



2009

Development and Outcomes of Investment Treaty Arbitration

Susan D. Franck

Washington and Lee University School of Law, francks@wlu.edu

Follow this and additional works at: <https://scholarlycommons.law.wlu.edu/wlufac>



Part of the [International Law Commons](#)

Recommended Citation

Susan D. Franck, *Development and Outcomes of Investment Treaty Arbitration*, 50 Harv. Int'l L. J. 435 (2009).

This Article is brought to you for free and open access by the Faculty Scholarship at Washington and Lee University School of Law Scholarly Commons. It has been accepted for inclusion in Scholarly Articles by an authorized administrator of Washington and Lee University School of Law Scholarly Commons. For more information, please contact christensena@wlu.edu.

VOLUME 50, NUMBER 2, SUMMER 2009

Development and Outcomes of Investment Treaty Arbitration

Susan D. Franck*

The legitimacy of investment treaty arbitration is a matter of heated debate. Asserting that arbitration is unfairly tilted toward the developed world, some countries have withdrawn from World Bank dispute resolution bodies or are taking steps to eliminate arbitration. In order to assess whether investment arbitration is the equivalent of tossing a two-headed coin to resolve investment disputes, this Article explores the role of development status in arbitration outcomes. It first presents descriptive, quantitative research about the developmental background of the presiding arbitrators who exert particular control over the arbitration process. The Article then assesses how (1) the development status of the respondent state, (2) the development status of the presiding arbitrator, and (3) the interaction of these variables affect the outcome of investment arbitration. The results demonstrate that, at the macro level, development status does not have a statistically significant relationship with outcome. This suggests that the investment treaty arbitration system, as a whole, functions fairly and that the eradication or radical overhaul of the arbitration process is unnecessary. The existence of two statistically significant simple effects—namely that tribunals with presiding arbitrators from the developing world made smaller awards against developed states in particular circumstances—suggests that particularized reform could enhance the procedural integrity of arbitration. Irrespective of whether future research replicates the results, reforms targeted to redress possible imbalance in the system have the potential to enhance procedural justice and the perceived legitimacy of arbitration in an area with profound political and economic implications.

I. INTRODUCTION

Within the last decade, investment treaty arbitration has moved from a matter of peripheral academic interest to a matter of vital international concern. Investment treaty arbitration permits foreign investors to sue host governments for damages those governments allegedly caused to their investments. A typical claim might involve an investor demanding over US\$300 million from a host state for governmental action such as regulating financial markets or instituting environmental protection measures.

With a four-fold increase in the number of disputes, billions of dollars at stake, and national sovereignty and international relations on the line, investment treaty arbitration has become a vital aspect of the debate about the international political economy. Comments from U.S. presidential candi-

* Associate Professor of Law, Washington & Lee University School of Law. The author thanks José E. Alvarez, Christopher R. Drahozal, Mark A. Drumbl, David A. Gantz, Nuno Garoupa, Tom Ginsburg, Laurence R. Helfer, Timothy S. Jost, Clint Pinehardt, W. Michael Reisman, Jennifer K. Robbenolt, Amr Shalakany, Steve L. Willborn, and Jason Yackee for their comments. The author is grateful to Professor Calvin P. Garbin for his patience and assistance in refining the statistical analyses performed in this Article. The author also thanks the faculty of the University of Cincinnati College of Law and the participants at the Law and Society Association in Montreal and the Society of International Economic Law's Inaugural Conference in Geneva for their thoughts in developing this work.

dates regarding trade “time-outs,”¹ possible reconsideration of the North American Free Trade Agreement (“NAFTA”),² withdrawals from the World Bank’s International Centre for the Settlement of Investment Disputes (“ICSID”),³ and articles in the *Financial Times*⁴ put investment treaty arbitration at the core of the globalization debate. Part of the concern originates from the potential impact upon the developing world and the implications for sustainable global economic development.

There are sharp disagreements related to the legitimacy of investment treaty arbitration. The president of Bolivia asserts that developing countries in Latin America “never win the cases. The transnationals always win.”⁵ Likewise, Nicaragua openly advocates withdrawal from ICSID,⁶ potentially because of concerns related to impartiality. Similarly, Ecuador is on track to

1. Associated Press, *Clinton to Give Trade Deals a “Time Out,”* MSNBC, Nov. 12, 2007, <http://www.msnbc.msn.com/id/21756848/>.

2. *Clinton, Obama Talk Economy, War on the Stump*, CNN, Mar. 21, 2008, <http://www.cnn.com/2008/POLITICS/03/20/dems.campaign/> (discussing Clinton’s position on the North American Free Trade Agreement (“NAFTA”) and her statement that “I have been very clear about what I would do to renegotiate NAFTA”); Cathleen Decker & Mark Z. Barabak, *Obama, Clinton take on NAFTA*, L.A. TIMES, Feb. 27, 2008, at A12 (“I will say, we will opt out of NAFTA unless we renegotiate it, and we renegotiate it on terms that are favorable to all of America,” Clinton said . . . ‘I will make sure that we renegotiate in the same way that Sen. Clinton talked about, and I think actually Sen. Clinton’s answer on this one is right,’ Obama said.”). *But see* Michael D. Shear, *Obama Hoping to Reinforce U.S. Trade Relationship with Canada*, WASHINGTON POST, Feb. 19, 2009, at A3 (observing that President Obama “warned against a ‘strong impulse’ toward protectionism while the world suffers a global economic recession and said efforts to renegotiate NAFTA will have to wait”); Jonathan Weisman, *Obama, in Canada, Warns Against Protectionism*, WALL STREET JOURNAL, Feb. 20, 2009, at A4 (observing that Obama “sought to reassure Canada that he had no intention of turning some of his campaign rhetoric on trade into actual barrier”).

3. News Release, Int’l Ctr. for Settlement of Inv. Disputes (ICSID), *Bolivia Submits a Notice Under Article 71 of the ICSID Convention* (May 16, 2007), *available at* <http://www.worldbank.org/icsid/highlights/05-16-07.htm> (withdrawing from ICSID); News Release, ICSID, *Ecuador’s Notification under Article 25(4) of the ICSID Convention* (Dec. 5, 2007), *available at* <http://icsid.worldbank.org/ICSID/FrontServlet?requestType=CasesRH&actionVal=OpenPage&PageType=AnnouncementsFrame&FromPage=Announcements&pageName=Announcement9> (withdrawing from ICSID in sectors related to oil, gas, and mining); *Ecuador Withdraws from ICSID?*, BRETTON WOODS PROJECT, Dec. 4, 2007, <http://www.brettonwoodsproject.org/art-558781> (discussing Ecuador’s withdrawal from ICSID); Luis Núñez Salmerón, *Posible Salida de Nicaragua de Ciadi*, LA PRENSA, Apr. 14, 2008, <http://www-usa.laprensa.com.ni/archivo/2008/abril/14/noticias/economia/253728.shtml> (discussing Nicaragua’s potential withdrawal from ICSID); *Threats to Withdraw from Bank’s Investment Tribunal*, BRETTON WOODS PROJECT, July 2, 2007, <http://www.brettonwoodsproject.org/art-554233> (discussing potential withdrawals by Bolivia, Venezuela, and Nicaragua).

4. Alan Beattie, *Concern Grows over Global Trade Regulation*, FINANCIAL TIMES, Mar. 12, 2008, at 9.

5. James M. Roberts, *If the Real Simon Bolivar Met Hugo Chavez, He’d See Red*, BACKGROUND NO. 2062 (The Heritage Foundation, Washington, D.C.) Aug. 20, 2007, at 12; *see also* *Bolivia resolvió problemas al recuperar recursos, dijo Evo*, ABC DIGITAL, July 30, 2007, <http://www.abc.com.py/2007-06-30/articulos/340261/Bolivia%20resolvi%F3%20problemas%20al%20recuperar%20recursos,%20dijo%20Evo>; *CIADI: Justicia a la Medida de las Transnacionales*, TUNUPA: CARTA INFORMATIVA DE LA FUNDACIÓN SOLÓN, Apr. 2007, at No. 33, *available at* <http://www.funsolon.org/publicaciones/TUNUPA33CIADI.pdf> [hereinafter TUNUPA].

6. *See* Skadden, Arps, Slate, Meagher & Flom LLP, *Nicaragua Advocates Withdrawal from ICSID: Implications for Investors*, Apr. 24, 2008, http://www.skadden.com/content%5CPublications%5CPublications1391_0.pdf.

eliminate investment arbitration.⁷ Meanwhile, commentary from non-governmental organizations (“NGOs”)⁸ and academics⁹ highlights the concern about using arbitration to resolve investment treaty disputes. These concerns are motivated by apprehension about arbitration’s potential disparate impact on the developing world¹⁰ and fear that development status¹¹ might inappropriately affect outcome.¹²

These concerns about the integrity of investment treaty arbitration are worthy of consideration. Unfair treatment of respondent states on the basis of whether they are part of the developed or developing world raises tangible issues about the legitimacy and long-term viability of arbitration. Similarly, if participants believe that a dispute’s outcome depends in some part upon whether an arbitrator comes from the developing or developed world, they may question the procedural integrity of arbitration.

Systemic bias is unacceptable. As a normative matter, the resolution of international investment disputes through legalized adjudication should not depend upon the development status of the respondent state or arbitrators. Instead, international arbitration should involve unbiased, depoliticized adjudication. Robert Hudec encouraged the judicialization of international trade disputes.¹³ In the same vein, investment treaty arbitration should in-

7. *El Gobierno Terminará Contratos con Petroleras que Insistan en Llevar Sus Reclamos al Ciadi: Galo Chiriboga*, EL COMERCIO, Aug. 6, 2008, available at http://www2.elcomercio.com/noticiaEC.asp?id_noticia=212082&id_seccion=6 (quoting Ecuador’s petroleum minister as explaining withdrawals from ICSID by saying “Tener una sede para arbitrajes en Chile es una garantía para ambos, cosa que nosotros no tenemos . . . en el CIADI porque efectivamente dudamos de su imparcialidad”) [hereinafter EL COMERCIO].

8. Food and Water Watch has asserted that investment treaty arbitration “rules are weighted heavily in favor of global corporations and against the mostly poor countries caught up in disputes . . . 93 percent of the cases [at ICSID] involve low- or middle-income developing countries . . . [and] ICSID tribunals have ruled in favor of the investor and ordered the government to pay compensation in nearly 70 percent of cases.” Press Release, Food and Water Watch, World Bank Court Grants Power to Corporations (Apr. 30, 2007), available at <http://www.foodandwaterwatch.org/press/releases/world-bank-court-grants-power-to-corporations-article12302007> [hereinafter Food and Water Watch]; see also 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 (2007), available at <http://www.ips-dc.org/reports/070430-challengingcorporatinvestorrule.pdf> [hereinafter ANDERSON & GRUSKY].

9. See generally ANDERSON & GRUSKY, *supra* note 8.

10. See generally TUNUPA, *supra* note 5.

11. The present research defined the development dimension in terms of membership in the Organization for Economic Cooperation and Development (“OECD”) and the World Bank’s four-point development scale. See *infra* Part IV.

12. See *infra* Part III.

13. Robert E. Hudec, *The Judicialization of GATT Dispute Settlement*, in IN WHOSE INTEREST? DUE PROCESS AND TRANSPARENCY IN INTERNATIONAL TRADE (Michael M. Hart & Debra P. Steger eds., 1992); see also JOHN H. JACKSON ET AL., LEGAL PROBLEMS OF INTERNATIONAL ECONOMIC RELATIONS 12 (5th ed. 2008); Susan D. Franck, *Foreword to A Symposium Exploring the Modern Legacy of William Jennings Bryan*, 86 NEB. L. REV. 142, 144–47 (2007) (discussing the shift away from gunboat diplomacy and war and the increasing use of international arbitration to resolve international disputes involving governments, particularly those in an economic context); David M. Trubek, *Transcending the Ostensible: Some Reflections on Bob Hudec as Friend and Scholar*, 17 MINN. J. INT’L L. 1, 3–4 (2008); Andrea Kupfer Schnei-

volve the neutral application of facts to mutually agreed-upon legal principles. This should lead to outcomes based on law rather than factors unrelated to the merits of the case. There will be political and economic implications if there is a reliable, let alone causal, link between investment treaty arbitration and variables associated with development status. If governments believe the deck may be stacked against them or that arbitration is the equivalent of tossing a two-headed coin, they may refuse to negotiate investment treaties altogether or eliminate arbitration in treaties they do negotiate. Likewise, investors may lose faith in the arbitration process and the commercial value of predictable dispute resolution, which may in turn affect decisions to invest or increase the cost of investment. Meanwhile, because the system is perceived to lack integrity, civil society groups may articulate concerns through methods ranging from organized programmatic critique to civil unrest. It is therefore vital to understand the development dimensions of the current arbitration process in order to examine the integrity of the dispute resolution system. This will also enable assessment of the appropriate level of confidence to place in a system with profound public and international implications.

Empirical methodologies can help assess linkages between development status and outcome. Although the use of empirical perspectives can be controversial, carefully conducted and transparently described empirical research is a powerful tool to test assertions about the efficacy of the arbitration process, draw inferences about the broader population, and contextualize examples.¹⁴ For investment treaty arbitration, empiricism permits the isolation of variables in order to gauge the relationship (or lack thereof) with outcome.

This Article evaluates whether there is a statistically significant¹⁵ relationship between development status and arbitration outcome, development status and amounts awarded, or both.¹⁶ The study's objective is to explore whether arbitration inappropriately favors either the developed or the devel-

der, *Not Quite a World Without Trials: Why International Dispute Resolution is Increasingly Judicialized*, 1 J. DISP. RESOL. 119, 119–24 (2006) (arguing that certain international disputes are increasingly judicialized).

14. Susan D. Franck, *Empiricism and International Law: Insights for Investment Treaty Dispute Resolution*, 48 VA. J. INT'L L. 767 (2008) [hereinafter Franck, *Empiricism*].

15. Statistical significance refers to "the likelihood, or probability, that a statistic derived from a *sample* represents some genuine phenomenon in the *population*." TIMOTHY C. URDAN, STATISTICS IN PLAIN ENGLISH 58 (2d ed. 2005) [hereinafter URDAN]. It also involves an assessment of whether phenomena are the result of chance. See Wendy Parker, *Lessons in Losing: Race Discrimination in Employment*, 81 NOTRE DAME L. REV. 889, 914 n.116 (2006) ("[S]tatistical significance indicates the likelihood that a particular outcome occurred by chance . . ."); see also Lee Epstein & Gary King, *The Rules of Inference*, 51 U. CHI. L. REV. 1, 109–13 (2002) (discussing the role that chance can play in empirical legal research).

