAs permit direct redress against the host state through the treaty's dispute resolution mechanism. This is the genesis of ITA, as IIAs permit investors to resolve their disputes directly and finally through arbitration. A cause of action under an IIA generally involves a foreign investor asserting that the host state's conduct violated the treaty and damaged the investment. If the dispute is not otherwise resolved through negotiation or mediation, the investor requires the state to arbitrate. While IIAs vary, investors can generally elect to arbitrate before: (1) an *ad hoc* tribunal organized under the United Nations Commission on International Trade Law (UNCITRAL) Arbitration Rules; (2) the Stockholm Chamber of Commerce (SCC); or (3) a tribunal organized through the World Bank's ICSID.[29] Despite some variations, these rules offer the procedural backbone for arbitration that operates as an individualized code of civil procedure subject to modifications by mutual agreement; they particularize baselines for the filing of pleadings, the selection and challenge of arbitrators, challenges to jurisdiction, the gathering and assessment of evidence, and the writing of awards.[30]

Arbitration mechanics are relatively straightforward. After complying with the necessary jurisdictional prerequisites, investors initiate arbitration by submitting a request for arbitration to one of the forums permitted by the treaty. Thereafter, the process of selecting a tribunal begins. Each party can generally select one arbitrator, and then either an arbitral institution or the two co-arbitrators typically appoint a presiding arbitrator or chair.[31] After signing statements of independence and impartiality, panels of three arbitrators resolve investment disputes[32] in accordance with the applicable law,[33] which generally includes the procedural rules, the applicable

---

(Nov. 21, 2011), *available at* http://www.italaw.com/sites/default/files/case-documents/ita0665.pdf.

29. There is some evidence that IIAs permit investors to arbitrate at the International Chamber of Commerce (ICC). *See* Jason Webb Yackee, *Bilateral Investment Treaties, Credible Commitment, and the Rule of (International) Law: Do BITs Promote Foreign Direct Investment?*, 42 LAW & SOC'Y REV. 805, 808, 812 (2008); UNCTAD, *IIA Issues Note No. 1: Latest Developments in Investor-State Dispute Settlement*, at 2, UNCTAD Doc. UNCTAD/WEB/DIAE/IA/2012/10 (2012) [*hereinafter* UNCTAD, *IIA Issues Note 2012*], *available at* http://unctad.org/en/PublicationsLibrary/webdiaeia2012d10_en.pdf (indicating only seven known ICC cases).

30. Chiara Giorgetti, *Who Decides Who Decides in International Investment Arbitration?*, 35 U. PA. J. INT'L L. 431, 438–54 (2013).

31. GARY B. BORN, INTERNATIONAL COMMERCIAL ARBITRATION 1696 (2d ed. 2014); Franck, *Legitimacy Crisis, supra* note 5, at 1543–44.

32. *See* RUDOLF DOLZER & MARGRETE STEVENS, BILATERAL INVESTMENT TREATIES 124 (1995); Susan D. Franck, *Empirically Evaluating Claims About Investment Treaty Arbitration*, 86 N.C. L. REV. 1, 77 (2007) [hereinafter Franck, *Empirically Evaluating*] (stating that "out of 102 awards, 100 awards had three-member tribunals, and sole arbitrators rendered two awards").

33. Toby Landau, *Report, Composition and Establishment of the Tribunal: Articles 14 to 36*, 9 AM. REV. INT'L ARB. 45, 52–53 (1998); Andreas F. Lowenfeld, *The Party-Appointed Arbitrator in International*

substantive provisions of the IIA, and other relevant principles of international law.[34]

Parties then marshal their facts and legal arguments to address different phases of the dispute, namely, jurisdiction, merits, quantum, and costs. The investor must first establish that it meets the three primary jurisdictional thresholds: (1) a qualifying investor, and (2) a qualifying investment that is (3) brought under an enforceable treaty within a proper time frame. If those prerequisites are not established, the case terminates and the state is not subject to liability under the IIA. Otherwise, the dispute continues. The merits phase involves a tribunal's determination of whether the respondent breached the treaty's substantive obligations and the specific protections the host state promised to foreign investors. If there is no substantive breach, the case terminates; otherwise, the dispute continues. At the quantum phase, the parties establish the value of the substantive treaty breach. Cost awards and assessments can be made at any point in the course of the lifetime of the dispute; and ultimately, although several awards or interim decisions on these core areas can occur, the tribunal renders a final award.[35]

The final doctrinal foundation relates to the international treaty that provides for recognition and enforcement of underlying awards. The two primary treaties permitting enforcement of arbitral awards are the ICSID Convention and the New York Convention. Although there are slight differences in the methodology, both treaties generally provide narrow bases for challenges to the awards and permit final awards to be enforceable worldwide.[36]

II. THE DEBATE ABOUT ITA

Debates related to the utility of ITA suggest that, thus far, its perceived success has exhibited variation akin to swings of a pendulum.[37] During its early phase, many commentators gave ITA "overwhelming praise" for its

---

Controversies: Some Reflections, 30 TEX. INT'L L.J. 59, 65–66 (1995); Claudia T. Salomon, Commentary, *Selecting an International Arbitrator: Five Factors to Consider*, 17 MEALEY'S INT'L ARB. REP. 1, 2–3 (2002).

34. Yas Banifatemi, *Mapping the Future of Investment Treaty Arbitration as a System of Law*, 103 AM. SOC'Y INT'L L. PROC. 323, 323–24 (2009) (summary of panel remarks); Andrea K. Bjorklund, *Mandatory Rules of Law and Investment Arbitration*, 18 AM. REV. INT'L ARB. 175 (2007).

35. *See, e.g.*, LUCY REED ET AL., GUIDE TO ICSID ARBITRATION (2d ed. 2010); ANDREW NEWCOMBE & LLUÍS PARADELL, LAW AND PRACTICE OF INVESTMENT TREATIES: STANDARDS OF TREATMENT (2009); *see also* Susan D. Franck, *Rationalizing Costs in Investment Treaty Arbitration*, 88 WASH U. L. REV. 769 (2011) (discussing costs awards).

36. *See* Franck, *Legitimacy Crisis, supra* note 5, at 1539–55.

37. *See generally* Sergio Puig, *Emergence & Dynamism in International Organizations: ICSID, Investor-State Arbitration & International Investment Law*, 44 GEO. J. INT'L L. 531 (2012) (discussing historical shifts in ITA).

"unmitigated" success.[38] Supportive commentary came from international law luminaries such as Judge Stephen M. Schwebel, who has identified the benefits of ITA for both capital-importing and exporting states.[39] This optimism may have been a byproduct of an increase in the foreign investment flows and in the number of treaties with a small number of disputes. In this period, the benefits seemed tangible and the risks appeared minimal. Yet cognitive illusions may have contributed to this perception. For example, the primacy effect may mean that the initial cases left a lasting impression that formed the basis for conscious or subconscious comparisons;[40] heuristics may facilitate initial cases being easier to recollect, in turn more readily discussed, more widely available, and ultimately perceived as representative of the larger population.[41] Early adjudication under the North American Free Trade Agreement (NAFTA) may have been viewed as favorable when contrasted with somewhat contemporaneous or prominent ICJ litigation. A few years before NAFTA, for instance, the United States was unable to establish an expropriation claim on behalf of investors in *ELSI*.[42] Unlike the nearly six decades required to secure full compliance with the ICJ damage award in *Corfu Channel*,[43] ITA cases — like *Metalclad*,[44] where U.S. investors received a prompt payment from Mexico after Mexico unsuccessfully challenged

38. David P. Riesenberg, Note, *Fee Shifting in Investor-State Arbitration: Doctrine and Policy Justifying Application of the English Rule*, 60 DUKE L.J. 977, 985 (2010) (discussing the satisfaction and praise for the existing ITA system and citing Jeswald Salacuse, Stephen Schwebel, Thomas Waelde, Ian Laird, and Joel Beauvais, who have extolled the virtues of ITA).

39. Stephen M. Schwebel, *The Overwhelming Merits of Bilateral Investment Treaties*, 32 SUFFOLK TRANSNAT'L L. REV. 263 (2008).

40. For example, flashbulb memories highlight detailed and emotionally vivid memories that seem fresh, precise, and easy to recollect. They can, however, be inaccurate. Should experience with ITA generate a flashbulb memory, this could theoretically facilitate inaccurate impressions and perceptions. CHRISTOPHER CHABRIS & DANIEL SIMONS, THE INVISIBLE GORILLA: AND OTHER WAYS OUR INTUITIONS DECEIVE US 68–78 (2010).

41. *See* FRANCK, MYTHS & REALITIES, *supra* note 14 (discussing heuristics including primacy, availability, and representativeness); DANIEL KAHNEMAN, THINKING, FAST AND SLOW (2011) (discussing similar heuristics); *see also* Timur Kuran & Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 STAN. L. REV. 683, 683 (1999) (identifying how the availability heuristic can "trigger[ ] a chain reaction that gives [a] perception [of] increasing plausibility through its rising availability in public discourse").

42. Elettronica Sicula S.p.A. (ELSI) (U.S. v. It.), 1989 I.C.J. 15 (July 20).

43. *Compare* Corfu Channel (U.K. v. Alb.), Judgment, 1949 I.C.J. 4 (Apr. 9) (awarding the United Kingdom for damages to two ships in 1946), *with* David J. Bederman, *Jurisprudence of the Foreign Claims Settlement Commission: Albania Claims*, 106 AM. J. INT'L L. 271, 272 n.11 (2012) (noting Albania paid the ICJ judgment in October 1996); *see also* Barry E. Carter et al., *Comparative Analysis of International Dispute Resolution Institutions*, 85 AM. SOC'Y INT'L L. PROC. 64, 64 (1991) (describing how prior to payment "over forty years after the decision in the *Corfu Channel* case, Albania has yet to pay Great Britain").

44. Metalclad v. United Mexican States, ICSID Case No. ARB(AF)/97/1, Award of the Tribunal (Aug. 30, 2000), 5 ICSID Rep. 212 (2002).

the award in a Canadian court[45] — were comparatively attractive. Some commentators continue to praise ITA and suggest that "[s]tates and investors both stand to benefit from international arbitration."[46]

Despite the initial wave of satisfaction, perception has shifted: Some commentary expresses discontent with ITA. Concerns come from various quarters, including international organizations, non-governmental organizations (NGOs), academics, and states. UNCTAD's World Investment Report 2012, for example, discussed "perceived shortcomings" of the system, emphasizing opportunities for

> reigning in the growing number of [ITA] cases, fostering the legitimacy and increasing the transparency of [ITA] proceedings, dealing with inconsistent readings of key provisions in IIAs and poor treaty interpretation, improving the impartiality and quality of arbitrators, reducing the length and costs of proceedings, assisting developing countries in handling [ITA] cases, and addressing overall concerns about the functioning of the system.[47]

Others suggest ITA involves considerable time and costs and that cases "run, on average, several years and entail large costs for both claimants and respondent States,"[48] which presumably has a disparate impact on developing states with fewer fiscal resources.

Meanwhile, NGOs generated controversy by publishing reports with tantalizing headlines — sometimes with minimal support or unbalanced representations — suggesting that ITA amounts to "[p]rofiting from [i]njustice," since "agreeing to arbitration [means] states have indeed accepted to be sued by the devil in hell."[49] Others have suggested ITA is a "legal monster."[50] Against this backdrop, South Korea — which has over

---

45. John Nagel, *Mexico Pays $16 Million to Metalclad, Ending First NAFTA Chapter 11 Dispute*, 24 Int'l Env. Rep. (BNA) 997 (Nov. 7, 2001); *see also* Terri L. Lilley, *Keeping NAFTA "Green" for Investors and the Environment*, 75 S. CAL. L. REV. 727, 730 n.13 (2002).

46. Joshua B. Simmons, *Valuation in Investor-State Arbitration: Toward a More Exact Science*, 30 BERKELEY J. INT'L L. 196, 202 (2012); *see also* Brower & Blanchard, *supra* note 5; Sebastian Perry, *Supporters of Investment Arbitration Launch EU Think Tank*, GLOBAL ARB. REV., July 7, 2014, http://globalarbitrationreview.com/news/article/32780/supporters-investment-arbitration-launch-eu-think-tank/.

47. UNCTAD, WIR 2012, *supra* note 4, at xxi, 88; *see also* UNCTAD, WIR 2013, *supra* note 1, at 110–17.

48. Kyla Tienhaara, *Third Party Participation in Investment-Environment Disputes: Recent Developments*, 16 REVIEW EUR. COMMUNITY & INT'L ENVTL. L. 230, 240 (2007).

49. EBERHARDT & OLIVET, *supra* note 5, at 3, 11; *see also* George Monbiot, *This Transatlantic Trade Deal Is a Full-Frontal Assault on Democracy*, GUARDIAN: COMMENT IS FREE (Nov. 4, 2013, 3:31 PM), http://www.theguardian.com/commentisfree/2013/nov/04/us-trade-deal-full-frontal-assault-on-democracy (suggesting ITA is a "monstrous assault").

50. *See* Mahnaz Malik, *The Legal Monster that Lets Companies Sue Countries*, GUARDIAN: COMMENT IS   FREE   (Nov.   4,   2011,   2:23   PM), http://www.theguardian.com/commentisfree/2011/nov/04/bilateral-investment-treaties; EBERHARDT & OLIVET, *supra* note 5; *Discover the Dark Side of Investment*, TRANSNATIONAL INST.

seventy IIAs with ITA provisions and no adverse awards — experienced dramatic protests in November 2011 that even included physical fights and the use of tear gas by legislators within the South Korean parliament.[51] One of the key issues of concern was, oddly, not the substantive provisions of the Korea-U.S. Free Trade Agreement, but rather, the reciprocal right of investors to sue states in arbitration.[52]

Some academics entered the fray in 2010 and issued a Public Statement. They claim that ITA

> is not a fair, independent, and balanced method for the resolution of investment disputes and therefore should not be relied on for this purpose. There is a strong moral as well as policy case for governments to withdraw from investment treaties and to oppose investor-state arbitration, including by refusal to pay arbitration awards [rendered] against them . . . .[53]

Another group of scholars opposing the creation of ITA in the TPP, the multilateral free trade agreement among a dozen Asia-Pacific states, argues that ITA should be abandoned to "reassert[ ] the integrity of our domestic legal processes."[54] The TTIP, involving a multilateral trade agreement with the United States and Europe, is similarly contentious.[55]

Some states have voted with their feet. Russia has withdrawn from the Energy Charter Treaty after an investor established jurisdiction over disputes related to the dissolution of Yukos Oil, potentially subjecting Russia to over US$50 billion in damages.[56] Bolivia, Ecuador, and

---

(June 27, 2012), http://www.tni.org/article/discover-dark-side-investment; *A Response to the Critics of "Profiting From Injustice"*, CORPORATE EUR. OBSERVATORY (Jan. 4, 2013), http://corporateeurope.org/blog/response-critics-profiting-injustice.

51. *South Korea Passes U.S. Free-Trade Agreement, Lawmaker Sets off Tear Gas Canister in Protest*, FOXNEWS.COM (Nov. 22, 2011) [hereinafter *South Korea Passes U.S. Free-Trade Agreement*], http://www.foxnews.com/world/2011/11/22/south-korea-passes-us-free-trade-agreement-lawmaker-sets-off-tear-gas-canister/; Alison Ross, *Arbitration Clause Sparks Protests in Korea*, GLOBAL ARB. REV., Nov. 29, 2011, http://globalarbitrationreview.com/b/29995/.

52. *See South Korea Passes U.S. Free-Trade Agreement*, *supra* note 51.

53. Gus Van Harten et al., *Public Statement on the International Investment Regime*, ¶ 8, Aug. 31, 2010, *available at* http://alainet.org/active/40578&lang=es; *see also id.* ¶ 14 (suggesting that investment treaty disputes be resolved by national court judges). *See generally* GUS VAN HARTEN, INVESTMENT TREATY ARBITRATION AND PUBLIC LAW 152–85 (2007) (stating ITA is "bias[ed] in favour of allowing claims and awarding damages against governments" and generally discussing aspects of this assertion). *But see* William W. Park, *Arbitrator Integrity: The Transient and the Permanent*, 46 SAN DIEGO L. REV. 629, 658–59 (2009) (identifying concerns about perceived partiality of having national court judges adjudicate disputes involving the state).

54. Elizabeth A. Evatt et al., *An Open Letter from Lawyers to the Negotiators of the Trans-Pacific Partnership Urging the Rejection of Investor-State Dispute Settlement* (May 8, 2012), *available at* http://tpplegal.wordpress.com/open-letter/.

55. *See* Jos Dings & Pieter de Pous, Letter to the Editor, *As It Stands, the TTIP Could Threaten Democracy*, FIN. TIMES, Feb. 18, 2014, http://www.ft.com/intl/cms/s/0/d49b7bb6-94de-11e3-af71-00144feab7de.html.