16. This research does not evaluate differences in whether investors come from the developed or developing world because approximately 10% of investors were from the developing world. Susan D. Franck, *Empirically Evaluating Claims About Investment Treaty Arbitration*, 86 N.C. L. REV. 31 (2007) [hereinafter Franck, *Evaluating Claims*]. This small percentage hinders the creation of reliable statistical models. As the population expands, future research should explore the development dimension of investors.

oping world, and whether a presiding arbitrator's developmental status is associated with outcome. Such an analysis offers the opportunity to consider whether (1) the arbitration process inappropriately favors the developed or the developing world, (2) arbitrators from the developed or the developing world exert undue influence on the process, or (3) these factors apply in combination.

Part II provides a background on investment treaty arbitration. Part III considers the existing literature on the role of the development dimension on investment treaty arbitration. Part IV then discusses the research hypotheses and methodologies.

Part V describes the initial results and explains the research limitations. Part V.A discusses the development status of presiding arbitrators, who control the arbitration process and are either the chairs of three-member tribunals or sole arbitrators. There were arbitrators from both the developed and developing world,¹⁷ but there were no presiding arbitrators from countries the World Bank classifies as Low Income.

Part V.B describes the lack of a statistically significant relationship between development status and the ultimate winners of investment arbitration. Part V.C likewise describes the lack of a statistically significant relationship between amounts awarded and the development status of respondents, the development status of presiding arbitrators, or even an interaction between those two variables. Two statistically significant simple effects,¹⁸ found in only one follow-up analysis with a small subset of potentially non-representative cases, suggested that tribunals with presiding arbitrators from Middle Income countries awarded different damages in limited circumstances: where there was a presiding arbitrator from a Middle Income country, High Income countries experienced statistically lower awards than both (1) Upper Middle income countries, and (2) Low Income countries.¹⁹ Part V.D synthesizes these results to suggest that, in general, development variables do not inappropriately affect the outcome of investment arbitra-

17. Frequently, investors choose one arbitrator and the government selects the second. The presiding arbitrator, however, is selected in a different manner. The ICSID Convention allows the parties to agree on the presiding arbitrator. Absent such agreement, the Convention provides a default rule. Convention on the Settlement of Investment Disputes Between States and Nationals of Other States art. 37(2)(b), Mar. 18, 1965, 17 U.S.T. 1270 [hereinafter ICSID Convention]. In ad hoc United Nations Commission on International Trade Law ("UNCITRAL") arbitration, the party-appointed arbitrators choose the chair. UNCITRAL Arbitration Rules art. 7(1), Apr. 28, 1976, 15 I.L.M. 701, 705.

18. A simple effect "is the effect of one factor conditioned on the level of a second factor." ANDREW F. HAYES, *STATISTICAL METHODS FOR COMMUNICATION SCIENCE* 439 (2005). In this research, it involved looking beyond the main effects of the development status of the presiding arbitrator and respondent state. Follow-up pairwise comparisons considered particular combinations (i.e., simple effects) to analyze how the amounts awarded varied and to assess whether the main effects masked otherwise meaningful differences.

19. See *infra* Part V.C.3 (describing two simple effects in which Middle Income arbitrators rendered different awards against: (1) a High Income versus an Upper-Middle income respondent state (higher awards rendered against Upper-Middle Income respondents) and (2) a High Income versus a Low Income respondent state (higher awards against Low Income respondents)).

tion. It cautions, however, that proper contextualization and replication of the research is necessary.

Part V.E then describes the limitations of the research. This study is one of the first steps in the larger process of empirically assessing arbitration outcomes. Yet this initial work makes several contributions. First, it serves as the basis for future research that involves a larger dataset, introduces more variables, and offers more sophisticated models to tease out nuances that create variation in the results. Second, it provides an instructive baseline that may permit future researchers to assess population parameters. Third, recognizing the limited nature of the inferences, the research can provide information to stakeholders, such as government officials negotiating treaties, who may wish to consider the potential implications for the design of their dispute resolution systems.

Part VI makes preliminary recommendations for the design of dispute resolution systems. Given that the data suggest the system is functioning reasonably well and development status is generally not associated with disparate outcomes at the macro level, Part VI.A argues that radical overhaul or rejection of the international arbitration system is unwarranted. In light of the two micro results (the simple effects), Part VI.B suggests that there is room for improvement and urges ongoing monitoring of the arbitration process. Changes could involve improving the application of the existing arbitration by building the capacity of arbitrators from the developing world. It might also involve more normative solutions. This might include reconsidering the optimal dispute resolution mechanism(s) for resolving disputes under investment treaties. It may also involve either the creation of structural safeguards, such as an arbitrator database or legal advice center, or legislative reforms to revise the terms of investment treaties in order to minimize arbitrator discretion and provide greater guidance about how to award damages. Irrespective of whether the two micro results are replicable, Part VI.C offers cautionary guidance for stakeholders to consider during the arbitration process. Recognizing the limitations of the research, this Part argues that reform has the benefit of promoting procedural legitimacy by addressing concerns related to perceptions about the system's fairness.

The Article ultimately concludes that, although further research is necessary, there are reasons to be cautiously optimistic about the system of investment treaty arbitration and its relationship with development variables. Addressing possible areas for targeted improvement may also increase the perceived legitimacy of the dispute resolution process. Ultimately, the use of empirical methodologies can gather information to test assertions about the integrity of arbitration and permit stakeholders to make more informed choices in an area with profound political and economic implications.

II. INVESTMENT TREATIES AND INVESTMENT TREATY ARBITRATION

Foreign investment is a vital aspect of the international political economy. Projections suggest that foreign investment inflows will be close to US\$1.4 trillion by 2010.²⁰ Foreign investment has a critical impact on the world economy and development,²¹ and there is keen competition in the developed and developing world²² to attract investment.²³ Traditional methods to lure foreign investment involve liberalizing an economic sector, providing tax incentives,²⁴ and improving dispute resolution systems.²⁵ Another potential method of promoting foreign investment involves signing an investment treaty.²⁶

An investment treaty is an agreement between two or more governments that safeguards investments made by qualifying investors in the territory of

20. ECONOMIST INTELLIGENCE UNIT, *WORLD INVESTMENT PROSPECTS TO 2010: BOOM OR BACKLASH?* 6 (2006), available at http://www.cpii.columbia.edu/pubs/documents/WIP_to_2010_SPECIAL_EDITION.pdf. Data from 2006 suggests that foreign investment inflows to the developed world were approximately US\$857 billion, approximately US\$379 billion to developing countries, and approximately US\$69 billion to transition economies. U.N. Conference on Trade and Dev. [UNCTAD], *World Investment Report 2007: Transnational Corporations, Extractive Industries and Development*, ¶1, U.N. Doc UNCTAD/WIR/2007 (2007), available at http://www.unctad.org/en/docs/wir2007overview_en.pdf.

21. R. DOAK BISHOP ET AL., *FOREIGN INVESTMENT DISPUTES: CASES, MATERIALS AND COMMENTARY* 2–7 (2005); *ENCYCLOPEDIA OF PUBLIC INTERNATIONAL LAW*, VOL. II 435–37 (1995).

22. Between 1992 and 2001, foreign investment flows to the developed world appeared higher than the developing world. Tom Ginsburg, *International Substitutes for Domestic Institutions: Bilateral Investment Treaties and Governance*, 25 *INT'L REV. L. & ECON.* 107, 108–09 (2005). It is not clear whether the data controlled for currency and valuation fluctuations. One might imagine, for example, that the same cement factory worth US\$10 billion in the United States may have a significantly different value in Liberia.

23. See Robert O. Keohane & Van Doorn Ooms, *World Politics and International Economics: The Multinational Firm and International Regulation*, 29 *INT'L ORG.* 169, 170 (1975) (“[A]most every government in the world . . . attempts to entice foreign capital”); Malcolm J. Rogge, *Towards Transnational Corporate Accountability in the Global Economy: Challenging the Doctrine of Forum Non Conveniens* in *In Re: Union Carbide, Alfaro, Sequihua, and Aguinda*, 36 *TEX. INT'L L.J.* 299, 314 (2001) (“[G]overnments in both rich and poor nations compete in a ‘race to the bottom’ to attract needed foreign investment.”). There is evidence that this competition is particularly prevalent among developing countries. See Teresa Edwards, *The Relocation of Production and Effects on the Global Community*, 13 *COLO. J. INT'L ENVTL. L. & POL'Y* 183, 190 (2002) (“[G]overnments of these less-developed countries, in an effort to attract foreign investment, compete with one another to make their standards the most attractive to the investing companies.”); Lawrence Jahoon Lee, *Barcelona Traction in the 21st Century: Revisiting Its Customary and Policy Underpinnings 35 Years Later*, 42 *STAN. J. INT'L L.* 237, 265 (2006) (“[D]eveloping states compete with each other to attract foreign investment.”).

24. The World Bank suggests that governments might provide fiscal incentives (like tax concessions or subsidies), improve domestic infrastructure, promote a skilled labor force, establish agencies to promote foreign investment, improve the regulatory environment, or enter into international agreements. See THE WORLD BANK, *WORLD DEVELOPMENT REPORT 2005: A BETTER INVESTMENT CLIMATE FOR EVERYONE* 20 (2004), available at http://siteresources.worldbank.org/INTWDR2005/Resources/complete_report.pdf.

25. *Effective ADR Mechanism Can Fetch More FDI than China*, EXPRESS INDIA, Nov. 5, 2005, available at <http://www.expressindia.com/fullstory.php?newsid=57809> (discussing how a law minister in India believed that an effective alternative dispute resolution system would help them attract more foreign investment than China).

26. Susan D. Franck, *Integrating Investment Treaty Conflict and Dispute Systems Design*, 92 *MINN. L. REV.* 161, 171 (2007) [hereinafter Franck, *Dispute Systems Design*].

other signatories.²⁷ For example, the United States and Ukraine might enter into a bilateral investment treaty,²⁸ or a group of countries might sign a regional trade agreement such as the Central American Free Trade Agreement (“CAFTA”).²⁹ These treaties grant reciprocal investment rights, both procedural and substantive, to private investors from the signatory countries.

Substantively, governments guarantee investors certain treatment, such as the right to be free from expropriation without just compensation, the right to be free from discrimination on the basis of nationality, the right to fair and equitable treatment, or the guarantee that states will honor their contractual commitments.³⁰ Procedurally, the existence of an investment treaty means that if investors believe their substantive rights have been violated they can seek redress against the host state through the treaty’s dispute resolution mechanism. The objective of these procedural rights is to move beyond war, gunboat diplomacy, and politicized forms of dispute resolution to provide a neutral forum for the resolution of investment conflicts.³¹ Investors might have an opportunity to engage initially in non-binding dispute resolution or to resolve disputes finally through arbitration.³²

27. *Id.* at 171.

28. See Treaty Between the United States of America and Ukraine Concerning the Encouragement and Reciprocal Protection of Investment, U.S.-Ukr., Mar. 4, 1994, S. TREATY DOC. NO. 103-37, 103 (1994).

29. Multilateral agreements, like NAFTA and Central American Free Trade Agreement (“CAFTA”) function in the same way as bilateral investment treaties (“BITs”) but provide investment protection on a multilateral basis. Dominican Republic-Central America-United States Free Trade Agreement art. 10, Aug. 5, 2004, available at http://www.ustr.gov/Trade_Agreements/Bilateral/CAFTA/CAFTA-DR_Final_Texts/Section_Index.html [hereinafter CAFTA-DR]; North American Free Trade Agreement, U.S.-Can.-Mex. ch. 11, Dec. 17, 1992, 32 I.L.M. 605 (1993); see also Antonio R. Parra, *Provisions on the Settlement of Investment Disputes in Modern Investment Laws, Bilateral Investment Treaties and Multilateral Instruments on Investment*, 12 ICSID REV. FOREIGN INVESTMENT L.J. 287 (1997) [hereinafter Parra, *Provisions*]. Multilateral agreements tend to address issues beyond investment protection and may address issues such as rules of origin, customs obligations, sanitary and phytosanitary measures, and cross-border trade in services. See North American Free Trade Agreement, U.S.-Can.-Mex. arts. 4-5, 7, Dec. 17, 1992, 32 I.L.M. 289 (1993); North American Free Trade Agreement, U.S.-Can.-Mex. ch. 15, Dec. 17, 1992, 32 I.L.M. 605 (1993); CAFTA-DR, *supra*, arts. 4, 5, 7, 11.

30. Franck, *Dispute Systems Design*, *supra* note 26, at 172.

31. HOWARD MANN & KONRAD VON MOLTKE, NAFTA’S CHAPTER 11 AND THE ENVIRONMENT 5-6 (1999) (suggesting that the “presumption behind the [investor-state arbitration] process is that foreign investors do not generally receive fair treatment in domestic courts in developing countries when complaining of a government action”); Charles N. Brower & Lee A. Steven, *Who Then Should Judge?: Developing the International Rule of Law Under NAFTA Chapter 11*, 2 CHI. J. INT’L L. 193, 196 (2001) (demonstrating that “the fundamental reason that the great majority of modern investment protection treaties have opted for international adjudication is that domestic courts are often in fact, and just as important, usually are perceived to be, biased against alien investors”); see also Catherine A. Rogers, *The Arrival of the “Have-Nots” in International Arbitration*, 8 NEV. L.J. 341, 356-57 (2007) [hereinafter Rogers, *Have-Nots*]; Andrea Kupfer Schneider, *Getting Along: The Evolution of Dispute Resolution Regimes in International Trade Organizations*, 20 MICH. J. INT’L L. 697, 717 (1999) (stating that investors who are “concerned with the potential bias, inefficiency, or unfamiliarity of foreign courts” are likely to prefer the investor-state arbitration regime).

32. Franck, *Dispute Systems Design*, *supra* note 26, at 172-73, 192-94.

R

R

While it varies by treaty, investors can generally elect to arbitrate before one or more of the following: (1) an *ad hoc* tribunal organized under the United Nations Commission on International Trade Law (“UNCITRAL”) Arbitration Rules, (2) the Stockholm Chamber of Commerce, or (3) a tribunal organized through the World Bank’s ICSID.³³

The mechanics of arbitration are relatively straightforward. Investors initiate arbitration by submitting a Request for Arbitration to their selected forum. Then, the process of selecting a tribunal begins. Typically panels of three arbitrators resolve investment disputes.³⁴ The investor selects one arbitrator and the respondent state picks a second arbitrator. The default rules for selecting the final arbitrator, the presiding arbitrator or chair, vary according to the institution chosen.³⁵ At ICSID, parties can agree on the appointment of the presiding arbitrator, and where the parties cannot agree, ICSID makes the final appointment.³⁶ In contrast, in *ad hoc* UNCITRAL arbitration, party-appointed arbitrators agree on the presiding arbitrator.³⁷

All arbitrators are generally required to be impartial and to contribute to the adjudicatory outcome.³⁸ Nevertheless, the presiding arbitrator “performs a different role than the party-appointed arbitrator,”³⁹ and his or her ap-

33. Parra, *Provisions*, *supra* note 29, at 288 (remarking that investors’ options may vary based on the rights enumerated in the applicable treaty as some treaties provide investors with a full range of options when seeking recourse, including the national courts, *ad hoc* arbitration under the UNCITRAL Rules, arbitration before the ICSID, the International Chamber of Commerce (ICC), or the Stockholm Chamber of Commerce (SCC)). Certain investment treaties have limited mechanisms for resolving disputes. *See e.g.*, Agreement Concerning the Encouragement and Reciprocal Protection of Investments, P.R.C.-Ghana, art. 10, Oct. 12, 1989, http://www.unctad.org/sections/dite/ia/docs/bits/china_ghana.pdf (providing that certain investment disputes are subject to *ad hoc* arbitration, the SCC is the default appointing authority, and the tribunal can use either the SCC or ICSID rules “as guidance”); Treaty Concerning the Encouragement and Reciprocal Protection of Investments, U.S.-Morocco, art. 6, July 22, 1985, S. TREATY DOC. No. 99-18 (noting that ICSID has exclusive jurisdiction over investment disputes); Treaty Concerning the Reciprocal Encouragement and Protection of Investment, U.S.-Haiti, art. 7, Dec. 13, 1983, http://www.unctad.org/sections/dite/ia/docs/bits/us_haiti.pdf (authorizing only the ICC to resolve investment disputes); Agreement for the Promotion and Protection of Investments, art. 8, Gr. Brit.-Sing., July 22, 1975, 1018 U.N.T.S. 175 (providing that ICSID has exclusive jurisdiction over investment disputes). Other treaties allow parties to select the arbitration method. *See e.g.*, Agreement for the Reciprocal Promotion and Protection of Investments, Egypt-Pol., art. 8(4), July 1, 1995, http://www.unctad.org/sections/dite/ia/docs/bits/egypt_poland.pdf (allowing disputes to be settled by SCC arbitration, ICC arbitration, *ad hoc* arbitration under UNCITRAL Rules, or ICSID arbitration); Treaty Concerning Business and Economic Relations, U.S.-Pol., art. 9, Mar. 21, 1990, http://www.unctad.org/sections/dite/ia/docs/bits/US_poland.pdf (giving investors the choice of arbitrating through the ICSID Additional Facility Rules, through *ad hoc* UNCITRAL based arbitration, or using the rules of “any arbitral institution” to which both parties agree).

34. Franck, *Evaluating Claims*, *supra* note 16, at 77.

35. RUDOLF DOLZER & MARGRETE STEVENS, *BILATERAL INVESTMENT TREATIES* 124 (1995).

36. Under the ICSID Convention, parties can agree on the appointment of the tribunal president. ICSID Convention, *supra* note 17, art. 29.

37. UNCITRAL Arbitration Rules art. 7(1), Apr. 28, 1976, 15 I.L.M. 701 (1976).

38. Claudia T. Salomon, *Selecting an International Arbitrator: Five Factors to Consider*, 17 MEALEY’S INT’L ARB. REP. Nos. 10, 2, 3 (Oct. 2002), available at <http://www.arbitralwomen.org/files/publication/0405202743129.pdf>.

39. *Id.*; see Toby Landau, *Composition and Establishment of the Tribunal*, 9 AM. REV. INT’L ARB. 45, 52–53 (1998) [hereinafter Landau]; see also Andreas Lowenfeld, *The Party-Appointed Arbitrator in International Controversies: Some Reflections*, 30 TEX. INT’L L.J. 59, 65 (1995) (discussing how party-appointed

R

R

R

pointment is a matter of vital importance.⁴⁰ The presiding arbitrator can “influence the style of an international arbitration”⁴¹ and make critical procedural decisions.⁴² Some suggest that presiding arbitrators resolve disputes between party-appointed arbitrators⁴³ and, in some cases, become the ultimate decision makers.⁴⁴ For these reasons, the role of the presiding arbitrator is of particular interest.

Once the tribunal is constituted fully, the parties gather evidence and present arguments. The tribunal then renders an award on the merits of the dispute that is enforceable worldwide.⁴⁵

Some investment treaty conflicts can involve political elements, such as the “Cochabamba Water Wars,” in which the privatization of water and sewer services in Bolivia led to social unrest, protests, deaths, and the imposition of martial law.⁴⁶ Other investment treaty disputes might seem more private, such as the governmental revocation of a private banking license or a breach of a commercial contract to which the government is a party.⁴⁷ Irrespective of how politicized the conflict may appear, the cause of action generally involves: (1) foreign investors asserting that a host government has behaved in a manner that violates the treaty and has damaged their investment, and (2) if the dispute is not otherwise resolved, investors seek redress by requiring the government to arbitrate the dispute.

arbitrators must carefully consider the representations of the appointing party and also serve as translators of the parties’ legal culture).

40. See generally Landau, *supra* note 39.

41. Urs Martin Laeuchli, *Civil and Common Law: Contrast and Synthesis in International Arbitration*, 62 DISP. RESOL. J. 81, 82 (2007).

42. See Michael Black et al., *Arbitration of Cross Border Disputes*, 27 CONSTR. L. 5, 17 (2007) (observing how the chair of a tribunal permitted the introduction of documents that were arguably privileged); Winston Stromberg, *Avoiding the Full Court Press: International Commercial Arbitration and Other Global Alternative Dispute Resolution Processes*, 40 LOY. L.A. L. REV. 1337, 1367 (2007) (observing how tribunals may rely on “testimonial summaries prepared by the presiding arbitrator and presented to the witness for approval and signature”).

43. Thomas Carbonneau explains that presiding arbitrators sometimes “[resolve] the disagreement between the two party-appointed arbitrators.” *The Exercise of Contract Freedom in the Making of Arbitration Agreements*, 36 VAND. J. TRANSNAT’L L. 1189, 1211–12 (2003).

44. *Id.*; Lawrence W. Newman, *A Practical Assessment of Arbitral Dispute Resolution*, in LEX MERCATORIA AND ARBITRATION: A DISCUSSION OF THE NEW LAW MERCHANT 5–6 (Thomas E. Carbonneau ed., rev. ed. 1998) (suggesting that when parties appoint arbitrators who “blatantly will favor one side” this can polarize the tribunal and “leav[e] the chair to decide”); see also YVES DEZALAY & BRYANT G. GARTH, DEALING IN VIRTUE 8–9 (1996) (discussing how the selection of the chair is a “key decision in winning or losing”); Hans Smit, *Comments on Public Policy in International Arbitration*, 13 AM. REV. INT’L ARB. 65, 67 (2002) (recommending that “if no majority decision seems possible, the chair’s vote is decisive”).

45. Franck, *Dispute Systems Design*, *supra* note 26, at 193–94.

46. OSCAR OLIVERA & TOM LEWIS, ¡COCHABAMBA! WATER WAR IN BOLIVIA 33–47 (2004). The social protests related to the privatization of the water sector may be an unrepresentative example of the possibility of social unrest. It is nevertheless an interesting case study.

47. Franck, *Dispute Systems Design*, *supra* note 26, at 185–86; Franck, *Evaluating Claims*, *supra* note 16, at 10.

R

R

R

III. THE DEVELOPMENT DIMENSION IN INVESTMENT TREATY ARBITRATION

There is a growing empirical literature examining whether investment treaties achieve their purported benefits, namely the proliferation of foreign investment and development.⁴⁸ In contrast, despite the interest,⁴⁹ there is little empirical analysis systematically assessing investment treaty dispute resolution.⁵⁰ Investment treaty arbitration is nevertheless ripe for empirical analysis.⁵¹ One area drawing attention is the impact of the development dimension on the outcome of disputes.⁵² The development dimension can manifest itself in various ways, such as the development status of respondent states and arbitrators.

The development dimension raises concerns about both respondents and decision makers. Regarding respondents, there are concerns about whether the host government's development status inappropriately influences the outcome of investment treaty arbitration. The concerns relate primarily to whether the developing world is unfairly harmed or whether the developed

48. See generally Franck, *Empiricism*, *supra* note 14, at 793; Jason Webb Yackee, *Bilateral Investment Treaties, Credible Commitment, and the Rule of (International) Law: Do BITs Promote Foreign Direct Investment?*, 42 L. & SOC'Y REV. 805 (2008); see also Peter Egger & Michael Pfaffermayr, *The Impact of Bilateral Investment Treaties on Foreign Direct Investment*, 32 J. COMP. ECON. 788 (2004); Eric Neumayer & Laura Spess, *Do Bilateral Investment Treaties Increase Foreign Direct Investment in Developing Countries?*, 33 WORLD DEV. 1567 (2005); Jeswald W. Salacuse & Nicholas P. Sullivan, *Do BITs Really Work?: An Evaluation of Bilateral Investment Treaties and Their Grand Bargain*, 46 HARV. INT'L L.J. 67 (2005); Deborah L. Swenson, *Why Do Developing Countries Sign BITs?*, 12 U.C. DAVIS J. INT'L L. & POL'Y 131 (2005); Mary Hallward-Driemeier, *Do Bilateral Investment Treaties Attract FDI? Only a Bit . . . and They Could Bite*, 1 (World Bank, Policy Research Working Paper No. 3121, 2003); Jennifer Tobin & Susan Rose-Ackerman, *When BITs Have Some Bite: The Political-Economic Environment for Bilateral Investment Treaties* (Nov. 14, 2006) (unpublished article, on file with The Yale Law Journal). But see M. Sornarajah, *Power and Justice: Third World Resistance in International Law*, 10 SING. Y.B. INT'L L. 19, 32 (2006) [hereinafter Sornarajah, *Third World Resistance*] (suggesting that there is no empirical foundation that investment treaties achieve their purported objectives, but failing to cite any empirical literature for that assertion).

49. Some political scientists express interest in empirical research related to investment treaty arbitration. See, e.g., Franck, *Empiricism*, *supra* note 14, at 802; Clint Peinhardt & Todd Allee, *The International Centre for Settlement of Investment Disputes: A Multilateral Organization Enhancing a Bilateral Treaty Regime* (Apr. 14, 2006) (paper presented at the Annual Meeting of the Midwest Political Science Association), available at <http://www.utdallas.edu/~cwp052000/mpsa.peinhardt-allee.pdf>; Todd Allee & Clint Peinhardt, *Delegating Differences: Bilateral Investment Treaties and Patterns of Dispute Resolution Design*, INT. ORG. (forthcoming).

50. See Franck, *Evaluating Claims*, *supra* note 16. A recent study considers the interpretive modalities of investment treaty arbitral tribunals and compares the reasoning of those tribunals to that of other international adjudicative bodies. Ole Kristian Fauchald, *The Legal Reasoning of ICSID Tribunals—An Empirical Analysis*, 19 EUR. J. INT'L L. 301 (2008).

51. Posting of David Zaring to *Opinio Juris*, <http://opiniojuris.org/2008/07/03/empirical-research-and-international-economic-law-a-comment-on-susan-franck%e2%80%99s-essay> (July 3, 2008, 08:56 EST).

52. See Luke Eric Peterson, *UK Bilateral Investment Treaty Program and Sustainable Development: Implications of Bilateral Negotiations on Investment Regulation at a Time When Multilateral Talks Are Faltering*, BRIEFING PAPER (The Royal Inst. of Int'l Affairs, Sustainable Development Programme, London), Feb. 2004, at 2, available at http://www.chathamhouse.org.uk/files/3160_binvestfeb04.pdf [hereinafter Peterson, *UK BIT Program and Sustainable Development*] (“[I]t is apparent that the treaties may harbour wide and potentially serious implications for sustainable development.”).

R

R

R

world is unfairly privileged. There are also concerns about decision makers, particularly presiding arbitrators given their unique role in the process,⁵³ and whether those arbitrators are influenced by the development status of their home jurisdictions. In other words, there is concern that arbitrators from the developed world may assess and decide cases differently than those from the developing world. It is precisely these types of concerns that the current research intends to assess. The remainder of this Part explores the literature related to aspects of development status and its potential relationship to the outcome of investment treaty arbitration.

A. *The Development Status of Respondent States*

The effect of investment treaty arbitration on the developed and the developing worlds has been a subject of controversy. There are three aspects of the debate. The first is an assessment of which governments in the developed and developing worlds have actually participated in investment treaty arbitration. The second is an evaluation of whether governments or investors win cases. The third is an examination of whether amounts awarded vary according to whether the government is from the developed or developing world.

Some commentators suggest that “mostly poor countries [get] caught up in disputes . . . 93 percent of the cases [at ICSID] involve low- or middle-income developing countries.”⁵⁴ In contrast, the United Nations Conference on Trade and Development (“UNCTAD”) suggests that approximately sixty percent of respondent states came from the developing world.⁵⁵ Given methodological shortcomings, such as a failure to explain the origin of data or the basis for classifying development, researchers should use caution when making inferences related to these claims.⁵⁶

There is another quantitative empirical analysis that describes its research methodology and permits evaluation of the validity of the research conclusions.⁵⁷ That research suggests—as do UNCTAD’s findings—that the presence of more developed countries in investment arbitration is not *de minimis*. The research used two different pre-existing measures to assess development: membership in the Organization for Economic Co-Operation and Develop-

53. See *supra* notes 38–44 and accompanying text.

54. Food and Water Watch, *supra* note 8; see also ANDERSON & GRUSKY, *supra* note 8. Others echo this theme. See, e.g., Emad Mekay, *Bias Seen in International Dispute Arbiters*, INTER PRESS SERVICE, June 19, 2007, <http://ipsnews.net/news.asp?idnews=38229> (“74 percent of concluded and pending cases were filed against ‘middle-income developing countries’ and 19 percent against ‘low-income developing countries.’ Only 1.4 percent of all cases were filed against nations from the powerful Group of Eight most industrialised nations.”).