56. *See generally* Emmanuel Gaillard, Letter to the Editor, *Russia Cannot Walk away from its Legal*

Venezuela have withdrawn from the ICSID Convention.[57] The President of Ecuador has claimed ICSID "signifies colonialism, slavery with respect to transnationals, with respect to Washington, with respect to the World Bank."[58] In October 2012, Ecuador's Minister of Foreign Affairs, Trade and Integration similarly called an ICSID award "unjust, illegal, illegitimate and absurd."[59]

Other dissatisfied stakeholders have encouraged the "reemergence of the 'toothless' investment treaty" which, like older treaties of Friendship, Navigation, and Commerce, provides substantive rights but no forum for dispute resolution.[60] Australia claimed that efficiency reasons warranted rejecting ITA in future investment treaties. Having been threatened with arbitration by Philip Morris for its plain packaging legislation for cigarettes, Australia expressed concerns about costly and fractious disputes and the net value of treaties as "the economic value of Australia's preferential [IIAs] have been oversold."[61]

Yet not all states agree. Some have made different decisions related to the utility of ITA. In April 2012, after a contentious consultation process, the United States crafted a model IIA that included streamlined substantive and procedural rights and retained ITA.[62] China also changed its approach to ITA by *broadening* the scope of arbitration rights to permit

---

Obligations, FIN. TIMES, Aug. 18, 2009, http://www.ft.com/intl/cms/s/0/c63d918a-8b8d-11de-9f50-00144feabdc0.html#axzz3JLTCdwf8 (identifying Russia's announced withdrawal of the Energy Charter Treaty).

57. *See* Diane Marie Wick, *The Counter-Productivity of ICSID Denunciation and Proposals for Change*, 11 J. INT'L BUS. & L. 239, 241–43 (2012) (identifying the denunciations of ICSID by Venezuela, Ecuador, and Bolivia); Michael Waibel et al., *The Backlash Against Investment Arbitration: Perceptions and Reality, in* THE BACKLASH, *supra* note 7, at xxxvii (discussing the fallout of the ITA backlash).

58. Trakman, *supra* note 23, at 604 (quoting *ICSID in Crisis: Straight-Jacket or Investment Protection?*, BRETTON WOODS PROJECT (July 10, 2009), http://www.brettonwoodsproject.org/art-564878 (internal quotation marks omitted)); *see also* Charles N. Brower, *The Evolution of the International Judiciary: Denationalization Through Jurisdictional Fragmentation*, 103 AM. SOC'Y INT'L L. PROC. 171, 183 n.59 (2009) (providing same quote and stating that President "Correa referred to ICSID as an 'atrocity'" and that "withdrawal from [ICSID] was necessary for 'the liberation of our countries'").

59. Press Release, Ministry of Foreign Affairs & Human Mobility, Ecuador Will Require Invalidation of Fault of the International Center for Settlement of Investment Disputes (ICSID) in Favor of Occidental (Oct. 10, 2012) (Ecuador), *available at* http://cancilleria.gob.ec/ecuador-will-require-invalidation-of-fault-of-the-international-center-for-settlement-of-investment-disputes-icsid-in-favor-of-western-minister-ricardo-patino/?lang=en.

60. Riesenberg, *supra* note 38, at 980, 987; *see also* UNCTAD, WIR 2012, *supra* note 4, at 139–40; Trakman, *supra* note 23.

61. Trakman, *supra* note 23, at 649–51 (discussing and quoting the Australian Productivity Commission); *see also* DEP'T OF FOREIGN AFFAIRS & TRADE, GILLARD GOVERNMENT TRADE POLICY STATEMENT: TRADING OUR WAY TO MORE JOBS AND PROSPERITY (2011) (Austl.), *available at* http://www.acci.asn.au/getattachment/b9d3cfae-fc0c-4c2a-a3df-3f58228daf6d/Gillard-Government-Trade-Policy-Statement.aspx.

62. *See* Press Release, U.S. Dep't of State, Model Bilateral Investment Treaty (Apr. 20, 2012), *available at* http://www.state.gov/r/pa/prs/ps/2012/04/188199.htm.

arbitration on both the merits and damages of treaty claims.[63] Canada, which has had its investors bring claims and has been a respondent, retained its ITA mechanism when it revised its model IIA in 2003 in light of its experience in NAFTA arbitration.[64] The May 2012 trilateral investment treaty among China, Japan, and Korea retained ITA for dispute settlement.[65] Even Australia, after a national election leading to a change in government, included ITA in a treaty negotiated with South Korea.[66] Meanwhile, states continue to renegotiate investment treaties, and ITA appears to be a viable dispute resolution option in the ongoing TPP and TTIP negotiations.[67]

Part of the debate about ITA derives from concerns about the development dimension and whether ITA exhibits a bias against developing states. The President of Bolivia, for example, has gone on record suggesting that no country, except the United States, ever wins — even though that assertion is empirically refutable.[68] Likewise, some organizations focus on the number of claims brought against developing states but do not necessarily provide holistic analyses of outcomes as a function of development status.[69] Meanwhile, others take remarks out of context without providing alternative analyses that demonstrate the

---

63. *See* Mark Feldman, Remarks, *in China-Africa Investment Treaties and Dispute Settlement: A Piece of the Multipolar Puzzle*, 107 AM. SOC'Y INT'L L. PROC. 225, 231, 231–32 (2013) (discussing the expansion of aspects of China's investment treaty provisions); *see also* NORAH GALLAGHER & WENHUA SHAN, CHINESE INVESTMENT TREATIES: POLICIES AND PRACTICE (2009).

64. *See Canada's Foreign Investment Promotion and Protection Agreements (FIPAs)*, FOREIGN AFFAIRS, TRADE & DEV. CAN., http://www.international.gc.ca/trade-agreements-accords-commerciaux/agr-acc/fipa-apie/fipa-apie.aspx?lang=eng (last modified Nov. 4, 2013); *Canada's FIPA Program: Its Purpose, Objective and Content*, FOREIGN AFFAIRS, TRADE & DEV. CAN., http://www.international.gc.ca/trade-agreements-accords-commerciaux/agr-acc/fipa-apie/fipa-purpose.aspx?lang=en (last modified May 27, 2014).

65. Agreement Among the Government of Japan, the Government of the Republic of Korea, and the Government of the People's Republic of China for the Promotion, Facilitation and Protection of Investment, May 13, 2012, *available at* http://www.mofa.go.jp/announce/announce/2012/5/pdfs/0513_01_01.pdf.

66. Press Release, Minister for Trade & Inv., Australia Concludes FTA Negotiations with the Republic of Korea (Dec. 5, 2013) (Austl.), *available at* http://trademinister.gov.au/releases/2013/ar_mr_131205.html; Free Trade Agreement Between the Government of Australia and the Government of the Republic of Korea arts. 11.15–11.28, [2014] ATNIF 4 (Austl.) (providing the text of the recently negotiated Korea-Australia investment agreement that provides for ITA).

67. *EU Draft Text on ISDS Contains Similarities, Differences to U.S. Approach*, INSIDE U.S. TRADE'S WORLD TRADE ONLINE (Apr. 5, 2013), http://insidetrade.com/Inside-Trade-General/Public-Content-World-Trade-Online/eu-draft-text-on-isds-contains-similarities-differences-to-us-approach/menu-id-896.html (suggesting the retention of ITA but possible variations in enforcement mechanisms).

68. Franck, *Development and Outcomes, supra* note 6, at 436.

69. UNCTAD, *ITA Issues Note 2012, supra* note 29; *see also* UNCTAD, *ITA Issues Note 2014, supra* note 4. *But see* Gallagher & Shrestha, *supra* note 6 (providing a more holistic explanation of potential bias of ITA against developing states). The gap may derive from limited institutional resources in international organizations.

holistically harmful impact on developing states,[70] or they may have a selective focus.[71] Nevertheless, given the potential impact on ITA's legitimacy and its value as a neutral rule of law mechanism, arguable problems related to development dimensions and ITA outcomes are a serious matter worthy of focused and careful attention.

Earlier research was unable to establish a reliable link between the outcomes of ITA disputes and respondent states' developmental background.[72] Even when re-defining development — using definitions spontaneously generated by critics — those null results were robust across all models.[73]

Concerns about the latent effect of development should not be dismissed, however. It is possible that, over time, analyses could reveal systematically biased outcomes. Yet, it is also possible that ITA outcomes are a function of variables other than development — such as internal domestic politics or some derivative variable.[74] Political science literature supports the plausibility of this inference. Robust scholarship has identified, for example, links between domestic political institutions and investment stability and focuses on how democratic protections and other

---

70. *See, e.g.*, DANIEL IKENSON, CATO INST., FREE TRADE BULLETIN NO. 57, A COMPROMISE TO ADVANCE THE TRADE AGENDA: PURGE NEGOTIATIONS OF INVESTOR-STATE DISPUTE SETTLEMENT (2014), *available at* http://object.cato.org/sites/cato.org/files/pubs/pdf/ftb57.pdf.

71. *See* Gus Van Harten, *Arbitrator Behaviour in Asymmetrical Adjudication: An Empirical Study of Investment Treaty Arbitration*, 50 OSGOODE HALL L.J. 211 (2012). For example, Professor Van Harten's intriguing analysis of investment treaty outcomes that purports to find a pro-investor bias evaluates only jurisdictional issues. Beyond the failure to control for awards that generate multiple issues — and thus double or triple counting cases where the same tribunal addresses multiple issues — he fails to account for ultimate outcomes of cases. Similarly, a recent study focusing primarily on "merit awards" asserts that states win nearly 80% of reported disputes and consolidates types of outcomes. CREDIBILITY CONSULTING LLC, STUDY OF DAMAGES IN INTERNATIONAL CENTER FOR THE SETTLEMENT   OF   INVESTMENT   DISPUTES   CASES   1   (2014),   *available   at* http://www.credibilityconsulting.com/wp-content/uploads/2014/06/Credibility-ICSID-Damages-Study-June-2014-1.pdf. Holistic research demonstrates that while the majority of cases ultimately generated outcomes favorable to states, those state-favorable outcomes occurred less frequently at the jurisdictional phase and more frequently in merits and/or damages phases. *See* FRANCK, MYTHS & REALITIES, *supra* note 14; *see also* Franck, *Empirically Evaluating, supra* note 32, at 52–53 (providing similar results in an earlier generation of research).

72. *See* Franck, *Development and Outcomes, supra* note 6.

73. Susan D. Franck et al., *Through the Looking Glass: Understanding Social Science Norms for Analyzing International Investment Law*, *in* YEARBOOK OF INTERNATIONAL INVESTMENT LAW & POLICY 2010–2011, at 883 (Karl P. Sauvant ed., 2012); *see also* Franck, *ICSID Effect, supra* note 14, at 875–93.

74. As the topic of the Symposium panel was the internal politics of host states and international investment, this Article focuses on one readily measurable indicator of politics, namely scores related to the host state's democracy. Politics and democracy need not be the focal points of analysis. Other variables might include the lawyers involved in the process, investor identity, arbitrator background, the nature of contested government activity, or some other permutation of factors; these factors could be potentially more worthwhile than a focus on politics alone. *See, e.g.*, FRANCK, MYTHS & REALITIES, *supra* note 14 (conducting multiple analyses on variables associated with outcomes); Susan D. Franck & Lindsey E. Wylie, *Demystifying Investment Treaty Arbitration* (Jan. 7, 2015) (unpublished manuscript) (on file with author) (same).

constraints on policymaking discretion are associated with lower economic risk and higher investment levels.[75] Jeremy Caddel and Nathan M. Jensen recently identified that executive branch action (rather than legislative or judicial) spawned the vast majority of ITA disputes.[76]

Yet the role of politics in investment activities — to say nothing of investment dispute outcomes — can be complicated.[77] Some scholars suggest that investment levels are better under authoritarian regimes[78] or that transitions to democratic institutions can detrimentally impact investment levels.[79] This research suggests that democracy might inhibit development or otherwise may be negatively related to positive investment outcomes. Others argue the contrary and explain that democratic institutions attract higher investment flows.[80] Similarly, recognizing the potential for conflating effects of political regimes and political risk, Jensen found that democratic regimes experienced lower political risk, as witnessed by the lower pricing of political risk insurance premiums.[81] Still others suggest that the degree of democratic government is unrelated to investment levels, and the real phenomenon is that institutions within democracies impact investment levels and encourage positive investment outcomes.[82]

When there is a link between political risk and the investment choices of foreign entities, the question is whether realized state liability is likewise

75. Nathan M. Jensen et al., *Unbundling the Relationship Between Authoritarian Legislatures and Political Risk*, 44 BRIT. J. POL. SCI. 655 (2013).

76. Jeremy Caddel & Nathan M. Jensen, *Which Host Government Actors Are Most Involved in Disputes with Foreign Investors?*, at 1, 2 (Vale Columbia Ctr. on Sustainable Int'l Inv., Columbia FDI Perspectives No. 120, 2014).

77. *See, e.g.*, ADAM PRZEWORSKI ET AL., DEMOCRACY AND DEVELOPMENT: POLITICAL INSTITUTIONS AND WELL-BEING IN THE WORLD, 1950–1990, at 146–53 (2000).

78. John R. Oneal, *The Affinity of Foreign Investors for Authoritarian Regimes*, 47 POL. RES. Q. 565 (1994).

79. Adam L. Resnick, *Investors, Turbulence and Transition: Democratic Transition and Foreign Direct Investment in Nineteen Developing Countries*, 27 INT'L INTERACTIONS 381 (2001).

80. *See* Quan Li & Adam Resnick, *Reversal of Fortunes: Democratic Institutions and Foreign Direct Investment Inflows to Developing Countries*, 57 INT'L ORG. 175 (2003) (finding that democracy decreases FDI flows, but only after controlling for the rule of law, where enhancing aspects can increase investment flows but other aspects, like regular policy change or antitrust laws, reduce investment flows); *see also* WITOLD JERZY HENISZ, POLITICS AND INTERNATIONAL INVESTMENT: MEASURING RISKS AND PROTECTING PROFITS (2002); NATHAN M. JENSEN, NATION-STATES AND THE MULTINATIONAL CORPORATION: A POLITICAL ECONOMY OF FOREIGN DIRECT INVESTMENT (2006); Matthias Busse & Carsten Hefeker, *Political Risk, Institutions and Foreign Direct Investment*, 23 EUR. J. POL. ECON. 397 (2007); Phillipp Harms & Heinrich W. Ursprung, *Do Civil and Political Repression Really Boost Foreign Direct Investments?*, 40 ECON. INQUIRY 651 (2002).

81. Nathan Jensen, *Political Risk, Democratic Institutions, and Foreign Direct Investment*, 70 J. POL. 1040 (2008).

82. Glen Biglaiser & Joseph L. Staats, *Do Political Institutions Affect Foreign Direct Investment? A Survey of U.S. Corporations in Latin America*, 63 POL. RES. Q. 508 (2010); Tim Büthe & Helen V. Milner, *The Politics of Foreign Direct Investment into Developing Countries: Increasing FDI Through International Trade Agreements?*, 52 AM. J. POL. SCI. 741 (2008).

driven by political factors or something else, such as tribunal animus against developing nations or a preference for relieving developed states of liability.

Ultimately, the debate about ITA is a live issue with political and economic ramifications. Claims of improper systemic bias are worth addressing, as outcomes untethered to legal merits and based purely upon a state's development levels provide a legitimate basis for critiquing dispute resolution. Yet, should stakeholders wish to have a debate based upon evidence, a balanced understanding of ITA's functionality is necessary. This requires data, gathered using sound *ex ante* research protocols that derive from established social science practices to inform the debate and the normative choices on the appropriate form of dispute resolution.

III. METHODOLOGICAL BACKGROUND

This research used archival data[83] to explore the potential links between the democracy levels of respondent states, the development levels of the respondent states, and ITA outcomes. The objective of the quantitative analysis was to explore potential reliable statistical relationships, identifying whether ITA favors respondents from the developed or developing world and how a respondent state's political institutions might (or might not) serve as an explanatory variable for outcomes.

A. Unit of Analysis

The unit of analysis was ITA awards that were publicly available as of January 1, 2012. This included all public awards irrespective of the language rendered, which necessarily included awards rendered in English, Spanish, and French. The method for gathering publicly available awards was the same methodology used in previous research.[84] As there is no single agreed-upon repository for ITA awards, the primary source of the public awards derived from Professor Andrew Newcombe's website, ITAlaw.com.[85] This was supplemented and cross-checked using a number of resources, including the UNCTAD website, the ICSID website, and the *ICSID Review*, among others.[86]

---

83. The data derives from a larger research project. *See* FRANCK, MYTHS & REALITIES, *supra* note 14. This Article abbreviates the methodology and incorporates those provisions as if fully set forth herein. The data from the larger project will be made publicly available upon publication; in the interim, data from pre-2007 cases already is available. *See Generation 1 Dataset and Related Materials*, WASHINGTON & LEE SCH. OF LAW, http://law2.wlu.edu/faculty/page.asp?pageid=1185 (last accessed Jan. 22, 2015).

84. Franck, *Empirically Evaluating, supra* note 32, at 16–20.

85. *See About ita*, ITA, http://www.italaw.com/about (last visited Jan. 7, 2015).

86. *See Database of Investor-State Dispute Settlement (ISDS) (reduced version)*, UNCTAD, http://unctad.org/en/Pages/DIAE/ISDS.aspx (last visited Jan. 7, 2015) (limited UNCTAD database

International arbitration, depending in part on the applicable law, experiences variations in the formal doctrinal definition of "award" and the distinguishing types of tribunal activity from other aspects of arbitration proceedings. Awards — namely elements of cases that resolve dispositive key legal issues — under the ICSID Convention[87] or under the UNCITRAL Model Law[88] have specific but slightly different meanings.