55. United Nations Conference on Trade and Dev. [UNCTAD], *Investor-State Disputes Arising from Investment Treaties: A Review*, 6, U.N. Doc. UNCTAD/ITE/IIT/2005/4 (2005) (prepared by Jörg Weber & Peter Muchlinski).

56. Franck, *Evaluating Claims*, *supra* note 16, at 31 nn.138–39.

57. See generally *id.*

ment (“OECD”) and World Bank development classifications.⁵⁸ Approximately one-third of respondent states were OECD Member States.⁵⁹ Using the World Bank classification system, respondents consisted of: (1) High Income countries (18%), (2) Upper-Middle Income countries (45%), (3) Lower-Middle Income countries (28%), and (4) Low Income countries (less than 9%).⁶⁰ This composition suggests both the developed and developing world have participated in arbitration.

The next issue is the arbitration outcome, or whether governments won or lost. Some express hope that investors’ treaty claims “may well succeed for good reasons, such as malfeasance, abuse or other mistreatment perpetrated by a host government.”⁶¹ Others disagree and imply that governments cannot win investment treaty arbitration.⁶²

Recent systematic, descriptive quantitative research makes several points. First, governments can and did win investment disputes. In fact, governments (57.7%) were more likely than investors (38.5%) to win cases and have no damages awarded for alleged treaty breaches.⁶³ Second, the average amount awarded (approximately US\$10 million) was a fraction of what investors typically requested (approximately US\$343 million).⁶⁴

This leads to the third issue. If respondents come from both the developed and developing worlds and governments have won, were the outcomes somehow linked to the respondent states’ development status in a manner that is unfair or systematically biased? In other words, is arbitration the equivalent of tossing a two-headed coin to resolve investment disputes?

There are no empirical studies on this issue, and it is precisely this gap in the literature that this research aims to fill. The lack of systematic analysis has not stopped commentators from claiming that there is a link between a state’s development status and the outcome of investment treaty arbitration. Some critics make normative statements, unsupported by data, that capital-

58. See *infra* notes 110–12 and accompanying text (providing definitions of these measures and explaining their relationship to development). R

59. Franck, *Evaluating Claims*, *supra* note 16, at 32. But see Ibronke T. Odumosu, *The Antinomies of the (Continued) Relevance of ICSID to the Third World*, 8 SAN DIEGO INT’L L.J. 345, 348, 364 (2007) (positing, without providing methodological support, that “only few arbitral proceedings have been initiated against developed states” and “ICSID tribunals have engaged mainly in settling disputes between foreign investors as claimants and Third World states as defendants,” and that “[a]n overwhelming majority of ICSID cases fall within this category”). R

60. Franck, *Evaluating Claims*, *supra* note 16, at 32. R

61. Peterson, *UK BIT Program and Sustainable Development*, *supra* note 52, at 6. R

62. See, e.g., Nathalie Bernasconi-Osterwalder, *Who Wins and Who Loses in Investment Arbitration? Are Investors and Host States on a Level Playing Field?*, 6 J. WORLD. INV. & TRADE 69, 69 (2005) (asserting that “host States cannot be winners in investment arbitration”); ANDERSON & GRUSKY, *supra* note 8, at ix (“Investors’ Odds of Winning are High.”). R

63. Franck, *Evaluating Claims*, *supra* note 16, at 49. R

64. *Id.* at 57–62.

importing countries “do not face unfair treatment or disadvantages out of BITs and ICSID arbitrations.”⁶⁵

Others disagree. The Bolivian President, Evo Morales, suggested that ICSID is an international organization where corporations bring their grievances and no country, except perhaps the United States, will win.⁶⁶ Meanwhile, Ecuador wishes to avoid investment arbitration at ICSID because of doubts about ICSID’s impartiality.⁶⁷ Others point to single cases with extreme damage awards as the norm without contextualizing these examples within the broader framework.⁶⁸ The British newspaper *The Guardian* reported on the *Biwater* case⁶⁹ to suggest, out of context, that the interests of the developing world were being “tramp[ed].”⁷⁰

Empirical research can and should assess these assumptions to determine whether there is a reliable link between the development status of respondent states and arbitration outcomes. This will enable commentators to make grounded assertions rather than relying on politically charged rhetoric. Likewise, with the aid of research, stakeholders will be able to make more informed policy choices when negotiating treaties or using the dispute resolution system.

B. *The Development Status of Arbitrators*

Arbitrator selection is a critical aspect of investment treaty arbitration. As Professor William Park suggests, “just as in real estate the three key elements are ‘location, location, location,’ so in arbitration the applicable trinity is ‘arbitrator, arbitrator, arbitrator.’”⁷¹

The objective in international arbitration is to have decision makers adjudicate disputes neutrally. Procedural justice is a fundamental aspect of the

65. Lars Glowinski, International Arbitration: Protection of Foreign Direct Investment and Foreign Investment Dispute Settlement Under ICSID and Bilateral Investment Treaties 75 (Feb. 15, 2007) (unpublished L.L.M. dissertation, University of Cape Town), available at <http://lawspace.law.uct.ac.za:8080/dspace/bitstream/2165/314/1/GLWLR001.pdf>.

66. BAI [Bolivian Agency of Information], *Palabras del Presidente de la República, Evo Morales Ayma, En La VI Cumbre de la Alternativa Bolivariana para los Pueblos de América y en Caribe (alba)* (26 de enero de 2008); see also *Evo Morales plantea buscar fórmula para neutralizar el CIADI*, EL ECONOMISTA, Apr. 29, 2007, <http://www.economista.es/empresas-finanzas/noticias/204986/04/07/Evo-Morales-plantea-buscar-formula-para-neutralizar-el-CIADI.html>; TUNUPA, *supra* note 5.

67. EL COMERCIO, *supra* note 7.

68. One example refers to “[i]nequitable and [e]xcessive” awards and then proceeds to cite cases that appear to be statistical outliers. Olivia Chung, Note, *The Lopsided International Investment Law Regime and Its Effect on the Future of Investor-State Arbitration*, 47 VA. J. INT’L L. 953, 965 (2007) [hereinafter Chung] (referring to the *CME (Netherlands) v. Czech Republic* award). It then refers to only a series of cases against Low Income countries without considering either the outcome of those cases or the outcome of other cases against more developed governments. See *id.* at 965–66.

69. *Biwater Gauff (Tanzania) Ltd. v. United Republic of Tanzania*, ICSID Case No. ARB/05/02 (Sept. 29, 2006), reprinted in 46 I.L.M. 15, 19 (2007).

70. Xan Rice, *The Water Margin*, THE GUARDIAN, Aug. 16, 2007, <http://www.guardian.co.uk/business/2007/aug/16/imf.internationalaidanddevelopment>.

71. William W. Park, *Income Tax Treaty Arbitration*, 10 GEO. MASON L. REV. 803, 813 (2002).

system.⁷² While a substantively correct result is desirable, it is also vital that parties and the public perceive the process to be procedurally fair, in order to maintain a legitimate dispute resolution system. The arbitrators are responsible for dispensing justice, so their backgrounds and methods of exercising authority are fundamental to systematic integrity. Arbitrator integrity has, however, been a historically thorny issue and is still an area of concern.

Given this concern, there has been a focus on variables that might affect, or be perceived to affect, arbitrators' neutrality and public confidence in arbitration.⁷³ For example, there is a historical focus on using arbitrator nationality to gauge neutrality.⁷⁴ Unfortunately, there is little theoretical analysis, let alone empirical assessment, of a link between nationality⁷⁵ and outcome.⁷⁶ The nationality of presiding arbitrators has been an area of particular focus, perhaps because of the decision makers' unique control over and role in the arbitration process. Redfern and Hunter suggest that "[i]n an ideal world, the nationality of a sole arbitrator, or of the presiding arbitrator, should be irrelevant."⁷⁷ Practitioners and parties, however, consider nationality of the presiding arbitrator to be crucial given concerns about the integrity of the dispute resolution process.⁷⁸

72. JOHN THIBAUT & LAURENS WALKER, *PROCEDURAL JUSTICE: A PSYCHOLOGICAL ANALYSIS* 94 (1975) (arguing that the dispute resolution process strongly influences the disputants' level of satisfaction with the ultimate resolution); TOM R. TYLER & STEVEN L. BLADER, *COOPERATION IN GROUPS: PROCEDURAL JUSTICE, SOCIAL IDENTITY AND BEHAVIORAL ENGAGEMENT* 77–80 (2000); TOM R. TYLER, *WHY PEOPLE OBEY THE LAW* 107–08 (1990) (reporting that in all types of dispute resolution, people are far more likely to obey the law if they are confident that decision-making procedures are fair); Tom R. Tyler, *Public Mistrust of the Law: A Political Perspective*, 66 U. CIN. L. REV. 847, 856–60, 867–70 (1998) (suggesting that people are more willing to comply with the law when they perceive it to be legitimate and deserving of compliance, and that the primary aspect of legitimacy is perceived procedural fairness and trust in legal authorities).

73. See Pierre LaLive, *On the Neutrality of the Arbitrator and of the Place of Arbitration*, in *SWISS ESSAYS ON INTERNATIONAL ARBITRATION* 23, 25–26 (1984) [hereinafter LaLive] (commenting on the importance of the perceived neutrality of the presiding arbitrator); see also *id.* at 25 (“[A]s a symbol of impartiality, the national neutrality of an arbitrator is a vital factor for the proper functioning of a good arbitral tribunal.”) (quoting Neil Pearson, *Nationalité et attaches de l'arbitre*, *SYMPOSIUM REV. ARBITRAGE* 239 (1970) (translation by Pierre LaLive)).

74. See Landau, *supra* note 39, at 74 (“Nationality, in the end, is simply one of several objective criteria to test impartiality.”); see also Ilhyung Lee, *Practice and Predicament: The Nationality of the International Arbitrator (With Survey Results)*, 31 *FORDHAM INT'L L.J.* 603, 613–14 (2008) [hereinafter Lee] (“The doubt and fear is of arbitrator bias toward the party with the same nationality, notwithstanding his obligation to be independent and impartial.”).

75. Nationality can be synonymous with citizenship. See Lee, *supra* note 74, at 608.

76. See *id.* at 604 (referring to commentary by LaLive and Redfern & Hunter).

77. ALAN REDFERN & MARTIN HUNTER, *LAW AND PRACTICE OF INTERNATIONAL COMMERCIAL ARBITRATION* 202 (4th ed. 2004).

78. See Lee, *supra* note 74, at 604 (“Consideration of the nationality of the [presiding] arbitrator is quite important then, to parties, counsel, and administering organizations; and when arbitrator nationality is disregarded, it could be of interest to reviewing courts that are requested to decline the award.”); *id.* at 612; LaLive, *supra* note 73, at 25 (“The question is whether he will thus be thought of or seen as more impartial by the parties, i.e. whether he will (rightly or wrongly) inspire more confidence because of his ‘neutral’ nationality.”); Salomon, *supra* note 38, at 2 (“Because people tend to share similar value systems when they identify with the same home country, many parties and institutions require the presiding arbitrator to share nationality with neither of the parties.”).

R

R

R

R

R

With the focus on development during the Doha Round of trade negotiations and UNCTAD's mandate to focus on the development dimensions of foreign investment, factors related to development are receiving enhanced scrutiny.⁷⁹ For example, in the international trade context, concerns about procedural integrity in adjudication were sufficiently critical that the WTO's Dispute Settlement Understanding, which creates legalized adjudication for trade disputes akin to investment arbitration, enshrines a right for a developing country to require the presence of one panelist from another developing country.⁸⁰ Therefore, decision makers from the developed world do not have exclusive control over cases involving the developing world.

In investment arbitration, there is a lurking concern that the development status of arbitrators, particularly presiding arbitrators who wield especially strong influence, may be inappropriately associated with certain outcomes.⁸¹ One author even explains that there is "some concern in developing countries over the selection of arbitrators" at entities such as ICSID, and such appointments may create a "systemic . . . bias in favor of Western legal concepts and the positions."⁸² The complaints have not been theoretically grounded but appear to suggest that arbitrators from the developed world treat the developing world unfairly and perhaps favor transnational entities, which are presumably Western in orientation.

Various theoretical narratives might help explain how the development status of arbitrators could affect or be perceived to impact outcome.⁸³ One narrative involves arbitrator over-identification with a party. The second involves alignment with the developed world to gain access to the Western

79. See DEZALAY & GARTH, *supra* note 44, at 95–98 (discussing the North/South divide); LaLive, *supra* note 73, at 26 (discussing the East/West divide); see, e.g., United Nations Conference on Trade and Development, *The Development Dimension of Foreign Direct Investment: Policies to Enhance the Role of FDI, in the National and International Context—Policy Issues to Consider*, ¶¶ 2–3, U.N. Doc. TD/B/COM.2/EM.12/2 (Sept. 23, 2002). Some concerns may come from the Cold War's East/West divide or the North/South divide between the economic haves and have-nots.

80. General Agreement on Tariffs and Trade—Multilateral Trade Negotiations (The Uruguay Round): Understanding on Rules and Procedures Governing the Settlement of Disputes art. 8.10, Dec. 15, 1993, 33 I.L.M. 112 (1994).

81. Ercus Stewart, Arbitration in the Developing World, Speech at Conferences 21: Cotrina 2008 Legal Conference 5–6 (Jan. 7, 2008), available at <http://www.conferences21.com/UserFiles/File/Ercus%20Stewart%20OSC%20-%20Arbitration%20in%20the%20Developing%20World.pdf>; see also *infra* notes 84–85. Variables connected with inadvertent or intentional bias may go beyond nationality or development status. They might involve domicile, residence, religion, regional origin, ethnicity, culture, or linguistic background. See Lee, *supra* note 74, at 616.

82. AMAZU A. ASOUZU, INTERNATIONAL COMMERCIAL ARBITRATION AND AFRICAN STATES: PRACTICE, PARTICIPATION AND INSTITUTIONAL DEVELOPMENT 404–05 (2001) (quoting THE ENERGY CHARTER TREATY: AN EAST-WEST GATEWAY FOR INVESTMENT AND TRADE 56–57 (THOMAS W. WÄLDE ED. 1996)). But see R. Rajesh Babu, *International Commercial Arbitration and the Developing Countries*, 4 ASIAN-AFRICAN LEGAL CONSULTATIVE ORG. Q. BULL. 385, 386–87, 399 (2006) (acknowledging that developing countries are using and complaining less about international arbitration, but arguing that the developing world should "dismantle the existing [arbitration] structure which is based on doctrines associated with neo-colonistic efforts at the preservation of economic dominance").