As the research cuts across ITA disputes, irrespective of whether cases were rendered at ICSID under the ICSID Convention or the Additional Facility Rules or some other forum under the New York Convention, the research adopted a precise definition of "award." Namely, it defined an "award" as a tribunal's written document that made a binding decision on one of the four substantive phases of an arbitration or another decision that was otherwise dispositive, namely awards on: (1) jurisdiction; (2) merits; (3) damages; (4) allocation of costs; or (5) settlement agreements or other orders indicating a dismissal or discontinuance.[89] This means that other decisions that might generally provide information or provide insights into the larger dispute resolution process — including interim measures, decisions on confidentiality and transparency, place of arbitration, arbitrator challenges, tribunals' interpretive decisions, *ad hoc* Annulment Committees, and subsequent decisions by national courts — were not generally coded. Disputes with only a public dispute notice or request for arbitration were also omitted.

There are three elements that must be kept in mind when understanding the dataset, the unit of analysis, and derivative inferences.

First, the dataset is a time-bounded population of ITA disputes that were derived from public awards. This feature necessarily means that it is

---

available while full database undergoes redesign); *Cases*, INT'L CTR. FOR SETTLEMENT OF INV. DISPUTES   [ICSID], https://icsid.worldbank.org/apps/ICSIDWEB/cases/Pages/AdvancedSearch.aspx (last visited Jan. 7, 2015) (searchable database of ICSID awards). For print versions of ICSID cases, see, e.g., 16 ICSID Rep. (James Crawford & Joanna Gomula eds., 2012). Likewise, certain cases at the SCC are accessible through its website, http://english.chamber.se/, or the *Stockholm International Arbitration Review*. The availability of public awards was crosschecked against other sources, including websites requiring a subscription, like the APPLETON-ISR database on Westlaw, the Oxford University Press website, http://oxia.ouplaw.com/home/ic, and the Investor-State Law Guide, http://www.investorstatelawguide.com/. The author also consulted NAFTAclaims.com for cases arising under NAFTA.

87. ICSID Convention, *supra* note 17, art. 48 (defining an award); ICSID, *Rules of Procedure for Arbitration Proceedings (Arbitration Rules)*, *in* ICSID CONVENTION, REGULATIONS AND RULES, at 119–21, ICSID Doc. ICSID/15 (2006) (providing in rules 41 and 43–45 for preliminary objections related to jurisdiction and other orders for discontinuances, settlements, or party inaction).

88. UNCITRAL MODEL LAW ON INTERNATIONAL COMMERCIAL ARBITRATION arts. 17–17J, U.N. Sales No. E.08.V.4 (2008) (defining interim measures as awards capable of enforcement and defining other enforceable awards).

89. This methodology is similar to other research analyzing ICSID awards addressing jurisdiction, merits, settlements, or other orders for discontinuances. Daphna Kapeliuk, *Collegial Games: Analyzing the Effect of Panel Composition on Outcome in Investment Arbitration*, 31 REV. LITIG. 267, 302 (2012).

possible that the population has continued to evolve, and the external
validity of this research is arguably limited. For example, awards rendered
after January 2012 generate variance in the population. Yet, the analysis in
this Article inevitably provides a core baseline for future consideration. It
is possible for ITA to change over time, and replication is critical.

Second, private awards may differ meaningfully from public awards,
thereby limiting the value of inferences. The growth of the current dataset
identified a small number of cases that were previously private but
ultimately entered the public domain. There did not appear to be
meaningful differences between previously non-public and public awards,
and there was rough parity in parties, claims, arbitrators, and outcomes. In
any event, it is not clear that case selection bias skews the data. Irrespective
of whether investors or states were successful, both have incentives to
disclose awards. Investors, for example, may disclose awards because
shareholders may positively view the information that the investor won a
claim against a state and created greater commercial certainty. Likewise,
states may believe that notifying the public that they have won a case will
restore confidence in the government, will be of political utility, or will
promote the confidence of other foreigners investing in the country.

Third, other investment-related conflicts were not coded. These
conflicts related to either formalized investment treaty disputes without
reliable publicly available information or conflicts that never crystalized
into formal disputes.[90] As Professor Michael Reisman explains, given the
"very large" scope of international investment, the 80,000 multinational
enterprises and their 100,000 affiliates, and the possibility of claims under
2,700+ treaties, "the number of actual disputes going to arbitration seems
to be a miniscule fraction of the universe" of investment and investment

---

90. There are two categories of formalized disputes that were not analyzed. First, ITA disputes
with a dispute notice — like those listed on the ICSID website — may be in the pipeline but
constitute disputes about which there is public knowledge. Yet the inability to reliably assess case
information without a verifiable legal record leads to omission, as even where ICSID publicly
registers a case, it does not indicate whether the case involves a breach of treaty, domestic contract
law, or domestic investment law. Second, there could be latent, non-formalized investment disputes.
Investment-related problems arise every day, and the author is aware of no datasets measuring the
full scope of investment treaty conflict lurking beyond formal disputes. States may implement
Dispute Prevention Policies to eliminate or minimize conflicts leading to formal dispute notices, or
parties may use alternative dispute resolution strategies — like negotiation, mediation, or ombuds —
to address conflicts prior to an investor submitting a formal dispute notice. Similarly, after client
counseling and exploring the costs and benefits of dispute resolution options, (1) an investor might
abandon its claim or seek alternative commercial redress, or (2) a state may address a disputed
measure or otherwise redress state activity. *See* Washington and Lee University and UNCTAD Joint
Symposium on International Investment and Alternative Dispute Resolution, Mar. 29, 2010, *Investor–
State Disputes: Prevention and Alternatives to Arbitration II*, UNCTAD Doc.
UNCTAD/WEB/DIAE/IA/2010/8 (2011) [hereinafter *Investor-State Disputes*]. The dataset therefore
focuses on the tip of the iceberg of investment-treaty conflict, and the full range of unreported and
informal disputes is likely larger and unknowable.

conflict.[91] By focusing on public arbitration awards, the research necessarily omits a larger context. It is, however, necessary to narrow the field of inquiry to generate a careful, yet contained, analysis of the information that is both reliable and available.

B. Generation of the Dataset

The current dataset grew from earlier research from 2006 ("Generation 1" or "G1" research). The five additional years of data expanded the number of publicly available awards by approximately 250% and increased the number of cases by 300%.[92] Using a Codebook of more than 100 pages, multiple coders coded raw data from the award with an initial 94.7% rate of inter-coder agreement. All divergences were then resolved in inter-coder reliability meetings, and a mutually acceptable code was agreed. The hope was that the rigor of the coding and the use of common standards in the Codebook would ensure that the underlying data were reliable, would minimize the risk of error, and would enhance the value of the data and related analyses.

C. Methodology and Variables

One of the most challenging aspects of social science is measurement validity, namely, finding a measure to capture the nature of the social phenomenon in which a researcher is interested. "It may be the case that the variable being explained cannot be measured accurately, either because of data collection difficulties or because it is inherently unmeasurable and a proxy variable must be used in its stead."[93] No attempt at operationalizing a variable will be perfect "because there are an infinite number of physical characteristics that might be addressed in any definition"; so, rather than striving for unattainable perfection, "researchers try to produce definitions that are adequate to permit a replication of all important respects by another researcher."[94]

In compliance with social science protocols, this Article first identifies three variables that required transformation from raw coding, specifically,

---

91. W. Michael Reisman, *International Investment Arbitration and ADR: Married but Best Living Apart*, 24 ICSID REV.–FOREIGN INVESTMENT L.J. 185, 186 (2009).

92. Coding occurred in three different time periods. G1 data were collected from the beginning of ITA until June 1, 2006. Generation 2 ("G2") data were collected from public awards between June 1, 2006, and June 1, 2009. Generation 3 ("G3") data were collected from awards within the public domain between June 1, 2009, and January 1, 2012. G2 and G3 data collection also looked for awards that, although once private, had become public. There was always at least one coder involved in earlier coding. All coders received extensive training by coding a core set of materials and using the Codebook for usually around four weeks prior to coding the live data.

93. PETER KENNEDY, A GUIDE TO ECONOMETRICS 3 (5th ed. 2003).

94. MILDRED L. PATTEN, UNDERSTANDING RESEARCH METHODS: AN OVERVIEW OF THE ESSENTIALS 17 (7th ed. 2009) (emphasis omitted).

definitions of development, politics, and fiscal amounts. Then, it defines how "outcomes" were operationalized.

1. Development Status

Defining "Development Status" is a perpetual challenge.[95] There is no consistent legal definition of this concept, and development can mean different things to different people. For example, the World Trade Organization does not offer a precise measurement for development. Rather, it permits member states to self-define development,[96] which has caused inconsistency and confusion in areas such as international environmental law.[97] Without a predefined or exclusive measure for a legal phenomenon, it becomes necessary to use measures from "judgments made for entirely different purposes by *other researchers*,"[98] so as to minimize selection and self-serving biases.

Similar to earlier research,[99] this research used two measures for development from established indexes. Out of an abundance of caution, this research used a third measure designed by an expert and reputable international organization.

First, development was operationalized as a binary categorical variable — Organisation for Economic Co-Operation and Development (OECD) status — that derived from state membership in the OECD at the date of the award. OECD membership is generally, but not always, associated with higher levels of development, and therefore is a blunt proxy for development. States become OECD members through a complicated accession process designed to demonstrate a willingness to embrace a market economy, democratic pluralism, and respect for human

---

95. Marc L. Busch & Eric Reinhardt, *Developing Countries and General Agreement on Tariffs and Trade/World Trade Organization Dispute Settlement*, 37 J. WORLD TRADE 719, 723 n.4, 726 (2003) (analyzing development dimensions in GATT disputes and observing the difficulty in making distinctions between developed and developing states).

96. *See Who Are the Developing Countries in the WTO?*, WORLD TRADE ORG., http://www.wto.org/english/tratop_e/devel_e/d1who_e.htm (last visited Jan. 8, 2015) ("There are no WTO definitions of 'developed' and 'developing' countries. Members announce for themselves whether they are 'developed' or 'developing' countries."); Anu Bradford & Eric A. Posner, *Universal Exceptionalism in International Law*, 52 HARV. INT'L L.J. 1, 32 n.159 (2011) ("WTO rules do not contain a definition of a 'developing country.' Instead, states self-designate themselves as developed or developing countries as part of a political calculus."); *see also* Kristin Bohl, *Problems of Developing Country Access to WTO Dispute Settlement*, 9 CHI.-KENT J. INT'L & COMP. L. 130, 132–33 (2009) (acknowledging the elasticity of defining developing countries); Andrew D. Mitchell & Joanne Wallis, *Pacific Pause: The Rhetoric of Special & Differential Treatment, the Reality of WTO Accession*, 27 WIS. INT'L L.J. 663, 696–97 (2009) (categorizing developing countries).

97. Benjamin L. Liebman, *Autonomy Through Separation?: Environmental Law and the Basic Law of Hong Kong*, 39 HARV. INT'L L.J. 231, 261–62 (1998).

98. GARY KING ET AL., DESIGNING SOCIAL INQUIRY: SCIENTIFIC INFERENCE IN QUALITATIVE RESEARCH 157 (1994) (emphasis in original).

99. *See* Franck, *Development and Outcomes, supra* note 6; Franck, *ICSID Effect, supra* note 14.

rights.[100] Although OECD membership has been fairly stable,[101] the possibility of change was addressed by coding for OECD membership on the date of the award, which was the date for other critical variables (i.e., currency conversions and inflation adjustments).[102]

Second, development was also operationalized using a four-category variable — World Bank Status — that derived from a World Bank classification system grouping states as High Income, Upper-Middle Income, Lower-Middle Income, and Low Income.[103] The World Bank's main criterion for classifying economies is gross national income per capita.[104] As World Bank classifications can and did change over time, World Bank Status was coded at the award date.[105]

Third, development status was operationalized using a continuous variable — HDI Status — derived from the United Nations Development Programme's (UNDP) Human Development Index (HDI). Developed by Mahbub ul-Haq and Amartya Sen to shift conceptions of development, the HDI evaluates elements including life expectancy, education, and income. HDI is a continuous variable and ranges from 0.0 (undeveloped) to 1.0 (completely developed).[106] Reflecting that states' HDI scores could shift over time, HDI scores were coded using the date of the award.[107]

---

100. Jan Wouters & Sven Van Kerckhoven, *The OECD and the G20: An Ever Closer Relationship?*, 43 GEO. WASH. L. REV. 345 (2011).

101. *See List of OECD Member Countries – Ratification of the Convention on the OECD*, OECD, http://www.oecd.org/about/membersandpartners/list-oecd-member-countries.htm (last visited Nov. 12, 2014).

102. This meant, for example, that in the jurisdictional decision in *CSOB v. Slovak Republic* (1999) that terminated the ITA dispute, Slovakia was coded as "0" to reflect that it was not an OECD member; but in *Eureko v. Slovak Republic* (2010), Slovakia was coded as "1" to reflect its December 2000 accession to OECD membership.

103. The World Bank uses specific means to classify and codify information. *See Data: Country and Lending Groups*, WORLD BANK, http://data.worldbank.org/about/country-and-lending-groups (last visited Nov. 12, 2014); *Data: What Is the World Bank Atlas Method?*, WORLD BANK, https://datahelpdesk.worldbank.org/knowledgebase/articles/378832-what-is-the-world-bank-atlas-method (last visited Nov. 12, 2014).

104. *Data: How Does the World Bank Classify Countries?*, WORLD BANK, https://datahelpdesk.worldbank.org/knowledgebase/articles/378834-how-does-the-world-bank-classify-countries (last visited Jan. 8, 2015).

105. This meant, for example, that in *Lauder v. Czech Republic* (2001), the Czech Republic was coded as an Upper-Middle Income state; but in *Phoenix Action v. Czech Republic* (2009), the Czech Republic was coded as High Income. Similarly, in *EnCana v. Ecuador* (2004), Ecuador was coded as a Lower-Middle Income state; but in *Murphy v. Ecuador* (2010), Ecuador was coded as an Upper-Middle Income state.

106. *See generally Human Development Reports: About Human Development*, U.N. DEV. PROGRAMME [UNDP], http://hdr.undp.org/en/humandev (last visited Jan. 22, 2015) (describing the evolution of the HDI measure); *see also* UNDP, HUMAN DEVELOPMENT REPORT 2011: SUSTAINABILITY AND EQUALITY: A BETTER FUTURE FOR ALL (2011), *available at* http://hdr.undp.org/sites/default/files/reports/271/hdr_2011_en_complete.pdf (applying the measure). As the methodology for coding HDI changed in 2011 and was applied to all the data retroactively, previously published Human Development Reports were not used to code HDI levels. The research used data directly provided by Dr. Milorad Kovacevic, Chief Statistician at the Human

2. Politics and Democracy

A state's political structure involves more than evaluation of a state's formal name.[108] Although the scale has a potentially narrow definition of democracy, the dataset transformed a state's respondent political system using a preexisting scale commonly used by political scientists.[109]

Using the date of the award, the respondent state was ranked pursuant to the Polity IV index. Polity IV is designed to identify the political regime of a state. Polity contains information on countries with a population of more than 500,000 people. The index places states on a scale in which states with more authoritarian and autocratic regimes have lower Polity scores, and states with more democratic regimes have higher scores.[110]

3. Fiscal Amounts

Some raw data were transformed using either preexisting scales offered by independent organizations or established statistical protocols. First, using the date the award was rendered, amounts claimed and awarded were converted to a common currency of U.S. Dollars. Second, data were adjusted for inflation using the Consumer Price Index on a common date.[111] Third, fiscal data that exhibited unacceptable statistical skewing were transformed to reduce skewing, to ensure that the data matched the assumptions of the underlying statistical tests, and to enhance statistical conclusion validity.[112]

Development Report Office of the United Nations Development Programme. All of the scores Dr. Kovacevic provided used the updated 2011 methodology to reevaluate the historical and current rankings of states.

107. This meant, for example, in *Tradex Hellas v. Albania* (1996), Albania's HDI Status was 0.662; but in *Pantechniki v. Albania* (2009), Albania's HDI Status was 0.734. Likewise, in *Sedelmayer v. Russian Federation* (1998), Russia had an HDI Status of 0.677; but in *RosInvestCo v. Russian Federation* (2010), Russia had an HDI Status of 0.751. In *Loewen v. United States* (2001), the United States' HDI Status was 0.898, and it was similar (0.906) in *Glamis Gold v. United States* (2009).

108. Sara McLaughlin Mitchell & Paul F. Diehl, *Caution in What You Wish for: The Consequences of a Right to Democracy*, 48 STAN. J. INT'L L. 289, 291 (2012).

109. *Id.* at 291 ("The Polity IV Project [is] the most widely used scholarly database on democracy and authoritarianism . . . ."); *see also* David. S. Law & Mila Versteeg, *Sham Constitutions*, 101 CALIF. L. REV. 863, 926 (2013); Sergey Mityakov et al., *International Politics and Import Diversification*, 56 J.L. & ECON. 1091 (2013) (using the Polity IV dataset); Mila Versteeg, *Unpopular Constitutionalism*, 89 IND. L.J. 1133 (2014) (same).

110. *See* Polity IV Project, *Polity IV Individual Country Regime Trends, 1946-2013*, CTR. FOR SYSTEMIC PEACE [CSP] & INTEGRATED NETWORK FOR SOCIETAL CONFLICT RESEARCH, http://www.systemicpeace.org/polity/polity4.htm (last updated June 6, 2014); *The Polity Project: About Polity*, CSP, http://www.systemicpeace.org/polityproject.html (last visited Jan. 22, 2015). Coding in this Article used 2012 scores prior to the release of the 2013 Polity IV data.

111. The Codebook contains detailed procedures for how fiscal data, whether or not in U.S. Dollars, should be converted to U.S. currency using the date of the award.

112. For example, winsorizing involves identifying and converting extreme values in data into the upper or lower bounds of the distribution of the normal curve. Franck, *Development and Outcomes*, *supra* note 6, at 456 n.116 (citing W.J. Dixon, *Simplified Estimation from Censored Normal Samples*, 31

4. Defining Outcomes

There are also thorny definitional issues about what it means to "win" and whether outcomes should be viewed through objective fiscal measures or subjective measures of satisfaction. Using subjective definitions of a "win" creates confusion as to what is being measured, injects variance, and impedes the reliable measurement of outcomes. Indeed, some investors who received financial compensations from states nevertheless subjectively viewed the outcome as a "loss."[113] Others suggest, even if a state is not liable, the mere existence of the claim is a "loss" to the state.[114] Still others might classify a settlement as a "loss" — even where parties reached an agreed outcome — because not everyone obtained everything they wanted.[115] The net result is that the measurement of a "win" or "loss" is both complex and emotionally evocative.

Recognizing the strong reactions to defining a "win" or a "loss," relying on parties' subjective experiences is likely to be unreliable and lead to invalid measures that are not generalizable. An objective, fiscally based metric offers a better opportunity to assess ITA through a single lens. This research therefore used an objective, verifiable measure — namely an award's *dispositif* that describes the award's core decisions — as the basis for categorizing whether a tribunal both identified a compensable breach of international law and provided any derivative damage assessment.

Given the lack of an agreed definition, this research used a nuanced approach to outcomes. The objective was to use multiple measures to provide three discrete proxies. First, it analyzed each case's ultimate outcome in terms of whether the tribunal awarded the investor any funds, whether the respondent avoided fiscal damage for a treaty breach, or whether the parties settled or otherwise discontinued the case. Excluding settlements, outcome was coded as a binary categorical variable as either a

---

ANNALS MATHEMATICAL STAT. 385, 385 (1960); John W. Tukey, *The Future of Data Analysis*, 33 ANNALS MATHEMATICAL STAT. 1, 18–19 (1962)). Transformations using natural logs and logarithmic transformation can also constructively minimize skewing. DAVID HOWELL, STATISTICAL METHODS FOR PSYCHOLOGY 347–52 (8th ed. 2013); *see also* Franck, *ICSID Effect, supra* note 14, at 853 n.132 (using logarithmic transformations); *infra* note 139 (discussing winsorized data).

113. *See* Jack J. Coe, Jr., *Toward a Complementary Use of Conciliation in Investor-State Disputes — A Preliminary Sketch*, 12 U.C. DAVIS J. INT'L L. & POL'Y 7, 8–9 (2005) (referring to remarks by Grant Kesler, former CEO of Metalclad, who successfully recovered over US$16 million from Mexico).

114. *See* Filip De Ly, *Who Wins and Who Loses in Investment Arbitration? Are Investors and Host States on a Level Playing Field?: The* Lauder/Czech Republic *Legacy*, 6 J. WORLD INV. & TRADE 69, 69 (2005); SARAH ANDERSON & SARA GRUSKY, FOOD & WATER WATCH, CHALLENGING CORPORATE INVESTOR RULE: HOW THE WORLD BANK'S INVESTMENT COURT, FREE TRADE AGREEMENTS AND BILATERAL INVESTMENT TREATIES HAVE UNLEASHED A NEW ERA OF CORPORATE POWER AND WHAT TO DO ABOUT IT ix, 4 (2007), *available at* http://www.foodandwaterwatch.org/doc/ICSID_print.pdf.

115. EBERHARDT & OLIVET, *supra* note 5, at 9, 13, 22, 30, 58.

respondent "win" or "loss."[116] Second, recognizing that a raw "win" or "loss" hides the degree of damage, the research used a continuous variable to code the amount awarded (if any) against the respondent, including settlements and discontinuances. Third, as the prior two definitions did not address relative success, another measure of outcome used a continuous variable that calculated the percentage of investors' relative success by comparing amounts investors claimed and were ultimately awarded against states.

5. Methodology

For analyses involving two categorical variables, the analyses used a Pearson's Chi-Square Test of Independence (χ^2) to identify the presence of meaningful patterns of relationship. Similarly, when analyzing a two-level categorical and a continuous variable, an independent samples t-test explored group differences; tests with multi-level categorical variables and a continuous variable used a one-way Analysis of Variance (ANOVA) (F) to compare group differences. For continuous variables, a Pearson product-moment correlation co-efficient (r) evaluated potential linear correlations between variables.[117] A partial correlation (r) measured the degree of association between two variables while eliminating and controlling for the effect of a third variable.

D. Research Hypotheses

The research explored three independent groups of questions. First, what link did respondent development status independently have to outcome? Second, what was the relationship between development status and democracy? Third, controlling for the effect of a respondent state's democracy levels, was there any link between the state's development status and outcomes?

In keeping with previous research, the research hypothesis was that neither development status nor democracy levels would affect outcome. The theoretical assumption was that international adjudicative processes should not depend upon variables related to development or political institutions but should instead be a function of neutral adjudication based upon facts and law. A legitimate alternative research hypothesis would have been that either development status or democracy level was reliably linked

---

116. Like previous research, the respondent was coded as the ultimate winner if the tribunal awarded $0 for a treaty breach, and the respondent did not win if the tribunal awarded more than $0 for a breach. Franck, *Development and Outcomes, supra* note 6, at 456.

117. *See* TIMOTHY C. URDAN, STATISTICS IN PLAIN ENGLISH 79, 93, 105, 161 (3d ed. 2010) (explaining use of statistical tests).

to outcome. As ITA can evolve,[118] it is useful to re-test hypotheses and initial findings against updated data. Replication through expanded analysis also permits assessment of new variables, including whether a state's political institutions generate compensable economic risk.

IV. RESEARCH RESULTS

This Part describes the results of the statistical analysis. Subpart A provides an assessment of outcomes as a function of development status. Subpart B assesses the link between development and democracy that may have conflated the results in Subpart A. Subpart C then reconsiders the link between outcomes and development status — but only after controlling for the effect of a respondent's level of democracy.

A. Development Status and Outcomes

The first hypothesis involves looking for outcomes as a function of development status. Outcomes were analyzed as a function of: (1) a binary win-loss approach; (2) amounts awarded; and (3) relative success.

1. Win or Loss and Development

A first set of tests evaluated the relationship between overall outcome (i.e., a raw win or loss) and respondent development status. No test was able to ascertain a reliable link between outcomes and development status. An analysis of *only* the World Bank classification, however, suggested that there may be a latent effect, whereby High Income or Lower-Middle Income states fared better than Upper-Middle or Low Income respondent states. Otherwise, irrespective of how development status was defined, it was not possible to identify a reliable pattern of relationship between whether a state won or lost ITA and its development status.

First, a 2x2 Chi-Square cross-tabulation analyzed the 144 cases with final decisions to see if there was a statistically significant pattern of relationship between a respondent's OECD status and winning or losing the ITA. As hypothesized and indicated by previous research, there was no statistically meaningful pattern of relationship between respondent OECD status and outcome of the dispute ($\chi^2(1)=1.598$; $p=.21$; $r=.10$; $n=144$). Table 1 provides a breakdown of the outcomes and indicates that, irrespective of whether the respondent was from an OECD or non-OECD state, respondents won more than claimants.

---

118. Susan D. Franck, The Public International Law Regime Governing International Investment. *By José E. Alvarez. The Hague: Hague Academy of International Law, 2011. Pp.502. $25, €18,* 106 AM. J. INT'L L. 890 (2012) [hereinafter Franck, *Review*] (reviewing ALVAREZ, *supra* note 11).

Table 1: Breakdown of Respondent State's OECD Status and Ultimate Outcome in Final ITA Cases

	Claimant Win	Respondent Win	Total
OECD	14	30	44
Non-OECD	43	57	100
Total	57	87	144

The pattern was not statistically significant at the .05, .01, or .001 levels.

Although OECD states facially won in a slightly higher proportion, it was not statistically meaningful. Given the statistically small effect size $(r=.10)$,[119] the test was underpowered, and a sample of more than 600 additional cases would be required to reach the conclusion that there is no effect for OECD status on outcome.[120] In the interim, the results are further evidence that OECD status was not reliably linked to outcome.

Table 2: Breakdown of Respondent State's World Bank Status and Ultimate Outcome in Final ITA Cases

	Claimant Win	Respondent Win	Total
High Income	4	18	22
Upper-Middle Income	31	35	66
Lower-Middle Income	16	29	45
Low Income	6	5	11
Total	57	87	144

The pattern was not statistically significant at the .05, .01, or .001 levels.

A second model conducted a 2x4 Chi-Square cross-tabulation to identify the effect of World Bank development status on outcome. As hypothesized and indicated by previous research, there was no statistically meaningful pattern of relationship between respondent World Bank status and which party won the dispute $(\chi^2(3)=7.054; p=.07; r=.22; n=144)$. Table 2 provides a breakdown of outcome as a function of development status.

---

119. *See* LOUIS COHEN ET AL., RESEARCH METHODS IN EDUCATION 113–16 (6th ed. 2007) (providing Cohen's conventions for understanding effect sizes and indicating a "small" effect is present when $r=.10$, a "medium" effect is present when $r=.30$, and a "large" effect is present when $r=.50$, whereas effect sizes below $r=.10$ are less than "small" and arguably of trivial impact).

120. *See* JACOB COHEN ET AL., STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES 3–6, 115 (2d ed. 1988) (explaining how to conduct *a priori* power analyses); Franck, *Development and Outcomes, supra* note 6, at 461 n.132 (explaining Cohen's power analysis in ITA). The *post hoc* power of an analysis is determined by using power tables to estimate the probability of committing a Type II error (Type II error rate = 1 − power). As the power of the OECD analysis is small ($r \leq .10$ and $n= 144$), there is theoretically a 70–80% risk of having incorrectly determined that there is no relationship. In those situations where there is less than a "small" effect, the social science literature does not generally perceive a power problem, because the non-significant potential effect is small. An *a priori* power analysis for an effect size of .10 (S=781) would require a sample of 781 final investment treaty cases to have a sufficient number in the two OECD conditions. Thus, in order to establish the generally accepted baseline of .80 power (i.e., a 20% risk of a Type II error), this study would require an additional 637 cases beyond the 144 analyzed here.

It indicates that, overall, states and investors won at roughly equal levels or that respondents generally won more.

Yet the model was on the cusp of significance and somewhat underpowered, indicating that it is not appropriate to conclude that a respondent's World Bank status had no relationship to outcomes. Although there was a facial disparity, whereby both High Income *and* Lower-Middle Income states experienced more advantageous results when compared to Upper-Middle and Low Income states, it is not yet possible to determine whether the pattern was attributable to other variable(s) or random chance alone.[121] In the interim, it was noteworthy that states either won in roughly equivalent levels to investors or fared slightly better than investors overall; furthermore, there was never a sub-category of state development levels where investors won in higher proportions than states.

A third model analyzed the impact of outcomes and HDI status. A t-test analyzed the relationship between whether investors or states won ITA and the states' HDI scores. The results failed to identify a statistically significant effect of HDI on ITA outcomes ($t(141)=-.779$; $p=.44$; $r=.07$; $n=143$).[122] The mean HDI score for cases where investors won was .715 ($n=57$; standard deviation (SD)=.118), and the mean HDI for cases where states won was .730 ($n=86$; SD=.112). To contextualize scores using the 2013 UNDP report on HDI, both successful and unsuccessful respondents were in the low end of the "high human development" category; the average unsuccessful state was scored similarly to Sri Lanka (.715), and the average successful state was scored similarly to Brazil or Jamaica (.730).[123]

Although the mean HDI score for successful states was facially higher by .015 than unsuccessful states, the difference was not statistically reliable. Given the trivial effect size ($r<.10$), it is possible the test was not underpowered; yet lack of power is a risk, and more than 600 additional cases would be required to conclude definitively that there was no link between HDI and ITA outcomes.[124] In the interim, the results provide

---

121. With the small-medium effect size and the 144 cases analyzed, the analysis has 60–70% power. An *a priori* power analysis for an effect size of .22 ($r=.20$; S=191) indicates 382 final investment treaty cases (N=$n(181/2)$ * k(4)) are required to have a sufficient power for the four World Bank conditions to reliably exclude the possibility of an effect. This necessitates an additional 238 final cases or a roughly 250% increase in the historical caseload.

122. One case, *Azpetrol v. Azerbaijan* (2009), was omitted from the HDI analysis, as an HDI score for Azerbaijan was unavailable.

123. UNDP, HUMAN DEVELOPMENT REPORT 2013: THE RISE OF THE SOUTH: HUMAN PROGRESS IN A DIVERSE WORLD, at 144–45 (2013), *available at* http://hdr.undp.org/sites/default/files/reports/14/hdr2013_en_complete.pdf.

124. The *post hoc* power analysis yields a 70–80% risk of a Type II error (Type II error rate = 1 – power) because of the small effect size ($r=.07$). An *a priori* power analysis for an effect size of .10 (S=781) requires 781 final investment treaty cases to generate 80% power, which necessitates 637 additional cases to establish the requisite power.

further evidence that respondent development status was not reliably linked to outcome.

2. Amounts Awarded and Development

The data related to amounts awarded was mixed. Unlike the binary win/loss variable, or even relative success, the amounts awarded did not generate uniform results about the link between amounts awarded and state development status. Models that defined development as OECD or HDI status were unable to ascertain a reliable link with outcome. In contrast, models using the World Bank classification *did* identify a meaningful difference in amounts awarded as a function of development. The key distinction in that model was that High Income states experienced comparatively lower amounts awarded than *only* their Upper-Middle Income counterparts. This Subpart therefore explores amounts awarded — both overall and for the subset of investor wins — by analyzing relationships between different measures of development status.

(a) Awards as a Function of OECD Status

A first set of models analyzed amounts awarded as a function of the respondent's OECD status. Both models indicated, irrespective of whether the focus was all final awards or only the subset of investor wins, that it was not possible to identify a reliable relationship between OECD status and amounts awarded.

For the set of all awards, three tests failed to identify a statistically significant link between OECD status and amount awarded. The first test used inflation-adjusted raw data with statistical outliers ($t(157)=.698$; $p=.49$; $r=.06$; $n=159$); a second test used winsorized data ($t(157)=1.371$; $p=.17$; $r=.11$; $n=159$) to minimize the impact of skewing and match the statistical assumptions of the test, but still used readily understandable U.S. Dollar amounts; and the third test used logged data ($t(1,157)=1.414$; $p=.16$; $r=.11$; $n=159$) with the least degree of skewing. Table 3 provides the mean amounts tribunals awarded against OECD and non-OECD states using inflation-adjusted winsorized data.[125] Although there was a facial pattern suggesting that the non-OECD states experienced somewhat larger awards than their OECD counterparts, that pattern was not reliable and could derive from random chance.

---

125. Winsorized data were presented because winsorizing offers data in U.S. Dollar values, which is a meaningful scale with practical significance, and it had the lowest levels of skewing for data in U.S. Dollars.

Table 3: Amounts Awarded in Final ITA Cases Using Inflation-Adjusted Winsorized Damages as a Function of Respondent State OECD Membership

Respondent State	Mean U.S. Dollar Award	SD	Total
OECD	2,176,168	4,250,501	49
Non-OECD	3,295,176	4,958,055	110
Total	2,950,324	4,766,128	159

The pattern was not statistically significant at the .05, .01, or .001 levels.

For the subset of those awards where investors were successful, it was not possible to identify a reliable link between OECD status and amounts awarded. This held true across three tests using inflation-adjusted raw ($t(56)=.057$; $p=.96$; $r<.01$; $n=58$), winsorized ($t(56)=.564$; $p=.58$; $r=.08$; $n=58$), and logged ($t(56)=.330$; $p=.74$; $r=.04$; $n=58$) data. For this subset, the mean amount awarded against OECD states using winsorized data was US$34.6 million (SD=52,705,728; $n=14$), whereas the mean award against non-OECD states was US$44.5 million (SD=58,722,315; $n=44$).[126]

Although the raw inflation-adjusted data include statistical outliers,[127] Figure 1 reflects the variance in amounts awarded as a function of whether the respondent was an OECD or non-OECD state. The thick black line, representing the median, indicates that when either an OECD or non-OECD state lost a case, the median amounts awarded were of a similar value and under US$40 million. This is reflected by the inflation-adjusted raw data. For OECD states, the median amount awarded was US$8,295,911.[128] For non-OECD states, the median amount awarded was US$11,878,972.[129] Figure 1 also reflects that although the highest quartile for non-OECD states was slightly higher than OECD states, both groups experienced statistical outliers.