83. See LaLive, *supra* note 73, at 26 (presuming that nationality may be important because "parties will generally assume without much further thought that a prospective arbitrator is likely, or even bound, to share his country's ideology and common values, if any").

R
RR
R

R

arbitration “club.” A third narrative, however, suggests that development status does not affect outcome, and that other variables may provide better indicators of potential arbitrator bias.

The concerns in the first narrative may stem from an assumption that a presiding arbitrator’s development status might lead to over-identification with parties with similar development backgrounds. This tendency might mean, for example, that arbitrators from the developed world might favor investors or governments from the developed world; likewise, presiding arbitrators from the developing world might favor developing countries.

In a 2005 speech, Roberto Dañino, then Secretary-General of ICSID, touched on this issue. He explained that there is a concern “expressed by a few . . . that ICSID arbitrators are predominantly nationals from developed countries, the implication being that they may be more favorably inclined towards investors” from the developed world and less favorably inclined towards governments from the developing world.⁸⁴ Presumably the concern is that where arbitrators share a party’s developmental background, and possibly legal culture, political perspectives, linguistic understandings, or loyalties, the arbitrator may be predisposed towards the party or inadvertently favor the party during the proceedings.⁸⁵ Dañino asserted that both investors and governments won arbitrations, and that such a result should address concerns about arbitrators’ development bias.⁸⁶ Dañino nevertheless did not address the underlying issue, namely, whether the success of investors or governments is somehow linked to the development background of the arbitrators.

The second narrative asserts that arbitrators, particularly those from the developing world, may seek to identify themselves with the developed world. Some have referred to this as the “question of Western bias.”⁸⁷ This bias, be it real or perceived, might manifest itself in different ways.

For instance, arbitrators might have a pro-investor protection approach⁸⁸ that favors the developed world.⁸⁹ Professor Sornarajah, for example, has criticized the appointment of arbitrators who have a pro-investor bias and

84. Roberto Dañino, Sec’y Gen, ICSID, Opening Remarks at the Symposium: Making the Most of Investment Agreements: A Common Agenda (Dec. 12, 2005) [hereinafter Dañino].

85. See Landau, *supra* note 39, at 73 (“A greater degree of confidence may be inspired on all sides if there is no chance that one party will get a better hearing because of some cultural or national identification between that party and the arbitrator.”).

86. See Dañino, *supra* note 84.

87. Stewart, *supra* note 81, at 2, 8; see also Dezalay & Garth, *supra* note 44, at 8 (referring to the “rather closed and arcane European club” of arbitrators); Arash S. Arabi, *Renegotiating Third World Debt*, 3 PEPP. DISP. RESOL. L.J. 251, 253–54 (2003) (suggesting that arbitration exhibits a pro-Western bias).

88. G. Kahale, III, *A Problem in Investor/State Arbitration*, TRANSNT’L DISP. MGMT (June 2008), http://www.cm-p.com/pdf/A_Problem_in_Investor_State_Arbitration-tdm.pdf [hereinafter Kahale] (discussing the problems of arbitrators with “expansive, pro-investor legal theories”).

89. This is presumably because commentators assume that favoring investors necessarily means favoring the developed world. There may be some credibility to the assumption. Data from one study found that approximately 90% of investors were from the developed world. Franck, *Evaluating Claims*, *supra* note 16, at 29.

R

R

R

R

would therefore presumably favor the developed world.⁹⁰ The purportedly “pro-Western” bias of arbitrators may also manifest itself through favoritism toward governments from the developed world. In both instances, the underlying concern is that arbitrators may align themselves with the developed, Western world to gain access to the prestige, stature, and financial opportunities⁹¹ of the arbitration “club.”⁹² Dezalay and Garth’s classic analysis of international commercial arbitration observes that “an arbitrator from the third world must find a way to gain access to and credibility with” the Western arbitration community.⁹³ This concern may be at the core of the perceived problems with “pro-Western bias [that] may be well ingrained within the culture of developing countries.”⁹⁴

Others reject the assertion that an arbitrator’s developmental background could create bias or otherwise affect outcomes. Without reference to data, some suggest that “ICSID arbitrators are not irredeemably tainted by institutional bias.”⁹⁵ Jan Paulsson suggests while there may have been historical concerns about arbitrator bias in investment-related arbitration, “the dice are loaded no longer.”⁹⁶ Instead, developing countries should recognize “international arbitration as it is: a neutral means for the resolution of conflicts . . . to be mastered rather than complained about.”⁹⁷

Professor Shalakany likewise rejects arguments of a pro-Western arbitrator bias⁹⁸ but suggests the phenomenon is more nuanced since different arbi-

90. Without recognizing empirical evidence that governments win more often than investors, Sornarajah posits that: “Though neutrality is the ideal subscribed to in international arbitration, the pattern of appointing arbitrators favourable to the articulation of norms that protect the interests of international business has existed for a long time. It is alleged that this pattern is inherent within the system of international arbitration itself, so that only persons known to be favourable to definite outcomes are chosen as arbitrators.” Sornarajah, *Third World Resistance*, *supra* note 48, at 33. This position is perhaps an outgrowth of his earlier critique that international commercial arbitration was historically biased against the developing world and “weighted in favour of the capital exporting States.” M. Sornarajah, *The Climate of International Commercial Arbitration*, 8 J. INT’L ARB. 47, 47 (1991). Sornarajah also asserts that “[a]rbitrators must subscribe to the tenets of the powerful if they are to remain in business.” Sornarajah, *Third World Resistance*, *supra* note 48, at 34.

91. DEZALAY & GARTH, *supra* note 44, at 50–51 (describing the prestige and financial benefits to “club” membership).

92. *Id.* at 91 (describing the “western domination in the market of north-south arbitration”); *see also id.* at 68–69, 96, 98.

93. *Id.* at 25–26, 89–91. While “lawyers from the third world can gain recognition as arbitrators, they must gain recognition from the individuals and institutions” in the arbitration club. *Id.* at 29. One of Garth and Dezalay’s interviewees expressed concern about “[characterizing] developing countries as ‘victims’ of an ‘international mafia’ of European arbitrators.” *Id.*

94. Stewart, *supra* note 81, at 8.

95. Sandra L. Caruba, *Resolving International Investment Disputes in a Globalized World*, 13 N.Z. BUS. L.Q. 128, 148 (2007).

96. Jan Paulsson, *Third World Participation in International Investment Arbitration*, 2 ICSID REV. FOREIGN INVESTMENT L.J. 19, 21 (1987) [hereinafter Paulsson, *Third World Participation*].

97. *Id.* at 20.

98. Amr A. Shalakany, *Arbitration and the Third World: A Plea for Reassessing Bias Under the Specter of Neoliberalism*, 41 HARV. INT’L L.J. 419, 430 (2000) (arguing that international arbitration is “not per se biased on an institutional level—that is, it is not inescapably predisposed to particular political interests or agendas” from either the developed or developing world).

R

R

R

R

trations may inhabit different political, economic, and legal spaces. Shalakany argues that arbitrators' approaches may vary,⁹⁹ positing that outcomes are instead associated with variables such as arbitrator attitudes towards appropriate levels of sovereign control.¹⁰⁰ One might develop this argument further to suggest that an arbitrator's approach may vary according to factors such as the arbitrator's background and educational training, the nature of the parties, the framework from which the legal rights originate, and the background of the particular disputes.

This research is the first step in isolating variables linked with arbitrators' decisions and modeling the complexities of Shalakany's theory of arbitrator bias. While some research explores the nationality of arbitrators,¹⁰¹ no quantitative empirical research has addressed the intersection of an arbitrator's development background and arbitration outcome. Given concerns that "[i]t is also crucial for arbitrators to not have a bias favoring developed countries,"¹⁰² research in this area could prove useful.¹⁰³ If nothing else, it could assess the validity of the perceptions that arbitrators over-identify with parties sharing similar developmental perspectives or align themselves with the developed world and exhibit a pro-Western bias.

It is no small matter that parties may perceive presiding arbitrators, with their unique role, to be inappropriately aligned with the developed world. That perception can foster concern about unfairness in the system. One commentator suggests the following:

[I]t is time to recognize that there is a perception of unfairness which can no longer simply be ignored. If the perception persists, it is to be expected that more states will withdraw from investment treaties and from ICSID, and more will simply refuse to agree to any international arbitration.¹⁰⁴

This is where empirical research has particular power to shed light on the integrity of arbitration and aid the assessment of the need for procedural reform.

99. *Id.* at 429.

100. *Id.* at 467–68; *see also supra* notes 81–85 and accompanying text (suggesting factors like shared religious or linguistic traditions may affect outcome).

101. Franck, *Evaluating Claims*, *supra* note 16, at 77–81.

102. *See* Chung, *supra* note 68, at 976.

103. Franck, *Evaluating Claims*, *supra* note 16, at 80–81.

104. Kahale, *supra* note 88, at 5.

IV. RESEARCH BACKGROUND: HYPOTHESES AND METHODOLOGY

This research used existing archival data¹⁰⁵ to explore whether there was an association between outcome and the development status of the respondent state, the development status of the presiding arbitrator, or an interaction between those variables. The objective of this quantitative research was to look for reliable statistical relationships to assess whether arbitration inappropriately favored the developed or developing world and whether the development status of the presiding arbitrator was associated with outcome.¹⁰⁶

A. *Research Hypotheses*

This research explored three questions. First, what kind of interaction effect¹⁰⁷ might exist between the development status of the government respondent and the development status of the presiding arbitrator that could influence the outcome? Second, how does the respondent's development status affect outcome, if at all? Third, how does the presiding arbitrator's development status affect outcome, if at all? In other words, is there a main effect¹⁰⁸ for either respondent state's development status or presiding arbitrator's development status?

The research hypothesis was that development status of presiding arbitrators and respondent states would not affect outcome, either as a main effect of each independent variable or through an interaction. The theoretical assumption was that international adjudicative processes need not depend upon spurious variables such as development status. Rather, arbitrators can make decisions neutrally on the basis of the facts and law. Using statistical methods, it is possible to evaluate whether reality mirrors that hope.

B. *Research Methodology*

This section defines and operationalizes the independent and dependent variables. It further explores and describes the type of statistical analyses used.

105. The data used to conduct the analyses came from the population of 102 investment treaty awards from 82 cases that were publicly available before June 1, 2006. Franck, *Evaluating Claims*, *supra* note 16, at 24, 52.

106. *Id.* at 52.

107. An interaction effect is "a pattern of data obtained from multifactorial analysis of variance (ANOVA) . . . in which the effect of an independent variable or factor [] varies across levels of another independent variable, or across combinations of levels of other independent variables or factors. When this occurs, variation in the dependent variable is not the result of a simple additive combination of the independent variables or factors." ANDREW M. COLMAN, A DICTIONARY OF PSYCHOLOGY 382 (2d ed. 2006) [hereinafter COLMAN]; URDAN, *supra* note 15, at 119–20.

108. A main effect is "data obtained from an analysis of variance (ANOVA) . . . the mean change in scores resulting from a change in the level of a single independent variable or factor, averaged over all combinations of levels of the other factor(s)." See COLMAN, *supra* note 107, at 437; see also URDAN, *supra* note 15, at 118–19.

R

R

R

R

1. *The Independent Variables: Development Status of Respondent States and Presiding Arbitrators*

There were two independent variables related to development. The first was the development status of the respondent state. The second was the development status of the presiding arbitrator of the tribunal.

“Development status” was operationalized in two ways. It used a more blunt, two-category variable (“OECD status”), as well as a more nuanced four-category variable (“World Bank status”).¹⁰⁹ OECD status was a binary, categorical variable defined by whether a country was a member or non-member of the OECD.¹¹⁰ Likewise, the development status of presiding arbitrators was derived from their nationalities, which were generally obtained from award descriptions.¹¹¹

World Bank status provided another gauge for defining development status. This more sensitive measurement scale used the World Bank’s categorical classifications: (1) High Income countries, irrespective of whether they are OECD Member States, (2) Upper-Middle Income countries, (3) Lower-Middle Income countries, and (4) Low Income countries.¹¹² For respondents, development status was equivalent to their World Bank classifications. The development status of each presiding arbitrator was coded using the World Bank status of his or her country of nationality.

109. Both OECD status and World Bank status were assessed in order to address different conceptions of what development can mean. It might mean gaining access to the OECD, whose members consist primarily of developed states. Alternatively, it may involve economic development measured by standard of living, which is related to gross national income (“GNI”). These two definitions do not necessarily overlap. For example, Mexico is a member of the OECD but is not classified as a High Income country. See *infra* notes 110, 112 and accompanying text.

110. The OECD Member States are Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, South Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom, and the United States. Ratification of the Convention on the OECD, http://www.oecd.org/document/58/0,2340,en_2649_201185_1889402_1_1_1_1,00.html (last visited Mar. 14, 2009).

111. The data sometimes came from institutional websites, such as ICSID and IAI, which publicly identify arbitrator nationality. For arbitrators whose nationality was not ascertainable in a public award or on an institutional website, nationality was coded according to publicly available information through internet searches. Where it was impossible to ascertain reliable information, the presiding arbitrator’s nationality was coded as unavailable. This accounts for the decrease in cell counts. For example, there are fifty-two awards finally resolving investment treaty claims. Since the nationalities of the presiding arbitrator in three cases are unavailable, the ANOVAs only analyzed forty-nine final awards. Similarly, the chi-squares only analyzed forty-seven final awards. The cell count decreased because the chi-square analyses omitted two awards that involved settlement agreements. The three cases for which arbitrator information was unavailable (*Nykomb Synergistics Technology Holding AB v. Latvia*, *CCL v. Kazakhstan*, and *Nagel v. Czech Republic*) all operated according to the Stockholm Chamber of Commerce Rules. This may implicate issues of external validity and case selection bias.

112. See The World Bank, Data and Statistics, Country Classification, <http://www.worldbank.org/data/countryclass/countryclass.html> (last visited Mar. 15, 2009). The World Bank’s main criterion for classifying economies is GNI per capita. *Id.*

2. *The Dependent Variables and Selected Statistical Models*

The dependent variable—outcome—was operationalized in two ways. Using data from publicly available final investment treaty arbitration awards, the research analyzed the two dependent variables separately. For a first set of analyses, outcome was a binary variable defined as whether the claimant or respondent ultimately won the arbitration. The respondent was coded as the ultimate winner if the tribunal awarded US\$0 for a treaty breach. The claimant was the ultimate winner if the tribunal awarded more than US\$0 for a breach. This dependent variable was then subjected to a Pearson's Chi-square Test of Independence,¹¹³ which assessed the two independent categorical variables¹¹⁴ to determine if different development categories had statistically different response patterns.