---

126. Out of an abundance of caution, a Mann-Whitney *U*-test evaluated median differences between two groups. Using inflation-adjusted raw data, that test also failed to reveal a meaningful group difference between median awards of OECD and non-OECD states in all cases ($U=2363.0$; $p=.15$) and the subset of investor wins ($U=284.0$; $p=.66$).

127. Although all test results were provided to address commentary on earlier research expressing confusion about why some analyses were omitted, the tests with logged data had the greatest statistical conclusion validity for analyzing amounts awarded. For the set of all awards, the skewing of raw data was 3.94, 1.21 for winsorized data, and .639 for logged data (SE=.19). For the subset of awards where investors won, the skewing was 2.16 for raw data, 1.47 for winsorized data, and -.72 for logged data (SE=.31).

128. Using inflation-adjusted raw data, the lowest quartile of amounts awarded against OECD states was US$3,750,878, and the highest quartile was US$47,718,081.

129. Using inflation-adjusted raw data, the lowest quartile of amounts awarded against non-OECD states was US$3,357,419, and the highest quartile was US$78,045,296.

Figure 1: Stem-and-Leaf Plot of Damages Awarded Using Raw Inflation-Adjusted Data for Subset of Final ITA Cases Generating Payment to Investor as a Function of Respondent State's OECD Status

Respondent's OECD Status at Date of Award (n=58)
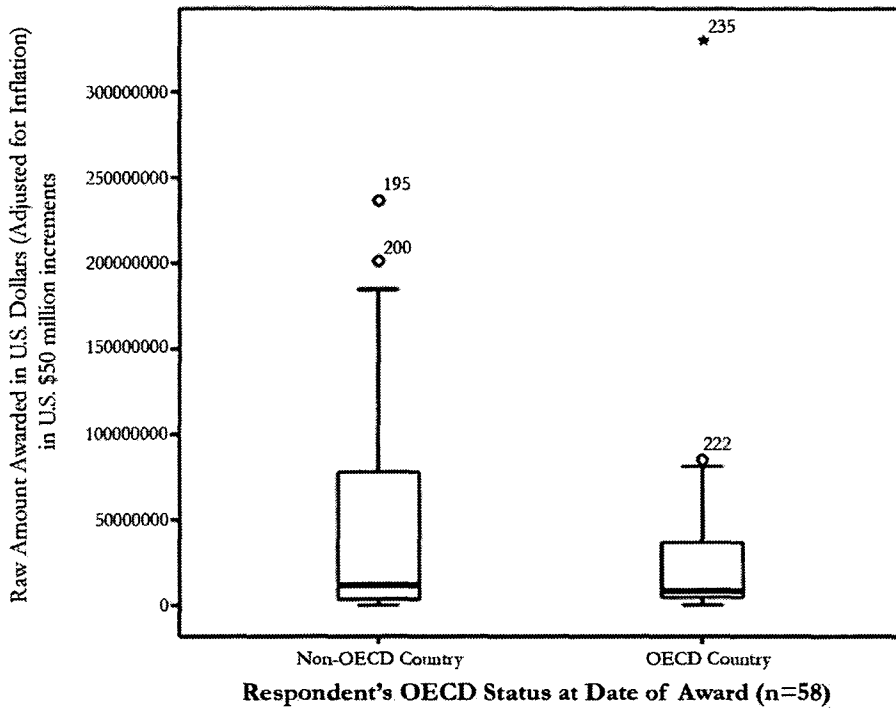
Although non-OECD states experienced slightly higher amounts awarded overall in the subset of investor wins, those results were not statistically meaningful. The effect sizes for all models were statistically small ($r=.10$ or $r<.10$). While there is a risk that tests were underpowered, given the potentially small latent effect, a sample of more than 600 additional cases would be required to conclude definitively that OECD status was unrelated to amounts awarded.[130] In the interim, the small effect sizes and non-significant results are an additional piece of evidence that OECD status was not reliably linked to outcomes.

(b) Awards as a Function of World Bank Status

A second set of models analyzed amounts awarded as a function of the respondent's World Bank status. In contrast to the findings related to OECD (and the next analysis of HDI), both models revealed a statistically

---

130. *See supra* note 120 (observing how, when analyzing the two-category OECD variable when the effect size is small or less than small ($r \leq .10$) a sample of 781 final ITA cases would be required to obtain sufficient levels of power, which requires 637 final cases to establish the requisite power).

meaningful link between respondent states' World Bank status and amount awarded.

For the set of all awards, three tests identified a statistically meaningful link between a respondent state's development status and the amount awarded. Irrespective of whether inflation-adjusted raw $(F(3,155)=4.223; p=.01; r=.28; n=159)$, winsorized $(F(3,155)=4.176; p=.01; r=.27; n=159)$, or logged $(F(3,155)=3.833; p=.01; r=.26; n=159)$ data were used, there was a reliable link between development status and outcomes. Table 4 displays the mean amounts awarded against states as a function of their World Bank classification.

Table 4: Inflation-Adjusted Winsorized Damages of Amounts Awarded in Set of All Final ITA Cases as a Function of Respondent State's World Bank Classification

Respondent State	Mean U.S. Dollar Award	SD	Total
High Income	626,310	2,410,144	28
Upper-Middle Income	3,986,975	5,356,584	71
Lower-Middle Income	2,359,771	4,195,751	46
Low Income	4,281,437	5,285,719	14
Total	2,950,324	4,766,128	159

The overall pattern was statistically significant at the .05 and .01 levels.

The main effect for World Bank status, however, was somewhat misleading. Not every follow-up comparison between categories generated the statistically significant result. Follow-up analyses using Tukey's HSD[131] on winsorized data revealed that *only* awards rendered against High Income states were reliably lower than awards against Upper-Middle Income states. All five other follow-up analyses[132] failed to identify a statistically meaningful difference attributable to World Bank status. Analyses using raw[133] and logged[134] data exhibited a similar pattern.

For the subset of awards where investors were successful, respondents' World Bank status was also reliably linked to amount awarded. This held

---

131. Tukey's honestly significant difference (HSD) test provides a follow-up for testing of a significant omnibus effect to determine whether there are statistically significant differences among different subsets of groups. FREDERICK J. GRAVETTER & LARRY B. WALLNAU, ESSENTIALS OF STATISTICS FOR THE BEHAVIORAL SCIENCES 376 (8th ed. 2013); Franck, *Development and Outcomes*, *supra* note 6, at 457 n.123, 468 n.151.

132. The other follow-ups involved: (1) High versus Lower-Middle; (2) High versus Low Income; (3) Upper-Middle versus Lower-Middle; (4) Upper-Middle versus Low; and (5) Lower-Middle versus Low.

133. Using inflation-adjusted raw data with statistical outliers, there was a significant effect whereby Lower-Middle Income states experienced *lower* mean awards (US$6,218,317; SD=17796366; $n=46$) than Upper-Middle states (US$30,479,694; SD=6449143; $n=71$). As the data included outliers, the results must be treated cautiously.

134. Using inflation-adjusted logged data, there was a significant effect. High Income states experienced lower mean awards (.89; SD=2.24; $n=28$) than Low Income states (3.92; SD=3.56; $n=14$).

true across the three models using inflation-adjusted raw $(F(3,54)=3.618;$ $p=.02;$ $r=.41;$ $n=58)$, winsorized $(F(3,54)=4.113;$ $p=.01;$ $r=.43;$ $n=58)$, and logged $(F(3,54)=4.457;$ $p=.01;$ $r=.45;$ $n=58)$ data. Unlike any other analysis in this Article, the effect size of this model reflects that the effect of World Bank status on amounts awarded was a medium to large effect using Cohen's conventions.[135] Table 5 reflects the mean amounts awarded in the subset of awards where investors obtained damages, whether by virtue of a settlement agreement or tribunal award. [136]

Table 5: Inflation-Adjusted Winsorized Damages of Amounts Awarded for Subset of Final ITA Cases Generating Payment to Investor as a Function of Respondent State's World Bank Classification

Respondent State	Mean U.S. Dollar Award	SD	Total
High Income	5,975,876	8,534,954	4
Upper-Middle Income	65,460,334	66,930,450	30
Lower-Middle Income	17,877,662	26,977,657	16
Low Income	21,151,169	37,607,021	8
Total	42,120,095	57,037,330	58

The overall pattern was statistically significant at the .05 level.

Similar to the earlier discussion, for the subset of cases where investors won, the main effect of a respondent's World Bank status was misleading. Not all follow-up analyses generated significant results. Follow-up analyses using Tukey's HSD tests on winsorized data showed that *only* the awards against Lower-Middle Income states were statistically lower than Upper-Middle Income states. All five other follow-up analyses did not detect a statistically meaningful difference,[137] and analyses using raw and logged[138] data exhibited similar results.

Although the raw inflation-adjusted data include statistical outliers, Figure 2 reflects the variance in amounts awarded as a function of a respondent's World Bank classification through a stem-and-leaf plot. The thick black line, which represents the median, indicates that all states experienced median awards under US$50 million. More specifically, using inflation-adjusted raw data, the median amounts awarded were

---

135. *See supra* note 119 and accompanying text.

136. A Kruskal-Wallace test evaluated median differences between four groups. Using inflation-adjusted raw data, that test revealed a meaningful group difference in median awards for all cases $(\chi2$ $(3)=11.452;$ $p=.01;$ $r=.26;$ $n=159)$ and the subset of investor wins $(\chi2(3)=10.815;$ $p=.01;$ $r=.39;$ $n=58)$.

137. The other follow-ups were: (1) High versus Lower-Middle; (2) High versus Low Income; (3) Upper-Middle versus Lower-Middle; (4) Upper-Middle versus Low; and (5) Lower-Middle versus Low.

138. The follow-up tests using the inflation-adjusted raw data exhibited the same pattern. Using inflation-adjusted logged data, however, there was also a significant effect whereby High Income states (M=6.24; SD=.89; $n=4$) experienced lower mean amounts awarded than Upper-Middle Income states (M=7.42; SD=.77; $n=30$).

US$2,658,294 for High Income, US$6,606,299 for Lower-Middle Income, and US$8,878,955 for Low Income states. In contrast, the median amount awarded was US$38,567,138 for Upper-Middle Income states. Nevertheless, Upper-Middle Income states experienced a wider distribution of damage awards than the three other classifications and also experienced a higher median award.

Figure 2: Stem-and-Leaf Plot of Damages Awarded Using Raw Inflation-Adjusted Data for Subset of Final ITA Cases Generating Payment to Investor as a Function of Respondent State's World Bank Classification
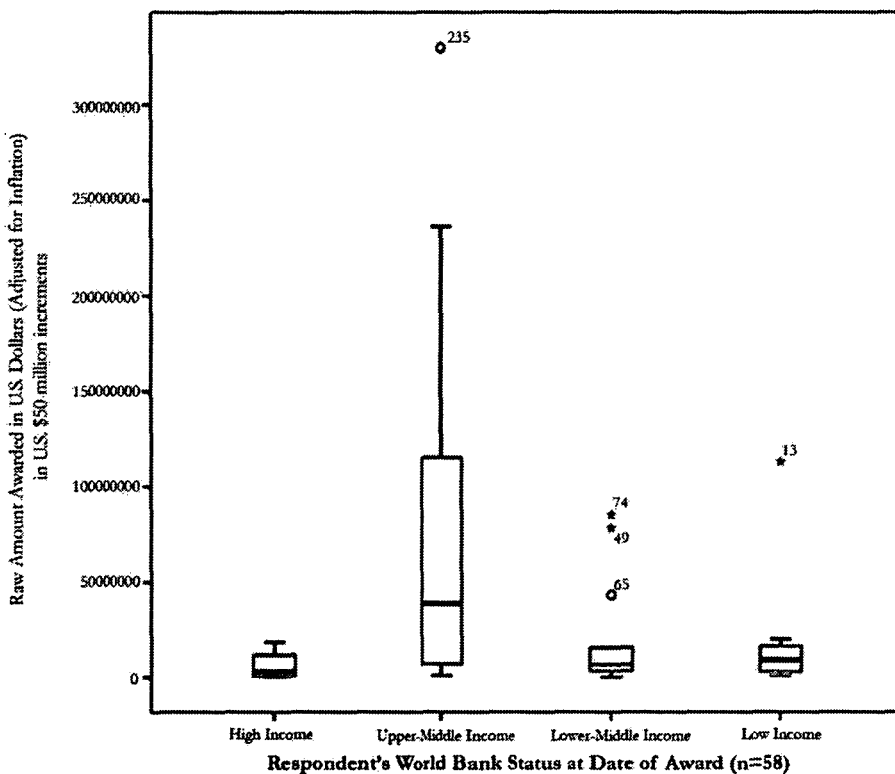
(c) Awards as a Function of HDI Scores

A third set of models analyzed amounts awarded to investors — whether by virtue of a tribunal or via a settlement agreement — varied as a function of a respondent state's HDI status. Like the results on OECD status, all models failed to indicate that there was a reliable relationship between HDI and amounts awarded. This held true for the set of all final awards and the subset of cases involving disputes with an investor recovery.

For the set of all awards, three tests failed to identify a statistically significant link between HDI status and amount awarded. Using a Pearson's Product-Moment Correlation Coefficient, it was not possible to identify a reliable linear relationship between HDI and amounts awarded using inflation-adjusted raw ($r(157)=.08$; $p=.35$), winsorized ($r(157)=-.09$; $p=.26$), or logged ($r(157)=-.12$; $p=.14$) data. Figure 3 uses inflation-adjusted winsorized data[139] to explore how amounts awarded vary according to a respondent's HDI score. Figure 3 reflects three aspects. First, a variety of states across the HDI scale were successful in avoiding any amounts awarded. Second, a series of states across the HDI scale experienced awards at the 75th percentile or higher. Third, for the group experiencing awards between those levels, there were a range of HDI scores.

Figure 3: Scatterplot of Damages Awarded Using Winsorized Inflation-Adjusted Data in Set of All Final ITA Cases as a Function of Respondent State's HDI Score

---
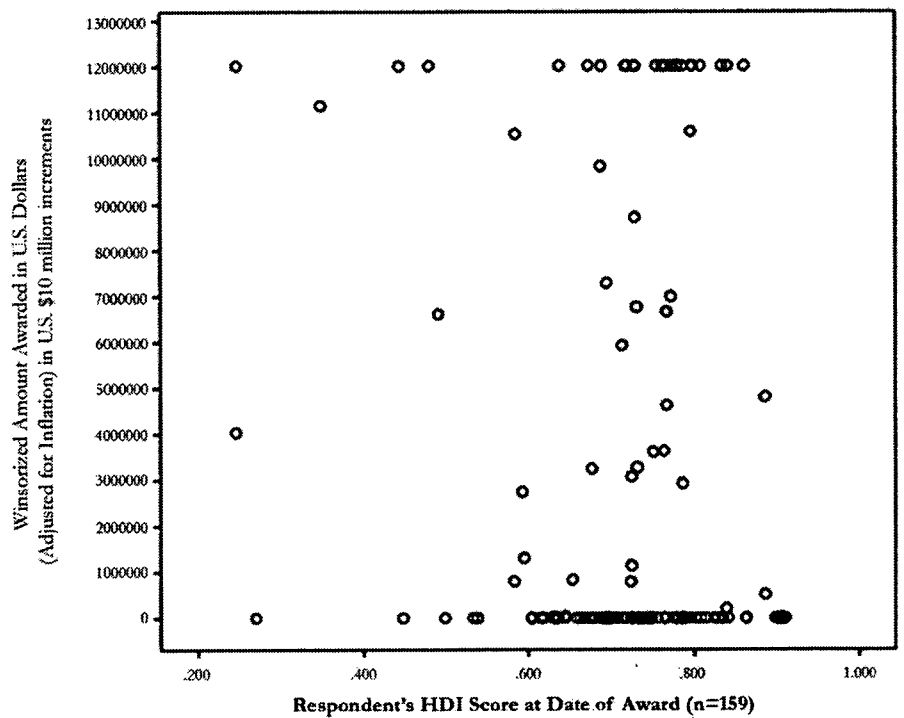
139. The process of winsorizing is an accepted practice in psychology to use Tukey's hinges to minimize the effect of statistical outliers. Winsorizing ensures that outliers beyond the seventy-fifth percentile of data distribution are capped at a common value to prevent data from being lost, while also ensuring that skewed data does not inadvertently disrupt the assumptions underlying the test. This accounts for the amounts awarded being capped at approximately US$120 million.

For the subset of awards where investors obtained damages, HDI was also not reliably linked to amounts awarded. This held true across the three tests using inflation-adjusted raw ($r(58)=.24$; $p=.08$), winsorized ($r(58)=.23$; $p=.08$), and logged ($r(58)=.16$; $p=.24$) data. The raw and winsorized data were on the cusp of significance but exhibited less acceptable levels of skewing than the logged data. The nature of the potential relationship was such that, the higher the amount awarded, the higher the respondent's HDI score (i.e., the more developed the respondent state). The logged data, which exhibited the most appropriate skewing levels, was not on the cusp of significance.

The effect sizes for models of all amounts awarded were either statistically small or trivial ($r=.10$ or $r<.10$) and not necessarily underpowered;[140] but for the subset of cases involving investor wins, the potential small- to medium-sized effects meant the analyses were underpowered.[141] To reliably ascertain the presence of the smallest effect, a sample of more than 600 additional cases would be required to conclude that HDI was definitively unrelated to amounts awarded. In the interim, the results provide evidence that HDI scores were not reliably linked to amounts awarded.

3. Relative Success and Development

Models evaluating relative investor success offer a uniform scale for analyzing outcomes irrespective of amount claimed and provide a proxy to partially control for a case's underlying merit.[142] As such, models analyzing

---

140. *See supra* note 120 (observing how, when analyzing a correlation using HDI for a small or less than small ($r \leq .10$) effect size, a sample of 781 final ITA cases would be required to obtain sufficient levels of power, which requires an additional 637 final cases).

141. For the small- to medium-sized effects ($r=.16–r=.24$), there was roughly 20–40% power and an unacceptable risk of error. To obtain sufficient power to identify the smallest of those effects ($r=.16$), 373 final cases where investors won would be required to establish sufficient power. In other words, this would require 315 new cases where investors won or, put differently, a roughly 600% increase in the historical number of cases in which investors obtained damage awards.

142. Given the lack of de jure precedent in ITA and the absence of an appellate body to create legal coherence, it is difficult to identify legally correct outcomes. *See* Nick Gallus, *Protection of Non-Governmental Organizations in Egypt Under the Egypt – U.S. Bilateral Investment Treaty*, 14 INT'L J. NOT-FOR-PROFIT   L.,   Sept.   2012,   at   62,   67,   *available   at* http://www.icnl.org/research/journal/vol14iss3/v14n3%20final.pdf ("[D]ecisions are certainly not always consistent and a decision on the interpretation of a provision is far from a guarantee that a similar provision, or even the same provision, will be interpreted the same way by another tribunal."); Michael D. Goldhaber, *The Global Lawyer: Arbitration Without Legitimacy*, LIT. DAILY, June 7, 2013, *available   at*   http://www.litigationdaily.com/id=1202603368340/The-Global-Lawyer-Arbitration-Without-Legitimacy (quoting Catherine Rogers as stating, "no study can control for the correct legal outcome"); Catherine A. Rogers, *The Politics of International Investment Arbitrators*, 12 SANTA CLARA J. INT'L L. 223, 234–35 & n.42 (2013) (noting that others attempted to control for a case's legal correctness "by having those challenges assessed by a panel of reputable investment arbitration specialists" after the dispute was decided, which creates a risk of hindsight bias and suggests that relative success provides a measure to control for legal merits).

this variable are more valuable than models analyzing amounts awarded. The last group of tests, therefore, evaluated development status as a function of relative investor success — namely the percentage difference between an investor's claim and the amount awarded.[143] Irrespective of whether OECD, World Bank, or HDI measures were used, it was not possible to identify a reliable link between respondent development status and relative investor success.

First, a t-test analyzed respondent's OECD status and relative investor success. The results failed to identify a reliable relationship between the respondent's OECD status and relative investor success ($t(101)=.707$; $p=.48$; $r=.07$; $n=103$). See Table 6. Even when focusing on the small subset of cases where investors won, it was also not possible to establish a reliable link between investors' relative success and respondents' OECD status ($t(51)=-.611$; $p=.54$; $r=.09$; $n=53$).

Table 6: Mean Rates of Relative Investor Success in Set of All Final ITA Cases as a Function of Respondent State's OECD Status

Respondent State	Relative Percent of Investor Success	SD	Total
OECD	15.35	28.12	33
Non-OECD	19.25	25.11	70
Total	18.00	26.04	103

The pattern was not statistically significant at the .05, .01, or .001 levels.

Table 6 reflects a facial trend where investors were generally more successful in obtaining relatively larger awards against non-OECD states. This facial difference, however, was not statistically meaningful, could have been due to chance alone, and was less than statistically small.[144] Moreover, when focusing on the subset of investor wins, the facial pattern flipped, and investors suing OECD states were relatively more successful (39%) compared to their non-OECD counterparts (33.7%). Given the trivial effect sizes ($r<.10$) for the non-significant effects, the tests may not have been underpowered; yet, given the risk of error, replication will require more than 600 additional cases to conclude reliably that there was no effect of respondent OECD status on relative investor success.[145] The

---

143. Because not all awards contained a reliable statement of the amount of an investor's claim, as compared to previous analyses, this decreased the cases analyzed.

144. Previous scholarship has described the importance of effect sizes in understanding the importance of statistically reliable relationships and the importance of understanding the potential size of potential latent effects. *See* Franck, *Development and Outcomes, supra* note 6, at 457–58, 461, 466, 475 (discussing the role and importance of effect sizes); Franck, *ICSID Effect, supra* note 14, at 856–60 (discussing the effects of various effect sizes that were small or less than small).

145. *See supra* notes 120 & 130 (observing how, when analyzing the two-category OECD variable when the effect size is small or less than small ($r \leq .10$) a sample of 781 final ITA cases would be required to obtain sufficient levels of power, which requires 637 new cases for sufficient power).

results provide additional evidence that ITA outcomes were not dependent on respondents' OECD status.

A second model used a one-way ANOVA to analyze respondent World Bank status and relative investor success.[146] The results failed to identify a reliable relationship between respondent World Bank status and relative success $(F(3,99)=.419; p=.74; r=.11; n=103)$. See Table 7. Even when focusing on the small subset of cases where investors won, it was not possible to establish a reliable link between investors' relative success and a respondent's World Bank status $(F(3,49)=1.025; p=.39; r=.24; n=53)$.

Table 7 suggests that investors may experience more relative success against Low Income states. This facial difference, however, was not statistically meaningful. The facial difference must be viewed cautiously, as the median rate of relative investor success was 1.9%.[147]

Table 7: Mean Rates of Relative Investor Success in Set of All Final ITA Cases as a Function of Respondent State's World Bank Development Status

Respondent State	Relative Percent of Investor Success	SD	Total
High-Income	14.53	32.53	15
Upper-Middle Income	17.40	22.69	47
Lower-Middle Income	17.95	26.78	31
Low Income	26.19	30.34	10
Total	18.00	26.04	103

The pattern was not statistically significant at the .05, .01, or .001 levels.

For the subset of investor wins, the facial pattern of relative investor success flipped. In that subset, the mean investor success rate was 35.98% $(n=53)$. High Income states were exposed to the highest relative mean recovery (54.5%; $n=4$), followed by Low Income states (37.4%; $n=7$), and then Lower-Middle Income States (37.1%; $n=15$). Upper-Middle Income states experienced the lowest levels of relative investor success (30.3%; $n=27$).[148] In other words, investors arbitrating against High Income states obtained awards for relatively more of their claims, whereas investors suing Upper-Middle Income states experienced the lowest levels of relative investor success. Yet inferences should be made cautiously, given the lack of statistical significance and the risk that effects could be due to chance alone. There was a risk that the tests were underpowered, particularly for

146. Because not all awards contained a reliable statement of the amount of an investor's claim, as compared to previous analyses, this decreased the number of cases analyzed in the model.

147. For the set of all cases, the median success rates were: (1) 0.0% for High Income; (2) 17.4% for Upper-Middle Income; (3) 0.0% for Lower-Middle Income; and (4) 14.7% for Low Income states.

148. This median rate of investor success was similar for the subset of investor wins. The overall median success rate was 28.7% for the subset of cases involving an investor win $(n=53)$. For the subset of awards where investors were successful, median success rates were: (1) High Income = 54.6%; (2) Upper-Middle = 27.3%; (3) Lower-Middle = 29.7%; (4) Low Income = 27.7%.

the subset of cases where investors won. Nevertheless, replication with nearly 1500 more cases should have sufficient statistical power to reliably identify whether there is a link between World Bank status and relative investor success.[149] Nevertheless, the results offer another piece of evidence that development status was not reliably linked to outcome.

Figure 4: Scatterplot of Investor's Relative Success in Set of All Final ITA Cases as a Function of Respondent State's HDI Score
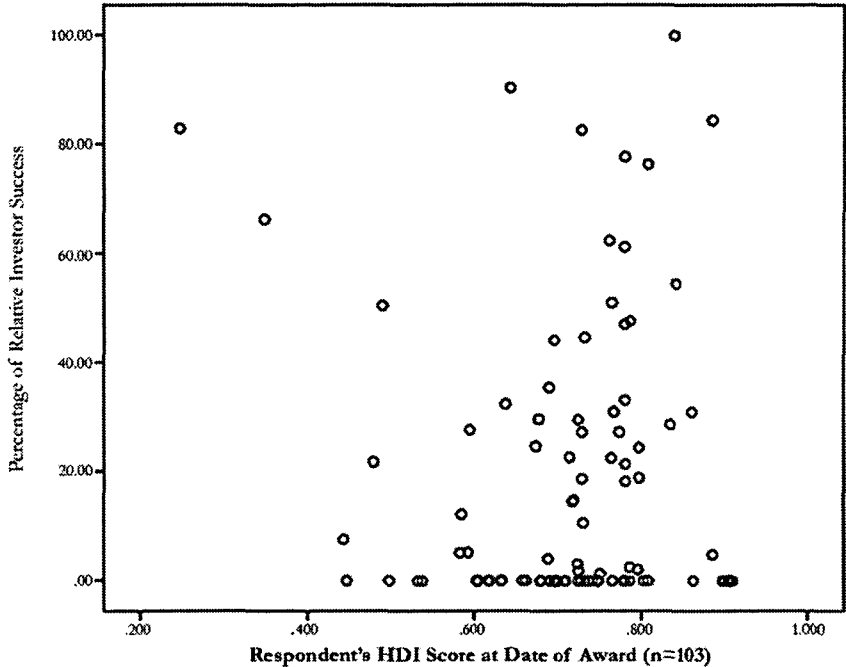
A third model evaluated the relationship between an investor's relative success and respondent HDI scores. Using a Pearson's Product-Moment Correlation Coefficient, it was not possible to identify a meaningful linear relationship between investor relative success and HDI ($r(103)=-.07$;

---

149. Where the effect size is small, as for relative investor success rates in all cases as a function of respondent World Bank status ($r=.11$; S=781), the power of the analysis is less than 20%, which may reflect the absence of the effect or a latent small effect. *A priori* power analysis would require a total of 1,562 final awards to provide sufficient power to detect a potential effect. This means it would be necessary to ascertain 1,459 final awards with complete information on relative success. For the subset of investor wins, there was a small- to medium-effect size ($r=.25$; S=120), which means a *post hoc* power analysis has only 50% power. *A priori* power analysis indicates that, to have sufficient power to reliably ascertain the impact of the four-category variable of respondents' World Bank status, there would need to be 240 total ITA awards rendered in investors' favor (N=n(120/2) * k(4)). As there were only 53 cases with complete information on investors' relative success where investors won, this requires 187 total new cases, which is a roughly 300% increase in cases. *See also supra* note 121 (observing how, when there is a small- to medium-effect size ($r=.20$), 382 total final treaty awards were required).

p=.47). Figure 4 reflects the distribution of relative investor success, which compared amounts claimed and awarded with respondent HDI status. Even for the subset of awards involving investor success, there was no significant link between investors' relative success and HDI (r(58)=-.03; *p*=.83). Although the results indicated that any potential relationship was close to zero, the facial trends both suggested that the lower respondents' HDI score, the higher investors' relative success. Given the statistically small effect size (r<.10), it is possible that the analyses were not underpowered. But the risk of error related to the potential latent effect means that more than 600 additional cases are needed to conclude definitively that relative success and HDI operated independently. Until then, the results are further evidence that a state's HDI status did not affect ITA outcomes.[150]

B. Development Status and Democracy Levels

Two different models — using amounts awarded[151] and relative investor success[152] — indicated a respondent state's democracy levels were linked to ITA outcomes at the *p*<.10 level.[153] The facial trend, which was non-significant at the traditional *p*<.05 level, was that higher levels of democracy were associated with positive ITA outcomes for states.[154]

Given those results, it is useful to explore the relationship between respondent democracy levels and development status. Irrespective of how development was defined, tests always identified a reliable link between a respondent state's development status and democracy level. This suggests

---

150. *See supra* notes 120 & 130 (observing how, when analyzing correlation between HDI and outcomes when there is a small or less than small (r ≤ .10) effect size, a sample of 781 final ITA cases would be required to obtain sufficient power to reliably exclude the possibility of a latent effect, which requires an additional 637 final cases).

151. A Pearson's Product-Moment Correlation Co-efficient between a respondent's Polity IV score and winsorized inflation-adjusted amounts awarded came closest to statistical significance (r(158)=-.15; *p*=.06). The direction of the potential relationship suggested that the lower a state's level of democracy, the higher the overall amount awarded; and the higher a state's level of democracy, the lower the amount awarded. Logged inflation-adjusted data reflected the same direction of the data (r(158)=-.11; *p*=.19).

152. A Pearson's Product-Moment Correlation Co-efficient of Polity IV score and an investor's relative success was not significant at traditional *p*<.05 levels (r(103)=-.18; *p*=.07). The direction of the potential relationship was that as a state's democracy level increases, the smaller the investor's relative success; and as a state's democracy level decreases, the higher the investor's relative success.

153. For the set of all awards, there was not a statistically significant relationship between whether an investor or state won and the democracy score (r(158)=.08; *p*=.35). The direction of the potential relationship was such that states with higher levels of democracy were linked with wins, whereas states with lower levels of democracy were associated with losses.

154. For the subset of investor wins, it was not possible to identify a significant link between outcomes and Polity IV scores. There was no reliable linear relationship for amounts awarded using winsorized (r(58)=-.14; *p*=.30) or logged (r(58)=-.03; *p*=.80) data; likewise there was no reliable relationship between investors' relative success and democracy levels (r(53)=-.16; *p*=.24).

that development status and democracy levels were related, and it is critical to control for potential co-linearity.

For all of the 202 individual cases in the dataset,[155] there was a strong, reliable, and robust relationship between a respondent state's level of development and its level of democracy. First, there was a link between a respondent's OECD status and Polity IV ($r(201)=.29$; $p<.001$). The direction was such that states that were OECD members were more likely to have high democracy levels, whereas non-OECD members were less likely to have high democracy levels. Second, there was a reliable link between respondent World Bank status and Polity IV ($r(201)=-.57$; $p<.001$). The direction was such that High Income states were more likely to have high levels of democracy, whereas less developed states were likely to have lower levels of democracy.[156] Third, there was a link between a respondent's HDI and Polity IV scores ($r(199)=.60$; $p<.001$). The direction was such that, the higher a state's human development, the more likely it was to experience high levels of democracy. In contrast, a lower human development score was linked with lower levels of democratic institutions.

For the set of final awards,[157] there was also a robust link between development status and respondent state democracy. First, there was a link between OECD status and Polity IV ($r(158)=.30$; $p<.001$). Second, there was a link between World Bank status and Polity IV ($r(158)=-.46$; $p<.001$). Third, there was a link between HDI scores and Polity IV ($r(156)=.59$; $p<.001$). For all measures of development status, as respondent development status increased, it was more likely to have increased levels of democracy. Likewise, the lower the respondent's development status, the less likely it was to experience high levels of democracy. There were similar results for the subset of cases in which investors were successful.[158]

This relationship, which generally reflected a large statistical effect, suggests that it is possible that the debate on ITA has perhaps confused development concerns with concerns that relate to democratic governance or a state's internal good governance practices. Should concerns about

---

155. It was not possible to identify a Polity IV and/or a HDI score for a small number of cases.

156. The negative statistical relationship is a by-product of how World Bank classifications were coded, with "1" as High Income, "2" as Upper-Middle Income, "3" as Lower-Middle Income, and "4" as Low Income.

157. For the 159 final cases, it was not possible to identify a Polity IV and/or a HDI score for respondents in 1–3 cases.

158. For the subset of investor wins, there was a strong and reliable relationship between: (1) Polity IV and World Bank Status ($r(58)=-.47$; $p<.001$); and (2) Polity IV and HDI ($r(58)=.65$; $p<.001$). The link between OECD status and Polity IV was non-significant ($r(58)=.22$; $p=.10$) at the traditional $p=.05$ level. The facial trend for the OECD analysis mirrored the significant results for World Bank and HDI classifications, namely, when respondent states' HDI levels increased, states were more likely to have higher levels of democracy. Likewise, when respondent states' HDI levels decreased, states were more likely to have lower levels of democracy.

development be conflated with those related to domestic governmental institutions and political infrastructure, concerns about the integrity of ITA may be attributed to the wrong variable. It is possible that outcomes may reflect a host state's underlying political situation, commitment to democratic governance, or some other explanatory variable(s), rather than a state's development status or poverty levels. To address the potential conflation, the next Subpart reevaluates outcomes — but only after controlling for respondents' internal levels of democracy.