For a second set of analyses, the research defined outcome quantitatively.¹¹⁵ After converting awards to a common currency, outcome was measured using the amounts awarded by investment treaty tribunals. Analyses used both (1) the actual raw data and (2) winsorized¹¹⁶ data that eliminated statistical outliers.¹¹⁷ This minimized positive skewing and provided data that adhered more closely to the assumptions of the statistical tests, allowing the analyses to operate more effectively.¹¹⁸ This dependent variable, namely the amount awarded, was then subjected to an Analysis of Variance

113. The Pearson's Chi-square Test of Independence evaluates whether there is a pattern of relationship between two categorical variables or whether the variables appear to be independent and unrelated. URDAN, *supra* note 15, at 161–63.

114. This research defined a categorical variable as a qualitative variable that breaks variance into different, discrete categories, as it represents qualitative differences rather than nominally scaled quantitative differences. *Id.* at 3–4.

115. A quantitative variable “is one that is scored in such a way that the numbers, or values, indicate some sort of amount.” *Id.* When the dependent variable “outcome” is a numerical value, it has a continuous, quantitative value.

116. Winsorizing is the process of identifying and converting extreme values in data into the upper or lower bounds of the distribution of the normal curve. See W.J. Dixon, *Simplified Estimation from Censored Normal Samples*, 31 ANNALS MATH. STAT. 385, 385 (1960); John W. Tukey, *The Future of Data Analysis*, 33 ANNALS MATH. STAT. 1, 18–19 (1962).

117. Outliers are data points too extreme to be part of the population of interest. In this case, they consist of extremely low and extremely high arbitration awards. The research identified outliers using cutoffs determined by Tukey's hinges, which is a method that mathematically computes the low and high cutoff scores for the variables of interest. After identifying the outliers, there were two possible methods for cleaning the data. The researcher can trim the scores, which entails removing all outliers and leaving missing data where these extreme values were once located. The other method is called winsorizing, which entails replacing the outliers' actual value with the upper and lower bounds provided by Tukey's hinges. This method reformulates the data to fit the assumptions of the statistical tests while retaining data points. See, e.g., DAVID SHESKIN, HANDBOOK OF PARAMETRIC AND NONPARAMETRIC STATISTICAL PROCEDURES 403–04 (3d ed. 2004). This model used winsorizing to create a better statistical approximation of the population of interest and to avoid losing information about award values.

118. This study used raw data as lawyers may be interested in raw dollar values related to the population of interest. However, it also used winsorized data, because winsorizing eliminates the strong positive skewing of the raw data, minimizes data loss by not eliminating statistical outliers, and permits the data to adhere to the underlying parameters of statistical tests to increase the likelihood that the tests are functioning effectively. See, e.g., THOMAS JANOSKI & ALEXANDER M. HICKS, THE COMPARATIVE POLITICAL ECONOMY OF THE WELFARE STATE 330–32 (1994); John W. Tukey, *supra* note 116.

R

R

(“ANOVA”)¹¹⁹ for independent groups, since it involved the assessment of categorical variables (that is, different development categories) to compare mean amounts awarded and assess whether there were statistically different response patterns.

The analyses’ objective was to determine whether there was (1) a main effect for respondent state development status (that is, a reliable statistical relationship between amounts awarded and the respondent’s development status), (2) a main effect for presiding arbitrator development status (that is, a reliable statistical relationship between amounts awarded and the presiding arbitrators’ development status), or (3) an interaction between the development status of respondents and presiding arbitrators that contributed to a statistically significant difference in amounts awarded.

V. RESEARCH RESULTS AND DISCUSSION

This Part describes the results of the statistical analyses. Section A provides quantitative, descriptive statistics about the development status of presiding arbitrators. It also examines presiding arbitrator appointments in relation to the development status of each case’s respondent state. Sections B and C provide the results of the associative hypothesis testing.¹²⁰ Section B considers the relationship among the two independent variables and the ultimate winner (that is, investors or states) of investment treaty arbitration. Section C considers the relationship among the two independent variables and the amounts awarded.

The subsections within sections B and C discuss the results of each statistical analysis in three ways. First, they describe the results of null hypothesis¹²¹ significance tests, which look for reliable statistical relationships between development status and the outcome of investment treaty cases.¹²² Second, they describe the effect sizes. A researcher can use effect sizes to estimate the risk of missing a reliable statistical relationship and to approximate the sample size necessary to assess reliably whether a statistical relationship exists. Effect sizes also measure the magnitude of a potential effect.¹²³ Cohen’s suggested conventions for understanding effect sizes indi-

119. An ANOVA analysis compares the means of two of more groups on a dependent variable to determine if the group means are significantly different from each other. URDAN, *supra* note 15 at 101–02, 117–18. R

120. Associative hypothesis testing analyzes the ways in which variables may relate to each other, but not necessarily in a causal manner. LOUIS COHEN ET AL., *RESEARCH METHODS IN EDUCATION* 519 (6th ed. 2007).

121. The “null hypothesis” is generally that there is no relationship or difference between the variables. URDAN, *supra* note 15, at 26; COHEN ET AL., *supra* note 120, at 520. R

122. URDAN, *supra* note 15, at 62. R

123. See FREDERICK J. GRAVETTER & LARRY B. WALLNAU, *ESSENTIALS OF STATISTICS FOR THE BEHAVIORAL SCIENCES* 198–201, 345–48 (6th ed. 2007) [hereinafter GRAVETTER & WALLNAU]; Robert Coe, *It’s the Effect Size, Stupid: What Effect Size Is and Why It Is Important* (Sept. 12–14, 2002) (paper presented at the British Educational Research Association annual conference), <http://www.cemcentre.org/>

cate that a “small” effect is present when $r = .10$, a “medium” effect is present when $r = .30$, and a “large” effect is present when $r = .50$.¹²⁴ Effect sizes below $r = .10$ are less than “small” and arguably of trivial impact.¹²⁵ By measuring the potential strength of a relationship between two variables, effect sizes aid assessment of whether, on a normative level, the size of a reliable statistical difference is a matter of practical concern or is so tiny as to be irrelevant.¹²⁶ Third, these subsections discuss and contextualize the individual analyses.

Section D then provides a general discussion that synthesizes and interprets the particularized results from sections B and C. Finally, section E describes the limitations of the analyses and related inferences.

A. Basic Arbitrator Demographics

As discussed earlier, parties have some, but not necessarily complete, control over the selection of the presiding arbitrator.¹²⁷ Irrespective of the level of control, there are lingering concerns that the pool of arbitrators is limited and there is a democracy deficit between those making the decisions and those stakeholders affected by the outcomes.

Previous research identified a pool of 145 arbitrators from 40 different countries. Even though approximately 70% of the cases in the study were from non-OECD countries, 109 arbitrators (75% of the population) were from OECD countries.¹²⁸ It is therefore useful to explore arbitrator nationality on a deeper level, namely by exploring the development status of presiding arbitrators and respondents.

As a general matter, there were presiding arbitrators from both OECD and non-OECD countries. Although there were more cases involving non-OECD states, arbitrations against OECD and non-OECD respondent states both had a greater number of presiding arbitrators from OECD countries than from non-OECD countries (Table 1).

Documents/CEM%20Extra/EBE/ESguide.pdf. See generally JACOB COHEN, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES (2d ed. 1988) [hereinafter COHEN].

124. COHEN, *supra* note 123, at 115.

125. *Id.* at 113–16; see also ROBERT J. GRISSOM & JOHN J. KIM, EFFECT SIZES FOR RESEARCH 85–87 (2005) (discussing effect sizes but acknowledging that the individual circumstances of each research area should be taken into account).

126. See COHEN, *supra* note 123, at 25 (“The terms ‘small,’ ‘medium,’ and ‘large’ are relative, not only to each other, but to the area of behavioral science or even more particularly to the specific content and research method being employed in any given investigation. . . there is a certain risk inherent in offering conventional operational definitions for these terms for use in power analysis. . . . This risk is nevertheless accepted in the belief that more is to be gained than lost by supplying a common conventional frame of reference.”).

127. See *supra* notes 35–37.

128. Franck, *Evaluating Claims*, *supra* note 16, at 77–79.

R

R

R

R

Table 1: Breakdown of Presiding Arbitrator's OECD Status and Respondent State's OECD Status

<i>Presiding Arbitrator</i>	Respondent State		Total
	Non-OECD	OECD	
<i>Non-OECD</i>	10	3	13
<i>OECD</i>	21	15	36
Total	31	18	49

A similar predominance of presiding arbitrators from higher income countries emerged with World Bank status. There were notable gaps. Namely, there were no presiding arbitrators from Low Income countries. Likewise, Low Income countries did not have tribunals with presiding arbitrators from Upper-Middle Income countries; and Upper-Middle Income countries did not have any presiding arbitrators from Lower-Middle Income countries (Table 2).

Table 2: Breakdown of Presiding Arbitrator's World Bank Status and Respondent State's World Bank Status

<i>Presiding Arbitrator</i>	Respondent State				Total
	High	Upper-Middle	Lower-Middle	Low	
<i>High</i>	9	13	12	2	36
<i>Upper-Middle</i>	1	4	3	0	8
<i>Lower-Middle</i>	1	0	2	2	5
<i>Low</i>	0	0	0	0	0
Total	11	17	17	4	49

Since there were so few developing-world presiding arbitrators to analyze, it was difficult to determine whether their presence made a difference. The small number of arbitrators from the developing world¹²⁹ created gaps in the design. To promote more statistically reliable results, the current research collapsed the World Bank status of presiding arbitrators into a binary variable. Presiding arbitrators were either from (1) a High Income country or (2) an Upper-Middle or Lower-Middle Income country.

B. Development Status and the Ultimate Winner of Investment Treaty Arbitration

The first analyses evaluated the relationship among the development status of the respondent state, the development status of the presiding arbitra-

129. The underlying data has limitations given that it comes from publicly available archives and dates only to June 1, 2006. There are now approximately three more years of data to collect and analyze. It is necessary to acknowledge this limitation, and future research should replicate the analyses.

tor, and the party (investor or state) that won the arbitration. This section analyzes the two chi-square tests. Both tests assessed whether there was a reliable pattern of relationship between the winner of investment treaty cases and the development status of respondent states and presiding arbitrators. The first test used the independent variable of OECD status, while the second test used World Bank status.

1. *OECD Status and Winning Investment Treaty Arbitration*

The research model utilized a $2 \times 2 \times 2$ cross-tabulation to see if there was a statistically significant pattern of relationship between the OECD status of respondent states, the OECD status of the presiding arbitrator, and winning or losing in investment treaty disputes.

a. Results of Null Hypothesis Significance Tests

As hypothesized, the results indicate three things. First, there was no statistically significant pattern among the OECD status of the respondent state, the OECD status of the presiding arbitrator, and winning a given investment treaty dispute. For the sample of forty-seven awards shown in Table 3, the number of winners and losers were statistically equivalent.

With regard to the simple two-way effect of presiding arbitrators from non-OECD countries, there was no pattern relationship between OECD and non-OECD respondent states' wins or losses in investment treaty arbitration ($\chi^2(1) = .258; p = .61; n = 13$). With regard to the simple two-way effect of presiding arbitrators from OECD countries, there was likewise no relationship between OECD and non-OECD respondent state status and success or defeat in investment treaty arbitration ($\chi^2(1) = .045; p = .83; n = 34$). Because the simple two-way tests were not significant, it is unlikely that there is a three-way interaction or that the two two-ways are different from each other.

Statistical significance generally requires an alpha (α) level of $p = .05$ or lower. A p value of $.05$ or lower means that there is a 5% chance (or less) that a detected effect is due to chance. The p values for this analysis were $.61$ and $.83$, which are far from being statistically significant. In other words, there was no association between the OECD status of the presiding arbitrator and the outcome of the arbitration.

The ultimate win frequencies, broken down by the OECD status of the presiding arbitrator and the OECD status of the respondent state, appear in Table 3.

Table 3: OECD Status of Presiding Arbitrator, OECD Status of Respondent State, and Frequency Breakdown of the Ultimate Winner of an Investment Treaty Arbitration Case

Presiding Arbitrator	Respondent State	Ultimate Winner		Total
		Claimant	Respondent	
Non-OECD				
	Non-OECD	5	5	10
	OECD	2	1	3
OECD				
	Non-OECD	7	12	19
	OECD	5	10	15
Total		19	28	47

* None of the observed frequencies differed from what would be expected due to chance.

b. Effect Sizes

The simple two-way effect of the chi-square test for the OECD status of presiding arbitrators on the ultimate winner had an effect size of .04, which is a potentially trivial value.¹³⁰ By contrast, the effect size for the simple two-way effect of chi-square analysis for the non-OECD status of presiding arbitrators on the ultimate winner was .14, which is slightly larger than a potentially “small” effect.¹³¹ An effect size calculation is a measure of the potential strength of the relationship that *might* exist, while the chi-square test is a measure of whether there is a statistically significant relationship.

The relationship between the development status of respondents, presiding arbitrators, and outcome was not significant, and the potential strength of that non-significant relationship was very weak. The simple two-way effect of OECD presiding arbitrators on the success rates of OECD and non-OECD respondent states does not appear to suffer from a power problem that requires a larger sample.¹³² Because the simple two-way effect for non-OECD presiding arbitrators potentially exhibits a small effect, replicating the analysis with a larger sample¹³³ eliminates a potential power problem.¹³⁴

130. COHEN, *supra* note 123, at 113–16.

131. *Id.*

132. The power of an analysis is determined by using power tables, such as those found in COHEN, *supra* note 123, which the researcher can then use to estimate the probability of having committed a Type II error (Type II error rate = 1 – power). *Id.* at 3–6. As the power of the OECD status of presiding arbitrators is less than .20 ($r < .10$ and $n = 34$), there is theoretically a greater than 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 very small. That would likely be the situation in connection with the simple two-way effect of OECD presiding arbitrators. However, if one wished to be conservative and assess the lack of a relationship in the population, an a priori power analysis for an effect size of .04 ($r = .10$; $S = 781$) would require a sample of 1,562 final investment treaty cases ($N = n(781/2) * k(4)$) to have a sufficient number in each of the four conditions. Generally, a result with .80 power (i.e., a 20% risk of having committed a Type II error) is acceptable in the social science literature.