C. Controlling for Democracy Levels — Re-Evaluating the Link Between Development Status and Outcomes

There is value in tracing the variance in ITA outcomes to meaningful elements, rather than misattributing the impact of spurious or confounding variables. The research therefore used a partial correlation to control for the effect of a respondent state's democracy levels on ITA outcomes. The tests explored whether there were any subsisting relationships between development status and outcomes after removing the variance attributable to Polity IV scores. A series of twelve partial correlations were completed to identify whether, overall, there was a link among: (1) OECD membership, (2) World Bank classification, and (3) HDI scores and outcomes. The four outcome variables analyzed were: (1) raw wins and losses; (2) inflation-adjusted amounts awarded using winsorized data; (3) inflation-adjusted amounts awarded using logged data; and (4) investor relative success.

Table 8 reports the results of those twelve tests. In every test the results were not statistically significant, and the effect size of the potential relationship was less than statistically small (i.e., $r < .10$). In other words, after eliminating the influence of host state democracy levels, the potential strength of the non-significant relationship was weak. Given the small effect size of any potential effect, the statistical power of the analysis was necessarily low, and technically, the risk of error for small effect sizes means it is not possible to conclude a relationship is impossible. With such a small effect size, however, a priori power analysis requires obtaining a sample of more than 600 additional final cases to reliably ascertain whether a latent effect is present or absent.[159] This is roughly equivalent to

---

159. For the set of all awards in Table 8, the largest effect size for the partial correlations was $r = .12$ and the smallest was $r = .01$. In order to identify the smallest potential effect, a priori power analysis necessitates roughly 80% power, or a 20% risk of a Type II error. With such a small .10 effect size, it is therefore necessary to gather a large group of cases, or 781, to identify the possibility of a reliable correlation among the variables. With 159 final awards at present, not all of which contain complete information on relative success, at a minimum, it will be necessary to obtain 637 extra final cases to establish the requisite power. See also supra notes 120, 130, & 150 (identifying the need for 781 final cases to conduct analysis and requiring 637 further final cases).

taking the first two decades of final investment treaty cases, where there were 159 final awards, and increasing that caseload by roughly 400%.

Table 8: All Cases: Partial Correlation Matrix Reflecting Relationships Among Respondent State Development Status and Outcome, Controlling for Respondent State Polity IV Score

Outcome and Type of State Development Status	Correlation Co-Efficient (r)	Statistical Significance (p)	Total
Win/Loss and OECD	.09	.27	139
Win/Loss and World Bank	-.07	.39	139
Win/Loss and HDI	.02	.87	139
Inflation-Adjusted Awards (Winsorized) and OECD	-.07	.36	153
Inflation-Adjusted Awards (Winsorized) and World Bank	.05	.57	153
Inflation-Adjusted Awards (Winsorized) and HDI	.01	.95	153
Inflation-Adjusted Awards (Logged) and OECD	-.09	.24	153
Inflation-Adjusted Awards (Logged) and World Bank	.12	.15	153
Inflation-Adjusted Awards (Logged) and HDI	-.07	.42	153
Relative Investor Success and OECD	-.02	.84	100
Relative Investor Success and World Bank	.01	.92	100
Relative Investor Success and HDI	.06	.53	100

None of the partial correlations was significant at the .05, .01, or .001 levels.

A second set of analyses focused only on the smaller subset of those cases where investors won. The objective was again to remove the variance attributable to a state's democracy levels and then re-assess the potential link between state development status and outcomes. A series of nine different partial correlations explored the potential links among: (1) OECD membership, (2) World Bank classification, and (3) HDI scores

and outcomes. The outcome variables analyzed were: (1) inflation-adjusted amounts awarded using winsorized data; (2) inflation-adjustment amounts awarded using logged data; and (3) investor relative success.

Table 9 provides the results of the nine tests. Even for the subset of cases reflecting some form of investor recovery, none of the tests could identify a statistically meaningful link between development and outcomes after controlling for the effect of a respondent state's democracy level.

Table 9: Subset of Cases Involving Investor Awards: Partial Correlation Matrix Reflecting Relationships Among Respondent State Development Status and Outcome, Controlling for Polity IV Score

Outcome and Type of State Development Status	Correlation Co-Efficient (r)	Statistical Significance (p)	Total
Inflation-Adjusted Awards (Winsorized) and OECD	-.10	.78	55
Inflation-Adjusted Awards (Winsorized) and World Bank	-.20	.14	55
Inflation-Adjusted Awards (Winsorized) and HDI	.23	.09	55
Inflation-Adjusted Awards (Logged) and OECD	-.04	.78	55
Inflation-Adjusted Awards (Logged) and World Bank	-.17	.19	55
Inflation-Adjusted Awards (Logged) and HDI	.24	.08	55
Relative Investor Success and OECD	.13	.37	50
Relative Investor Success and World Bank	-.10	.48	50
Relative Investor Success and HDI	.13	.38	50

None of the partial correlations was significant at the .05, .01, or .001 levels.

The potential relationships that came closest to generating a reliable link involved amounts awarded using either a World Bank- or HDI-based definition of development status. In each of those analyses, the direction of the potential relationship was notable; namely, as states' development

levels increased, the amounts awarded against states increased.[160] If the concern about ITA is that it disparately affects the *developing* world (rather than the developed world), the results potentially undercut the narrative. A reasonable alternative explanation for the results, however, is that the higher amounts awarded against developed states reflected the higher value of investments made in developed states as compared to developing states, which could cause similar conduct to create larger fiscal risk. The tests analyzing investors' relative success implicitly address the disparity generated by the different values of individual investments as the claimed value is reflected by the outcome variable. As such, those tests are particularly valuable, as they partially address case selection effects generated by the underlying relative value of the investment or otherwise provide a proxy for a case's merits.[161] All tests of relative investor success indicated that any potential effect was statistically small, and the results were not significant.

Given the analyses with statistically small effect sizes, the power of several tests was inevitably low.[162] Caution should therefore be used in making strong inferences about the lack of any effect, or the lack of bias within the system, as it is not possible to conclusively prove the lack of an effect. Nevertheless, the robust replication of previous research using more sophisticated methods to control for the potential impact of extraneous variables, and the existence of more than twenty analyses failing to identify a link, is additional evidence of the potential lack of a reliable relationship between development status and ITA outcomes. While the tests cannot prove the absence of a result, the results provide insight for future research and may promote understanding in the ongoing debate about ITA.

---

160. For World Bank status, the potential relationship is reflected by a negative correlation, as lower scores of development status reflected higher levels of state development. *See supra* note 156 and accompanying text. In contrast, for HDI scores, higher scores indicated higher levels of development.

161. *See supra* note 142 and accompanying text (identifying the value of focusing on relative investor success).

162. For the subset of cases where investors obtained awards in Table 9, five of the nine analyses had effect sizes that were small or less than small. This suggests that the results may not suffer from a power problem as the potential link between development status and outcomes, once democracy levels were controlled, was negligible. It also reflects that 781 total cases are necessary for sufficient statistical power to reliably identify the lack of an effect. *See also supra* note 159. Four of the nine analyses had small- to medium-sized effects. While the power of those analyses was stronger (30–40% power), it was still insufficient. To detect the largest potential effect, 120 cases involving investors who obtained damage awards would be required. As it is necessary to have a sufficiently large sample to identify the smallest potential effect, the larger sample would be required to have sufficient power to make reliable conclusions about the lack of effects.

V. DISCUSSION AND LIMITATIONS

In light of the quantitative empirical findings of Part IV, this Part first offers a synthesis of those results. It then argues that development status and democracy require disentanglement so that the policy debate does not unduly conflate the effects of politics and development. Only after separating the variance attributable to each variable and potential interactions will stakeholders be in a position to tailor their normative reforms to achieve their ultimate objectives. It then acknowledges the limitations of the analyses and recommends replication of this research.

A. Synthesis and Observations

Overall, the vast majority of the tests failed to identify any reliable link between ITA outcomes and the development status of respondent states. The only set of tests that identified a reliable link between development and outcomes used a World Bank definition and amounts awarded; but states' World Bank classifications nevertheless failed to generate a significant link to investors' relative success. When controlling for democracy levels, none of the twenty-one models analyzed could identify a reliable link between outcome and respondent development status. The results of those tests suggest that a host state's level of democracy, some aspect of domestic political infrastructure, or other variables or combinations of variables could exert more influence on ITA outcomes.[163] In keeping with the themes of the Symposium panel's focus on politics, one might hypothesize that ITA outcomes may be a function of a state's internal compliance with good governance practices, difficulties with corruption, or lack of a stable and transparent regulatory environment. To the extent that such variables relate to the cause of action and a claim's merits, a state may choose to settle or to successfully defend cases.

When analyzing ITA outcomes as a binary function of whether states won or lost, there was no reliable relationship between outcomes and development status. Analyzing development status three different ways (OECD, World Bank, and HDI) all generated uniform and robust results. Descriptively, the general trend was for states to win cases either at roughly equivalent levels to investors or at a slightly higher rate than investors. The only facial, but non-significant, pattern in the overall data was that developed states sometimes (but not always) won more than developing states. For example, High Income states won slightly more than Upper-Middle Income states, and Lower-Middle Income states won slightly more than their Upper-Middle income counterparts. The either roughly

---

163. Theoretically, there could be a series of variables affecting outcome, which was beyond the Symposium panel's focus on politics. There may, however, be a variety of other variables affecting ITA outcomes. *See supra* note 74.

equivalent or state-favorable outcomes provide intriguing evidence that
runs counter to claims of "pro-investor" bias in ITA. Rather, it suggests
that while both states and investors won at roughly equivalent levels,
perhaps some states — particularly High Income and Lower-Middle
Income states — obtained particularly "pro-state" outcomes. While some
state stakeholders may feel disadvantaged, the data suggest a more realistic
interpretation that all categories of states generally won at least half of
their cases, but two categories of states had somewhat better rates of
winning cases. While some stakeholders may perceive the comparative
advantage of multiple sub-groups of other states as "bias," it may only
reflect the general balance in ITA outcomes where states tended to win
about half the cases at a minimum. Yet it also should generate systemic
inquiry into why some categories of states obtained proportionately larger
state-favorable outcomes.

Focusing on states' relative success reveals a similar narrative. When
assessing ITA outcomes as a function of the percentage of a state's fiscal
liability (as compared to the investors' amounts claimed), it was not
possible to identify a reliable link between outcomes and development
status. The null results were robust across all models, irrespective of
whether development status was measured using OECD, World Bank, or
HDI definitions of "development status." While not conclusive, it is
further evidence that a state's relative success in ITA may operate
independently of development status, that investors' relative success may
be roughly equivalent across various different development classifications,
or there may be other variables that make an important contribution to
ITA outcomes. Beyond the inferential statistics, it is critical to observe that
investors' relative success rates were generally low, with a mean recovery
of less than 20% (a 2% median) in the set of all cases. This investor
success rate seems at odds with claims of a pro-investor bias. For the
subset of cases where investors won, there was a 36% mean
(approximately 30% median) of relative investor success. While non-
significant and statistically unreliable, one facial pattern suggested that
investors recovered relatively larger amounts against High Income states.
This result may derive from parties making more conservative claims
against developed states, but it may also reflect that investors were
unwilling to initiate ITA against a developed state or had other alternatives
to facilitate dispute settlement. A case selection effect could also create
incentives for investors to only pursue the most meritorious claims against
developed states. Although drawing definitive conclusions is premature,
the results suggest there is a phenomenon worthy of further inquiry. The
results raise questions about how investors decide to bring claims against
states and when to place their political and economic capital at risk in ITA.

The only area of divergence involved pure amounts awarded, which is a measure that neither accounts for amounts in dispute nor the potential merits of the claim. Like earlier analyses of binary outcomes or relative success, it was not possible to identify a reliable relationship between amounts awarded and two measures of development status, namely OECD membership and HDI scores. Using either of those metrics, the relative distribution of amounts awarded was reliably different. Those non-significant results were replicated even when analyzing the subset of cases where states were found liable.

Analyzing amounts awarded as a function of World Bank classification, however, generated a divergence in the data and did not replicate earlier research.[164] This suggests that, over time, the ITA system has evolved, the sample size has gained sufficient power, or previously latent effects were sufficiently large to become detectable. Irrespective of whether the group of all awards or only the subset of investor wins was analyzed, there was a statistically meaningful difference among states depending upon their World Bank classification. Namely, High Income states experienced statistically lower awards than their Upper-Middle income counterparts; other comparisons between groups had roughly equivalent amounts awarded.[165] These findings perhaps explain why there has been an emotive push to reject ITA. Heuristics may intuitively use raw amounts awarded and classification of easily discernable World Bank classifications as a "rule of thumb" to assess ITA. Focusing on World Bank development status analyses in isolation could create an instinctually credible perception of bias against the developing world. It reflects a reasonable "System 1" emotive response[166] and may contribute to the current dynamic debate about ITA's future.

Nevertheless, an over-emphasis on initial impressions can be misleading and can lead to sub-optimal decisionmaking about normative reforms. Rather, the results must be viewed in context and re-framed to override (or aid re-assessment of) intuitive judgments.[167] The vast majority of analyses — and all of the relative success analyses that partially control for

---

164. *See* Franck, *Development and Outcomes*, *supra* note 6.

165. There were, however, two potential latent effects. In one model, follow-up analyses identified that Upper-Middle Income states experienced statistically higher amounts awarded compared to Lower-Middle Income states; another model reflected that amounts awarded against Low Income states were higher than High Income states. *See supra* notes 133–34.

166. *See* Anne Van Aaken, *Behavioral International Law and Economics*, 55 HARV. INT'L L.J. 421 (2014); Tomer Broude, *Behavioral International Law*, 163 PENN. L. REV. (forthcoming 2015), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2320375##; *see also* FRANCK, MYTHS & REALITIES, *supra* note 14 (discussing the biases and heuristics literature and the potential impact upon the ITA debate); KAHNEMAN, *supra* note 41 (describing System 1 responses in depth).

167. *See* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES (Daniel Kahneman et al. eds., 1982); KAHNEMAN, *supra* note 41; Daniel Kahneman & Amos Tversky, *On the Reality of Cognitive Illusions*, 103 PSYCHOL. REV. 582 (1996).

legal merits — failed to reliably identify a significant result and generally found that any latent relationship was statistically small. Moreover, it is critical to recall the basic descriptive data that, overall, the average amount awarded using raw inflation-adjusted data was under US$20 million.[168] When excluding those cases where states were not found liable, the mean amount awarded was approximately US$45 million.[169] These figures are in contrast with assertions that "[w]hat it takes to distinguish yourself these days is a US$350 million award, minimum."[170] Moreover, when considering the distribution of data even across the World Bank categories, the distribution of liability for High Income, Lower-Middle, and Low Income states was roughly parallel in terms of median awards and general distributions.[171] Upper-Middle Income states, however, generated more variance and had a larger number of cases,[172] a higher median award, and a greater distribution of amounts awarded.

The results, however, raise important questions about how to explain the solitary set of significant results given the overall data. Procedural and substantive aspects likely contributed to the results.

Methodologically, the results could reflect the use of amounts awarded as the dependent variable; as previously identified, amounts awarded do not address amounts claimed, do not generate a standardized measure for analyzing cases, and do not provide a partial proxy to address legal merits. Additionally, the results might reflect the use of a categorical independent variable, like World Bank status, to define development status. A four-category variable compresses variance into a small set that facilitates the identification of meaningful differences, rather than lumping it bluntly into a binary variable (as with the OECD) that can mask variation. Nevertheless, the use of a more sensitive variable, namely the continuous HDI variable that also reflects broader goals of human development, never generated a meaningful link to outcomes. It is therefore possible that using the World Bank's preexisting categorical variable with four groups compressing a great deal of variance into a single category (namely Upper-Middle Income) facilitated the identification of the significant results. As a

---

168. Franck, *Conflict Management, supra* note 8, at 14 (explaining raw inflation-adjusted data generated average amounts awarded in all final cases of US$16.6 million); *see also supra* Tables 3 & 4 (indicating the winsorized inflation-adjusted data generated an average award of US$2.95 million).

169. Franck, *Conflict Management, supra* note 8, at 14 (explaining raw inflation-adjusted data generated average amounts awarded in this subset of final cases of US$45.6 million); *see also supra* Table 5 (indicating the winsorized inflation-adjusted data generated mean state liability of US$42.1 million).

170. EBERHARDT & OLIVET, *supra* note 5, at 14 (citation omitted) (internal quotation marks omitted).

171. *See supra* Figure 2.

172. This included the cluster of cases that involved a single government measure, namely Argentina's response to its 2001 currency crisis.

continuous variable, HDI is more nuanced and is likely a better proxy for assessing development status.

Substantively, the effects may also be attributable to Argentina, as more than 1/3 of the adverse awards involving Upper-Middle Income states involved Argentina, and eight of the sixteen largest awards in the dataset were against Argentina.[173] While Argentina undoubtedly generated important variance, it is not clear whether Argentina's experiences with ITA were representative or unrepresentative of ITA. Therefore, rather than attempting to answer a potentially unanswerable question, a better approach is to control for either the effect of Argentina or a related construct. As Argentina has a complex political history, it is useful to control for the real variable of interest — namely a state's internal political infrastructure. Democracy levels can be a proxy for domestic political institutions, and Polity IV scores are an established mechanism for identifying variance related to democracy levels. It was telling that, no matter which set of variables was analyzed, once democracy levels were held constant, it was impossible to identify a link between respondent development status and outcomes. Any potentially latent effects generally reflected a statistically small phenomenon.[174]

Recognizing the limitations identified earlier and in Part V.C. below, as well as the challenge of proving a negative, the vast majority of the analysis failed to find that a respondent's development status was meaningfully related to ITA outcomes. Where research cannot even establish reliable relationship, it is logically improper to suggest the existence of a causal relationship. In short, based upon the current evidence, it is incorrect to claim that respondent states' development status alone caused ITA outcomes. Future research may therefore wish to focus on other factors, such as the identity of investors, the level of experience of parties' lawyers, the impact of host country corruption, the level of good governance practices, or the scope of regulations related to Corporate Social Responsibility for investors.

B. Normative Implications

The results offer insights to inform the current ITA debate and to guide stakeholders considering the potential use of arbitration in either the TPP or TTIP negotiations. The results did not support claims that rejecting ITA is necessary given potential problems of bias against states in general

---

173. FRANCK, MYTHS & REALITIES, *supra* note 14; *see also* Franck, *Review*, *supra* note 118, at 895 (making a similar observation about the number of adverse cases against Argentina in an earlier iteration of the dataset); *supra* note 172 (identifying the cluster of Argentina currency cases).

174. In the subset of investor wins, the trend was for developed states, rather than developing states, to experience larger fiscal amounts awarded.

or against developing states in particular. Rather, the results suggest that ITA outcomes were not monolithic; a more nuanced approach to understanding outcomes is necessary to craft reform measures tailored both to generate enhanced international dispute settlement and to foster a domestic investment climate furthering sustainable development.

The data first demonstrated that claims of pro-investor bias and claims of disparate treatment of developing states — while well intentioned — may have been overly simplistic. Rejecting ITA solely on the basis of the impact upon states or potentially disparate treatment of the developing world was not supported by the data. Rather, the data offered evidence that disrupted claims of pro-investor bias in ITA. It seems odd to argue that a system of adjudication was biased against states when: (1) states won in equal or greater proportions than investors; (2) measures of central tendency indicated that investors won less than US$20 million on average overall;[175] and (3) in those cases where investors were successful, investors' relative success was roughly 30% of the amount claimed.

The data also replicated earlier analyses indicating that it was not possible to link a respondent state's development status to ITA outcomes. While the analysis still lacked sufficient power to definitively prove the lack of bias in ITA against the developing world, the overall results provided additional evidence suggesting ITA may operate independently from a respondent's development background or be attributable to variable(s) other than development status. Nevertheless, one metric suggested ITA outcomes varied according to development status — in a way that inured to the benefit of only High Income states, as compared with only Upper-Middle Income states; this finding occurred in the context of pure amounts awarded, which does not necessarily correspond with the amount claimed. The presence of one set of tests, however, suggested that, over time, the ITA caseload has evolved, and caution must be used when recommending reforms for a system in flux. Making sweeping changes — like outright rejection of ITA — while a relatively new system of international adjudication is still evolving seems somewhat extreme. A more subtle approach is to view the significant results in context, to understand the actual variance in ITA outcomes, and then to make tailored reforms to address related concerns.

Rejecting ITA outright would not fix the inevitable problems that derive from any system attempting to manage investment treaty conflict. Investment-related conflict will always exist. The question is how best to address that conflict. The core issues for consideration are: (1) how investment treaty conflict can be channeled to generate value; and (2)

---

175. Even for the subset of awards where investors won, measures of central tendency indicated typical state liability ranged from US$40–46 million. *See supra* text accompanying note 136.

should adjudication be necessary, which adjudicator(s) in the national or international regime are best situated to resolve disputes.

Regarding the first element, depriving investors of an adjudicative forum would undercut the utility of other dispute resolution systems — such as negotiation and mediation — as they exist in "the shadow of the law."[176] Without access to some adjudicative baseline, other forms of dispute resolution would likely wither as viable methods of conflict management. Rather than eliminating conflict, problems could either periodically erupt in a dramatic international relations crisis or generate downstream economic costs, possibly in the loss of foreign investment, foregone development opportunities, or higher costs to receive the investment.[177] As those possibilities are economically undesirable and may cause social and political disruption, the better course is to offer direct access to adjudication in appropriate cases, while simultaneously promoting dispute prevention.[178] The objective should be to design procedures in treaties to channel conflict into proper forums to minimize cost and maximize value.

In the second element, nothing prevents states from promoting their investment objectives through both national and international law paradigms. This is, in fact, a typical approach in international criminal law that recognizes that, although systemically complex, there is value in both domestic and international law incentives.[179]

Focusing on the complementarity[180] of national and international legal regimes offers an opportunity to strengthen investment law. Domestic

---

176. Robert Cooter et al., *Bargaining in the Shadow of the Law: A Testable Model of Strategic Behavior*, 11 J. LEGAL STUD. 225 (1982); Robert H. Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L.J. 950 (1979).

177. *See generally* ROGER FISHER & WILLIAM URY, GETTING TO YES: NEGOTIATING AGREEMENT WITHOUT GIVING IN (1st ed. 1981) (exploring ways to analyze conflict management and dispute settlement); *see also* Franck, *Dispute Systems Design*, *supra* note 8; Mariana Hernandez Crespo, *Building the Latin America We Want: Supplementing Representative Democracies with Consensus-Building*, 10 CARDOZO J. CONFLICT RESOL. 425 (2008); Mariana Hernandez Crespo G., *From Noise to Music: The Potential of the Multi-Door Courthouse (Casas de Justicia) Model to Advance Systemic Inclusion and Participation as a Foundation for Sustainable Rule of Law in Latin America*, 2012 J. DISP. RESOL. 335; Welsh & Schneider, *supra* note 8, at 3–4, 22–23.

178. Coe, Jr., *supra* note 113, at 8–9; Roberto Echandi, *Complementing Investor-State Dispute Resolution: A Conceptual Framework for Investor-State Conflict Management*, *in* PROSPECTS IN INTERNATIONAL INVESTMENT LAW AND POLICY: WORLD TRADE FORUM 270 (Roberto Echandi & Pierre Suavé eds., 2013) [hereinafter PROSPECTS]; Lucy Reed, *Synopsis of Closing Remarks, in Investor-State Disputes, supra* note 90, at 30; Margrete Stevens, *ICSID at The Crossroads: Some Thoughts and Recommendations for Improving the Dispute Settlement System, in* PROSPECTS, *supra*, at 243, 258.

179. *See* MARK A. DRUMBL, ATROCITY, PUNISHMENT, AND INTERNATIONAL LAW (2007); Alexander K.A. Greenawalt, *The Pluralism of International Criminal Law*, 86 IND. L.J. 1063 (2011); Oona Hathaway & Scott J. Shapiro, *Outcasting: Enforcement in Domestic and International Law*, 121 YALE L.J. 252 (2011).

180. *See generally* THE INTERNATIONAL CRIMINAL COURT AND COMPLEMENTARITY: FROM THEORY TO PRACTICE (Carsten Stahn & Mohamed M. El Zeidy eds., 2011) (discussing

remedies for investment disputes, offered by domestic court judges, have value — such as developing the legal capacity of local bench and bar. But this does not undermine the value of international adjudication. Where national court judges may be perceived to be (or actually are) biased in a state's or a local actor's favor, it is sensible to place decisionmaking in an adjudicator with enforceable duties of impartiality and potential civil liability for misconduct.[181] Likewise, where streamlined international enforcement is required, it is sensible to benefit from the established, efficient, and preexisting legal infrastructure of international arbitration.[182] Meanwhile, there is no empirical proof that international arbitrators make worse decisions than national court judges. Rather, there is evidence that national court judges, even those from developed legal systems, like the United States, fall prey to biases and heuristics deviating from the rational actor model.[183]

By rejecting ITA without a considered assessment of its net value in the larger international conflict management system, the desire to generate constructive reform could lead to externalities causing net harm to international economic relationships. Moreover, generating normative reform that fixes the wrong (or a non-existent) problem can lead to systemic harm that destroys the gains international investment seeks to produce.

---

complementarity in the context of international criminal law); SARAH M.H. NOUWEN, COMPLEMENTARITY IN THE LINE OF FIRE: THE CATALYSING EFFECT OF THE INTERNATIONAL CRIMINAL COURT IN UGANDA AND SUDAN (2014) (same).

181. *See, e.g.*, Susan D. Franck, *The Liability of International Arbitrators: A Comparative Analysis and Proposal for Qualified Immunity*, 20 N.Y.L. SCH. J. INT'L & COMP. L. 1 (2000) (identifying areas of arbitrator liability and immunity).

182. Reliance on an international law regime for enforcement of arbitral awards is efficient and effective. The New York Convention and ICSID Convention provide streamlined enforcement procedures that are generally unavailable for the enforcement of domestic court judgments. *See supra* notes 16–17 and accompanying text; Franck, *Legitimacy Crisis, supra* note 5, at 1547–48, 1554–55. The Hague Convention on the Choice of Courts Agreement has a narrow scope of covered agreements, provides for a narrow scope for internationally enforceable domestic court judgments, and contains multiple exceptions. *See* Jeffry Talpis & Nick Krnjevic, *The Hague Convention on Choice of Court Agreements of June 30, 2005: The Elephant that Gave Birth to a Mouse*, 13 SW. J.L. & TRADE AMERICAS 1 (2006). As there have only been three state parties that have signed or otherwise acceded since its inception, it is not in effect. *Status Table, Convention of 30 June 2005 on Choice of Courts Agreements*, HAGUE   CONFERENCE   ON   PRIVATE   INT'L   LAW, http://www.hcch.net/index_en.php?act=conventions.status&cid=98 (last updated Nov. 19, 2010).

183. *See* Chris Guthrie et al., *Blinking on the Bench: How Judges Decide Cases*, 93 CORNELL L. REV. 1 (2007); Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777 (2001); Chris Guthrie et al., *The "Hidden Judiciary": An Empirical Examination of Executive Branch Justice*, 58 DUKE L.J. 1477 (2009); Jeffrey J. Rachlinski et al., *Inside the Bankruptcy Judge's Mind*, 86 B.U. L. REV. 1227 (2006); *see also* Mark Schweizer, Kognitive Täuschungen vor Gericht (2005) (unpublished Ph.D. dissertation, University of Zurich), *available at* http://www.decisions.ch/dissertation.html#Thesen (finding similar evidence of cognition errors testing Swiss judges).

This is not to say that the current system of international investment dispute settlement is without flaws. Any system established by humans will generate error. The more critical issue is how to craft reform measures to minimize error, to address real problems, and to "fit the forum to the fuss."[184] This Article argues that generating useful normative solutions requires a careful appreciation of ITA outcomes. An over-emphasis on development status — to the exclusion of variables like state democracy levels — risks conflating the variables potentially impacting adjudication. Given that all measures of outcomes failed to reveal a link to any measure of development once the variance attributable to a host state's democracy was extracted, it seems inapposite to condemn ITA on the basis of disparate treatment. Instead, it may be prudent to re-focus the debate onto how domestic political institutions, transparency, and good governance practices contribute to the creation of ITA disputes and ultimate outcomes.[185] Alternatively, it may be appropriate to focus on other variables that have reliable and demonstrative explanatory power. Likewise, normative reforms for investment law might usefully focus on domestic regulation and other internal political reforms. In this way, states can exercise sovereignty to promote good governance practices and corporate social responsibility on the domestic level. Even if the quality of a state's internal political infrastructure does not influence ITA outcomes, focusing on domestic good governance offers critical benefits that are in a state's long-term interests, irrespective of its international law obligations.[186]

The objective in international investment law should be to create dispute resolution mechanisms that address actual, rather than perceived or intuited, problems. This, in turn, requires identifying the real variables driving ITA outcomes and restructuring conflict management processes to address derivative concerns. International investment law stakeholders should therefore consider focusing on other variables — perhaps the role of lawyers, the background of investors, the role of third-party funders, or

---

184. Frank E.A. Sander & Stephen B. Goldberg, *Fitting the Forum to the Fuss: A User-Friendly Guide to Selecting an ADR Procedure*, 10 NEGOTIATION J. 49 (1994); Deborah M. Kolb & Susan S. Silbey, *Enhancing the Capacity of Organizations to Deal with Disputes*, in NEGOTIATION THEORY AND PRACTICE 315 (J. William Breslin & Jeffrey Z. Rubin eds., 1991).

185. Such a focus suggests a host state's internal democracy or regulatory systems may ultimately be better predictors of ITA outcomes. One might imagine, for example, that more egregious government activity that violates principles of good governance — irrespective of whether this is generated from conduct of a developed or developing state — is more likely to generate a finding of state liability.

186. Beyond increasing compliance with international investment law and decreasing the risk of investors initiating ITA, focusing on good governance should increase the likelihood that states will comply with their domestic law. Moreover, predictable and transparent regulatory regimes could decrease commercial risk and thereby make a state more attractive to initial investment by foreign and domestic investors, create incentives for existing investors to expand their investment, and promote related development objectives.

the merits of individual cases — as potential factors to generate reforms to the current dispute resolution process. In the interim, as opposed to putting effort into objecting to ITA, states may wish to shift their focus to promoting and implementing good governance practices within their jurisdictions and preventing investment conflict from erupting in the first instance. An optimal result would involve states exercising their sovereignty to promote activity that complies with domestic and international law obligations, promotes a favorable investment climate, and limits exposure to unnecessary or unmeritorious claims.

Ultimately, an over-emphasis on development status alone means that implementing normative choices focused solely on a respondent state's development status misses the opportunity to construct normative solutions tailored to address real problems. The data demonstrated that, when holding a respondent state's democracy levels constant, it was not possible to reliably link ITA outcomes and development status. By focusing on those variables that meaningfully affect ITA outcomes, stakeholders can design more legitimate, targeted, and appropriate dispute resolution processes in a time of international economic transition.

C. Limitations

It is important to understand the results within their proper context, as it restrains the strength of the related inferences and scope of derivative normative suggestions. Part III.A identified a series of limitations related to case-selection bias, and Parts IV.A & B identified limitations related to statistical power deriving from null-results. It is important to reinforce these points and identify other relevant limitations.

First, there may be limitations to the strength of the inferences, as they may not reflect population parameters. Inferences should be made cautiously, particularly when considering whether the subsets analyzed are representative of ITA as a whole. Replication with additional data is necessary to avoid establishing a parameter that could reflect chance alone.

Second, there may be issues about the validity of the statistical conclusions. Effect sizes suggest that the power of the research is low, given the small size of the potentially latent effects. Yet, the statistically small effect sizes may reflect that the analyses were not underpowered, and that, even if future analyses were sufficiently powered, the relationship may be of low practical significance. Nevertheless, it would be prudent to establish a broader pool of data, based on an *a priori* power of analysis that would require more than 600 additional final ITA cases, to confirm, clarify, contradict, or supplement these findings.

Third, the statistical models used were blunt, and the variables were limited. This means that there would be utility in replicating and expanding

this research by using more complicated models and additional variables to refine both the research questions and the statistical conclusions. More complex designs analyzing additional variables and/or combinations of variables, multivariate regressions, or matching variables might help reflect the complexities of reality and simultaneously decrease the likelihood of statistical error.

Fourth, there may be issues related to external validity, since the results come from data based on publicly available awards. It is possible that publicly available awards do not represent the broader population of both public and private awards. The data used for the analysis were coded from awards that were publicly available as of January 1, 2012. Since there is now roughly two additional years' worth of data to gather and analyze, future research should replicate the analysis. Also, there are issues related to case-selection bias. While there have been anecdotal comments suggesting that the data is representative, this has not been empirically confirmed. It is possible that these analyses may be limited and/or systematically biased, particularly if there is a recent trend for investors to avoid ICSID, which may mean that fewer cases find their way into the public domain. Further research should consider the effect of this possible case-selection bias.

As a result of these cautionary considerations, more research is required to create sufficient power, stability, statistical control, and enhanced validity necessary to reach more definitive conclusions. It will likely take years — if not decades — before a sufficient pool of awards is available to run the requisite analysis. While it makes future research challenging, it does not diminish that this research replicated and expanded upon existing analyses. Nevertheless, it is fundamental to recognize the limitations of statistical power and related risks of error.

CONCLUSION

This Article has identified that democracy and development status can be conflated; once the variance from democracy is controlled for, there was not any evidence of a reliable relationship between ITA outcomes and a state's development status. Without a careful analysis of ITA outcomes, it is not possible to generate proper normative reforms. The analysis in this Article cannot be used as a basis for eradicating ITA in its entirety. While one model indicated critics of ITA have an intuitively grounded basis for concern, upon further reflection and control, those potentially concerning results dissipate. Viewing development variables in isolation can create an improper impression of systemic bias when concerns might be more appropriately focused on other variables — such as a state's internal political infrastructure or other variables.

While the dissatisfaction is understandable, care must be taken to avoid creating dispute resolution systems that generate sub-optimal outcomes or redress the wrong harm. As explaining ITA outcomes likely requires a complicated, multivariate set of other variables, international investment law should focus on designing dispute resolution systems tailored to generate meaningful solutions to remedy demonstrable problems. By shifting focus to variables that reliably contribute to the variance of ITA outcomes, ITA will generate a more legitimate and appropriate dispute resolution process in a time of international economic transition.

* * *