133. With an overall sample of forty-seven and an effect size of .14 ($n = 13$), the power was less than .20. This means that there is theoretically a greater than 80% risk of having incorrectly determined that

R

R

2. *World Bank Status and Winning Investment Treaty Arbitration*

The research used a $2 \times 4 \times 2$ cross-tabulation to see if there was a statistically significant pattern of relationship among (1) having a presiding arbitrator from either a High Income or Middle Income country, (2) being a High Income, Upper-Middle Income, Lower-Middle Income, or Low Income respondent state, and (3) winning or losing an investment treaty arbitration.¹³⁵

a. Results of Null Hypothesis Significance Tests

As hypothesized, the results suggest that there was no significant pattern of relationship between the World Bank status of the presiding arbitrator, the World Bank status of the respondent state, and the winner of an investment treaty arbitration. For the sample of forty-seven awards in Table 4, no statistically significant difference exists between winners and losers.

With regard to the simple two-way effect of presiding arbitrators from High Income countries, there was no relationship between the World Bank status of the respondent and the winner of an investment treaty arbitration (χ^2 (df > 1) = 2.216; $p = .53$). With regard to the simple two-way effect of presiding arbitrators from Upper-Middle or Lower-Middle Income countries, there was likewise no relationship between the World Bank status of the respondent states and the winner of an investment treaty arbitration (χ^2 (df > 1) = 1.130; $p = .77$).

Follow-up pairwise comparisons did not find any statistically significant relationships among the World Bank status of the respondent state, the World Bank status of the presiding arbitrator, and winning or losing. In other words, whether (1) the respondent state is a High, Upper-Middle, Lower-Middle, or Low Income country, or (2) the presiding arbitrator was from a High, Upper-Middle, or Lower-Middle Income country was unrelated to whether the respondent won or lost.

The p values of these analyses were .53 and .77, which are far from being statistically significant. In other words, the World Bank status of neither the respondent nor the presiding arbitrator was related to which party won the arbitration. The ultimate winners, broken down by the World Bank status of the presiding arbitrator and respondent, appear in Table 4.

there is no relationship between a non-OECD presiding arbitrator, the OECD status of the respondent state, and the ultimate winner of an investment arbitration.

134. An a priori power analysis suggests that, to capture an effect size of $r = .14$, a sample of 686 final arbitration awards would create the requisite power. In order to create a sample with .80 power (20% likelihood of a Type II error), and an effect size of .14 ($r = .15$; $S = 343$), a sample of 686 final arbitration awards would be necessary to ensure a number in each of the four conditions ($N = n(343/2) * k(4)$) that would be sufficient to find the smallest effect size in the simple 2×2 .

135. The sample size for this analysis was forty-seven. Two awards involving settlement agreements and three awards where the presiding arbitrator's nationality was unknown were omitted from the analysis.

Table 4: World Bank Status of Respondent State, World Bank Status of Presiding Arbitrator, and Frequency Breakdown of the Ultimate Winner of an Investment Treaty Arbitration Case

Presiding Arbitrator	Respondent State	Ultimate Winner		Total
		Claimant	Respondent	
<i>High Income</i>				
	High Income	3	6	9
	Upper-Middle Income	5	8	13
	Lower-Middle Income	3	8	11
	Low Income	1	0	1
<i>Upper-Middle and Lower-Middle Income</i>				
	High Income	1	1	2
	Upper-Middle Income	3	1	4
	Lower-Middle Income	2	3	5
	Low Income	1	1	2
Total		19	28	47

* None of the observed frequencies differed from what would be expected due to chance.

b. Effect Sizes

The effect sizes for these analyses all fell between what the literature deems a “small” and a “medium” effect.¹³⁶ More particularly, the potential effect size for High Income presiding arbitrators was $r = .25$. For the group of Upper-Middle and Lower-Middle Income presiding arbitrators the effect size was $r = .28$.

Although the chi-square test indicates that there was no effect of presiding arbitrator development status on outcome, the effect size suggests that the research may be underpowered. That is, a sample size of forty-seven is not large enough to distinguish reliably between significant and non-significant results.¹³⁷ Replicating the analysis with a larger sample could ascertain whether there is a detectable and reliable effect among the World Bank status of the respondent, the World Bank status of the presiding arbitrator, and the arbitration outcome.¹³⁸

136. COHEN, *supra* note 123, at 115.

137. With an effect size of .25 for a sample of thirty-four arbitrators, the power of the analysis is .30, and there is a 70% chance of making a Type II error. Likewise, with an effect size of .28 for a sample of thirteen arbitrators, the sample has less than .20 power and more than an 80% risk of making a Type II error. In other words, there is a risk that this analysis incorrectly retained the null hypothesis and that a larger sample is necessary to detect a statistically significant effect.

138. An a priori power analysis suggests that, to capture a potential effect size of $r = .25$, a sample of 480 final awards should create the requisite power. In order to obtain a sample with .80 power (20% likelihood of a Type II error), and an effect size of .25 ($r = .25$; $S = 120$), a sample of 480 final arbitration awards would be necessary to make sure there was a sufficient number in each of the eight conditions ($N = n(120/2) * k(8)$) to find the smallest effect size in the simple 2×2 .

3. *Discussion: Development and Winning Investment Treaty Arbitration*

None of the statistical analyses exhibited a statistically significant pattern among the development status of the presiding arbitrator, the development status of the respondent, and winning or losing an investment treaty arbitration.

The consistency in these results offers a powerful narrative that there is procedural integrity in investment arbitration. It undercuts the argument that development variables inappropriately affect outcome by unfairly harming the developing world or that arbitrators' decisions vary by virtue of their development backgrounds. Rather, the evidence suggests that neither the development status of the respondent nor the development background of the presiding arbitrator affects who wins investment arbitrations. Given the overall pattern of data, it is reasonable to infer that the system is functioning relatively well. This, in turn, undercuts the case for radical overhaul or elimination of investment treaty arbitration.

The results nevertheless merit further reflection. The nearly medium effect sizes connected with World Bank status suggest that more research, with a larger sample, is necessary before it will be possible to establish that development status has no reliable association with outcome. In contrast, future research would likely replicate the lack of a relationship between outcome and OECD status since the results were all non-significant and the effect sizes were either miniscule or small. In any event, more developed analysis, using additional data, further variables, and more sophisticated models, may reveal more information about the system's application.

C. Development Status and Amounts Awarded

Considering a pure win/loss scenario is a sensible first step in assessing arbitration outcomes. However, such a binary approach ignores variation in potential outcomes. This next section considers the variance in outcomes by exploring the intersection of the amounts tribunals awarded and the development status of respondent states and presiding arbitrators.

1. *OECD Status of Governments, OECD Status of Presiding Arbitrators, and Amounts Awarded by Tribunals*

Using the winsorized data,¹³⁹ a 2×2 between-subjects factorial design analyzed the effects of two independent variables, namely (1) the OECD status of the respondent state and (2) the OECD status of the presiding arbitrator, on the mean amounts that tribunals awarded.

139. The winsorized data form the focus of this analysis for two reasons. First, statistical outliers may influence the raw means. Second, using winsorized data, in which the parameters of the data are more likely to adhere to the statistical assumptions of the test, aids statistical conclusions validity.

a. Results of Null Hypothesis Significance Tests

The results indicated three things. First, there was no statistically significant interaction among the OECD status of the respondent state, the OECD status of the presiding arbitrator, and the amount that tribunals awarded ($F(1, 45) = .001; p = .97$). Second, the OECD status of the respondent state had no main effect on the amounts tribunals awarded ($F(1, 45) = .005; p = .94$). Third, there was no main effect of the OECD status of the presiding arbitrator on amounts tribunals awarded ($F(1, 45) = .900; p = .35$). Using raw data with statistical outliers, there were likewise no mean differences in the amounts awarded.¹⁴⁰

As hypothesized, the ANOVAs using winsorized data demonstrated that the mean amount awarded did not differ as a function of the OECD status of the respondent state, the OECD status of the presiding arbitrator, or an interaction between these two variables. Table 5 displays the mean amounts tribunals awarded against OECD and non-OECD countries, broken down by the OECD status of the presiding arbitrator using winsorized data.

Table 5: Winsorized Damage Amounts Awarded by Investment Treaty Tribunals as A Function of OECD Status of Respondent State and OECD Status of Presiding Arbitrators

<i>Presiding Arbitrator</i>	Respondent State	Amount		Total
		Mean	Standard Deviation	
<i>Non-OECD</i>				
	Non-OECD	1,808,556	2,676,995	10
	OECD	1,896,110	3,150,610	3
<i>OECD</i>				
	Non-OECD	1,073,470	1,830,895	21
	OECD	1,102,065	2,104,620	15
Total				49

No strong trends appeared in Table 5. Tribunals with presiding arbitrators from non-OECD countries awarded statistically equivalent amounts against both OECD and non-OECD respondent states (approximately US\$1.9 million and US\$1.8 million, respectively). Likewise, tribunals with presiding arbitrators from OECD countries awarded statistically equivalent amounts against OECD and non-OECD governments (approximately US\$1.1 million). The only potentially observable trend was that tribunals with presiding arbitrators from non-OECD countries made larger awards

¹⁴⁰. There was no statistically significant interaction among the OECD status of the respondent state, the OECD status of the presiding arbitrator, and the amount awarded ($F(1, 45) = 1.370; p = .25$). Likewise, there were no main effects for either the OECD status of the respondent state ($F(1, 45) = .002; p = .97$) or the OECD status of the presiding arbitrator ($F(1, 45) = .007; p = .93$).

than their OECD counterparts. However, this was not statistically significant.

The raw data, which included statistical outliers,¹⁴¹ did not reveal any significant mean differences, but the non-significant trends were slightly different. Raw data suggested a potential interaction among OECD status of the respondent, OECD status of the presiding arbitrator, and amount awarded. Tribunals with presiding arbitrators from OECD countries made higher average awards against OECD countries (approximately US\$19.5 million) and lower average awards against non-OECD countries (US\$1.2 million).¹⁴² In contrast, tribunals with presiding arbitrators from non-OECD countries made higher average awards against non-OECD countries (US\$21.5 million) and lower average awards against OECD countries (US\$1.9 million).¹⁴³

b. Effect Sizes

The effect sizes for the winsorized analyses were $r = .01$ (interaction), $r = .01$ (main effect of OECD status of respondent state), and $r = .14$ (main effect of OECD status of presiding arbitrator).¹⁴⁴

The effect sizes for the interaction and the main effect of OECD status of the respondent were tiny.¹⁴⁵ Since the statistical relationships were not significant and the strength of those non-significant relationships was potentially trivial, it is unlikely that these analyses suffer from a power problem requiring a larger sample. In other words, given the potential minor effect on outcome, the relationship between amounts awarded and the interaction

141. For example, *CME Czech Republic B.V. (Neth.) v. Czech Rep.*, Mar. 14, 2003, available at http://ita.law.uvic.ca/chronological_list.htm, positively skews the overall mean damages. *CME* involved a presiding arbitrator from an OECD country awarding US\$269 million against an OECD country.

142. For tribunals with presiding arbitrators from OECD countries, in awards against OECD countries, the mean award was US\$19,445,260 (standard deviation ("SD") = 69,396,386; $n = 15$). In awards against non-OECD countries, the mean award was US\$1,196,429 (SD = 2,193,461; $n = 21$).

143. For tribunals with presiding arbitrators from non-OECD countries, in awards against non-OECD countries, the mean award was US\$21,479,250 (SD = 45,126,901; $n = 10$). In awards against OECD countries, the mean award was US\$1,896,110 (SD = 3,150,610; $n = 3$).

144. Using the raw data, the effect sizes for the analyses were $r = .17$ (interaction), $r = .01$ (main effect of OECD status of respondent state), and $r = .01$ (main effect of OECD status of presiding arbitrator). The effect sizes for the non-significant main effects are less than small, which means that the analyses likely do not suffer from a power problem and would require a minimum sample of 781 cases to assess the effect reliably.

The effect size of the interaction was between small and medium. COHEN, *supra* note 123, at 115. Because of the potential size of the effect, there is a possibility that the analysis lacks power and that a sample size of 49 is insufficient to decide definitively whether there is no relationship among the OECD status of respondent states, the OECD status of presiding arbitrators, and the amounts awarded. For a sample with 80% power and a .20 effect size ($S = 191$), 382 awards would be necessary to ascertain reliably the potential scope of the effect in the four conditions of the interaction ($N = n(191/2) * k(4)$). The effect size $r = .20$ was selected because of (1) the effect size of the interaction ($r = .17$) and (2) the fact that the effect sizes of the simple effects were in the order of .20. Using a mean square error of 1,908,000,000,000,000, the simple effect for non-OECD presiding arbitrators for non-OECD and OECD respondent states was $r = .22$ and the simple effect for OECD presiding arbitrators for non-OECD and OECD respondents was $r = .20$.

145. COHEN, *supra* note 123, at 115.

R

R

or main effect of the respondent's OECD status does not necessarily warrant further analysis.¹⁴⁶

The main effect for the OECD status of the presiding arbitrator, however, was different from the other effect sizes for the winsorized data. It was also different from the effect size for the main effect of presiding arbitrator in the raw data. The effect size for the main effect of the OECD status of the presiding arbitrator was .14, namely a possibly small to medium effect.¹⁴⁷ Because of the potential size of the effect, there is a possibility that the analysis lacks power and a sample size of 49 was insufficient to decide definitively whether the OECD status of the presiding arbitrator affects the mean amount awarded.¹⁴⁸ Given the effect size, 343 awards would be necessary to ascertain whether the OECD status of the presiding arbitrator affects amounts awarded¹⁴⁹ with appropriate statistical confidence.

2. *World Bank Status of Governments, World Bank Status of Presiding Arbitrators, and Amounts Awarded by Tribunals*

A 2×4 between-subjects factorial design analyzed the effects of (1) the World Bank status of the respondent state and (2) the World Bank status of the presiding arbitrator on amounts tribunals awarded.

a. *Results of Null Hypothesis Significance Tests*

The results from winsorized data indicated three things. First, there was no statistically significant interaction among the World Bank status of the respondent, the World Bank status of the presiding arbitrator, and the amount tribunals awarded ($F(3, 41) = .536; p = .66$). Second, there was no main effect of the World Bank status of the respondent on amounts tribunals awarded ($F(3, 41) = 1.285; p = .29$). Third, there was no main effect of the World Bank status of the presiding arbitrator on amounts tribunals awarded ($F(1, 41) = .365; p = .55$). Using raw data with statistical outliers, there were likewise no mean differences in the amounts awarded.¹⁵⁰

146. When assessing the main effect of the development status of either the presiding arbitrator or the respondent state, with $r = .10$ and 80% power, a minimum sample of 781 cases would be necessary to assess any effect reliably.

147. COHEN, *supra* note 123, at 115.

148. With a sample of 49 and an effect size of .14 ($r = .15$), the power of this analysis is between .20 and .30, which means that there is a 70–80% chance of making a Type II error. As such, there is a 70–80% risk that this analysis incorrectly retained the null hypothesis, and a larger sample would be necessary to detect a statistically significant effect.

149. Under this estimate, the researcher would look for a sample with .80 power (20% likelihood of a Type II error) and an effect size of $r = .15$ ($S = 343$). A sample of 343 final arbitration awards would be necessary to ensure a sufficient sample for the two different conditions in the main effect of the OECD status of the presiding arbitrator ($N = n(343 / 2) * k(2)$).

150. There was no statistically significant interaction among the World Bank status of the respondent, the World Bank status of the presiding arbitrator, and the amount awarded ($F(3, 41) = .723; p = .54$). There was no main effect for either the World Bank status of the respondent ($F(3, 41) = .194; p = .90$) or the World Bank status of the presiding arbitrator ($F(1, 41) = .075; p = .79$).

As hypothesized, the mean damages that tribunals awarded did not differ as an independent function of either the World Bank status of the respondent state or the World Bank status of the presiding arbitrator. Likewise, as hypothesized, the mean damages awarded did not differ as a result of an interaction between these two variables. Table 6 displays the mean damage awards from the winsorized data against countries on the basis of their World Bank status and the World Bank status of the presiding arbitrator.

Table 6: Winsorized Damage Amounts Awarded by Investment Treaty Tribunals as a Function of World Bank Status of Respondent State and World Bank Status of Presiding Arbitrators

Presiding Arbitrator	Respondent State	Amount		Total
		Mean	Standard Deviation	
<i>High Income</i>				
	High Income	1,103,070	2,127,403	9
	Upper-Middle Income	1,318,088	2,131,167	13
	Lower-Middle Income	657,548	1,700,852	12
	Low Income	2,060,248	1,314,353	2
<i>Upper-Middle and Low-Middle Income</i>				
	High Income	77,657	109,824	2
	Upper-Middle Income	2,951,876	3,073,025	4
	Lower-Middle Income	1,227,107	2,494,713	5
	Low Income	2,837,769	4,013,211	2
Total				49

Post-hoc follow-up analyses for the interaction and main effects were conducted using Tukey’s HSD test¹⁵¹ to determine whether specific combinations were significantly different. In particular, pairwise comparisons assessed the presence of statistically significant simple effects at the micro level. Analysis of raw data did not isolate any statistically significant simple effects.¹⁵²

Analysis of winsorized data was slightly different. For the interaction, the HSD minimum mean difference was US\$3,968,695, and none of the smaller effects came close to reaching statistical significance. For the main effect of the World Bank status of the respondent, the minimum mean difference was US\$2,353,063. There were, however, two statistically significant effects in the follow-up comparisons. In particular, where the presiding arbitrator

151. Tukey’s honestly significant difference (“HSD”) test provides a follow-up for significance testing to determine, at a smaller level, whether there are statistically significant differences among different conditions. GRAVETTER & WALLNAU, *supra* note 123, at 365.

152. The HSD minimum mean differences for the raw data were (1) US\$81,595,507 for the interaction, and (2) US\$48,378,458 for the main effects. None of the follow-up comparisons amongst the means revealed statistically significant simple effects.

came from an Upper-Middle or Lower-Middle Income country, there were statistically significant differences between the awards against (1) High Income countries versus Upper-Middle Income countries (a US\$2,874,219 mean difference), and (2) High Income countries versus Low Income countries (a US\$2,760,112 mean difference). In both comparisons, tribunals with presiding arbitrators from Middle Income countries rendered statistically different—and lower—awards against High Income countries.

The results in Table 6 merit two further observations. First, there was a phenomenon whereby tribunals with presiding arbitrators from both High and Upper-Middle/Lower-Middle Income countries awarded greater damages against Low Income countries than against other countries. This phenomenon, however, was not statistically significant and was also not observable from the raw data.¹⁵³ Second, tribunals with presiding arbitrators from Upper-Middle/Low Income countries rendered the highest awards issued against Upper-Middle Income countries. Lower-Middle Income countries had the third largest awards.

The raw data exhibited slight differences. There, tribunals rendered the largest awards against respondents with the same development status as the presiding arbitrator. This meant that tribunals with presiding arbitrators from Upper-Middle or Lower-Middle Income countries rendered the largest awards against Upper-Middle and Lower-Middle Income respondents.¹⁵⁴ Likewise, tribunals with presiding arbitrators from High Income countries rendered the largest mean amounts against High Income respondents.¹⁵⁵ These later phenomena were not observable, however, in the winsorized data.

b. Effect Sizes

The effect sizes for the analyses were $r = .19$ (interaction), $r = .29$ (main effect of World Bank status of respondent state), and $r = .09$ (main effect of World Bank status of presiding arbitrator).

Since the statistical relationship for the main effect of the World Bank status of the presiding arbitrator was not statistically significant and the potential strength of the non-significant relationship was very weak ($r = .09$), it is unlikely that that this aspect of the analysis suffers from a power problem that requires a larger sample. This suggests that by itself, the exis-

153. See *infra* notes 154–55.

154. Tribunals with presiding arbitrators from Upper-Middle and Lower-Middle Income countries awarded the highest amounts against Upper-Middle Income states (US\$34,832,922; SD = 65,624,852; $n = 4$) and against Lower-Middle Income states (US\$14,398,710; SD = 31,939,967; $n = 5$). They rendered the smallest mean awards against High Income (US\$77,657; SD = 109,824; $n = 2$) and Low Income (US\$4,500,000; SD = 6,363,961; $n = 2$) states.

155. Where the presiding arbitrator was from a High Income country, the mean awards were: (1) US\$30,451,788 against High Income respondents (SD = 8,976,962; $n = 9$), (2) US\$2,180,030 against Upper-Middle Income respondents (SD = 4,688,136; $n = 13$), (3) US\$856,411 against Lower-Middle Income respondents (SD = 2,354,715; $n = 12$), and (4) US\$2,060,248 against Low Income respondents (SD = 1,314,354; $n = 2$).

tence of any effect from the main effect of development status of the presiding arbitrator on amounts awarded may simply be too minor to warrant further analysis.¹⁵⁶

The effect sizes of the interaction ($r = .19$)¹⁵⁷ and the main effect of the status of the respondent state ($r = .29$),¹⁵⁸ however, were larger and could have been between a small and a medium effect.¹⁵⁹ The sample size of this analysis was only 49, and the power appears to be low.¹⁶⁰ It might be possible to detect the main effects and interaction with acquisition and analysis of more cases. Given the effect size of the existing data related to the interaction, a sample size of 764 awards would be necessary to ascertain reliably the potential scope of the effect.¹⁶¹

3. Discussion: Development and Amounts Awarded

At the macro-level, none of the analyses revealed a statistically significant relationship among the development status of the respondent state, the development status of the presiding arbitrator, or an interaction between those variables, and amounts awarded.

First, neither the OECD nor the World Bank status of a respondent state had a statistically significant link to amounts awarded. The lack of a main effect for a respondent's development status stands in sharp contrast to the assertions that investment treaty arbitration unfairly privileges the developed world or improperly harms the developing world.

Second, neither OECD nor World Bank status of the presiding arbitrator had a statistically significant relationship with amounts awarded. The absence of a main effect for the presiding arbitrator's development background challenges the assertion that arbitrators from the developed and developing worlds assess cases differently. One might suggest that this may be a function of the small, perhaps homogenous, pool of arbitrators from the develop-

156. The main effect of World Bank status of the presiding arbitrator using raw data was likewise tiny ($r = .04$). In future research, looking for $r = .10$ and a power of .80, a sample of 781 final awards should be sufficient to replicate and isolate the main effect of World Bank status of the presiding arbitrator on amounts awarded.

157. The effect size estimate of the interaction from the raw data was similar ($r = .22$).

158. The effect size estimate of the World Bank status of the respondent using raw data was smaller ($r = .12$) but was still between a potentially small and medium effect.

159. COHEN, *supra* note 123, at 115.

160. With a sample of 49, the power of the interaction ($r = .19$) is between .20 and .30, which means that there is a 70–80% chance of making a Type II error. The power of the main effect for the World Bank status of the respondent state ($r = .29$) is between .50 and .60, which means that there is a 40–50% chance of making a Type II error. A larger sample could ascertain the potential existence of a statistically significant effect.

161. This estimate is based upon a sample with .80 power (20% likelihood of a Type II error) and an effect size of $r = .20$ ($S = 191$). A sample of 764 final arbitration awards would be necessary to ensure a sufficient number for all of the conditions of the interaction ($N = n(191/2) * k(8)$). A more conservative approach could use $r = .15$ ($S = 343$) as a baseline. The requisite number of awards would then be 1,372 ($N = n(343/2) * k(8)$). Although one might also use the effect size of the main effect of respondent state status ($r = .29$) and use a conservative estimate ($r = .25$; $S = 120$), the sample size would then be 240 [$N = n(120/2) * 4$] and would not be sufficient to capture the potential interaction.

ing world, which implies that future analysis may benefit from a broader pool of data. It may also suggest that future research may need to expand the measurement of development status in order to consider the development status of the tribunal as a whole. It may also be helpful to explore variables related to arbitrators' background, such as gender, country of legal education, professional background, and experience in different roles (for instance, private practice, government work, or affiliation with civil society groups).

Third, there was no interaction among development variables and amounts awarded. This undercuts the possibility that, while there might not be an overall effect for the development status of either respondents or presiding arbitrators, the two variables may operate together to affect investment arbitration outcome unfairly.

The absence of a reliable statistical relationship makes it more challenging to assert that development status improperly affects the outcome of investment arbitration. It also undercuts the case for radical overhaul or elimination of arbitration. Nevertheless, the results merit further consideration for two key reasons. First, follow-up pairwise comparisons using the winsorized data detected significantly different outcomes in two narrow situations. Second, the effect sizes of certain analyses suggest that further research, based upon a broader pool of data, is warranted before inferring that the current results establish population parameters.

Generally, the follow-up pairwise comparisons did not reveal any statistically significant simple effects. This held true for OECD status generally, irrespective of whether the research model employed raw or winsorized data. This was also generally true for World Bank status using raw data. There was, however, a small wrinkle as regards winsorized data. In keeping with the overall pattern, there was no statistically significant difference in amounts awarded by presiding arbitrators from High Income countries. That was not always so for presiding arbitrators from Middle Income countries. Awards from Middle Income presiding arbitrators against High Income and Lower-Middle Income respondents were statistically equivalent. There were two simple effects in which amount awarded differed depending upon the respondent's World Bank status. Specifically, tribunals made smaller awards against High Income countries as compared to (1) Upper-Middle Income countries, and (2) Low Income countries.¹⁶² The results suggest that very narrow, targeted reform of investment arbitration may be necessary to address a particularized problem.

162. The statistically significant effects came from winsorized data. While winsorization was used to ensure that the data adhered to the normality assumptions of the statistical tests in order to enhance the statistical conclusion validity, the raw data exhibited an interesting (but statistically non-significant) phenomenon. In particular, the raw data suggested that panels with presiding arbitrators from the developed world awarded greater amounts against developed countries, whereas presiding arbitrators from the developing world awarded greater amounts against developing countries.

In order to ensure proper understanding of these effects, it is important to keep three things in mind. First, they come from a narrow set of follow-up pairwise comparisons and have different results from the majority of other follow-up comparisons. Second, some cells had very limited data. For example, for tribunals with a presiding arbitrator from a Middle Income country, there were only two cases involving High Income respondent states and two cases involving Low Income respondents. As the two statistically significant pairwise comparisons involve these countries, inferences from the data are limited. Third, the two significant simple effects, namely the award differential when there is a presiding arbitrator from a Middle Income country, for (1) High versus Upper-Middle Income countries and (2) High versus Low Income countries, are influenced by only two cases. In particular, the two cases involving High Income countries with presiding arbitrators from Upper-Middle or Lower-Middle Income countries are (1) *Maffezini v. Spain*, with Francisco Orrego Vicuña as the presiding arbitrator,¹⁶³ and (2) *ADF Group, Inc. v. United States*, with Florentino P. Feliciano as the presiding arbitrator.¹⁶⁴ It very well may be, for example, that *Maffezini* is not a representative case. Although Vicuña has served as a presiding arbitrator in several arbitrations,¹⁶⁵ the facts of *Maffezini* were unusual. First, of all the cases in the sample, it involved the lowest damage amount claimed (approximately US\$155,314). Second, it was one of two cases where the respondent did not dispute damages.¹⁶⁶ If either *Maffezini* or *ADF* (or both) fail to represent the larger population, then the simple effects will be suspect.

The effect sizes also suggest that further research is warranted. Many analyses did not achieve statistical significance and had effect sizes that were so tiny as to be of little practical effect.¹⁶⁷ Nevertheless, there were several circumstances where, despite non-significant results, the effect sizes indicated the possible presence of between a small and a medium effect related to development variables.¹⁶⁸ Out of extra caution in this sensitive area, it would therefore be prudent to engage in more research, with a larger sample, before establishing a population parameter that development status is not reliably associated with outcome. Indeed, it is possible that the simple

163. *Maffezini (Argentina) v. Spain*, ICSID CASE NO. ARB/97/7 (Jan. 25, 2000), 16 ICSID REV.-FOREIGN INV. L.J. 212 (2001).

164. *ADF Group, Inc. (Canada) v. United States*, ICSID CASE NO. ARB(AF)/00/1 (Jan. 9, 2003), 18 ICSID REV.-FOREIGN INV. L.J. 195 (2003).

165. Franck, *Evaluating Claims*, *supra* note 16, at 79 n.362, identifies nine cases where Vicuña is an arbitrator. In each one of them—*Fedax*, *Camuzzi*, *Maffezini*, *Enron*, *PSEG Global*, *Occidental Exploration*, *CMS Gas Transmission*, *Joy Mining*, and *Sempra Energy*—Professor Vicuña was the presiding arbitrator.

166. Franck, *Evaluating Claims*, *supra* note 16, at 54, 57–58.

167. For instance, the effect sizes for OECD main effects of respondent states and presiding arbitrators, using raw data, were .01. Likewise, the effect sizes for OECD status related to the interaction and main effect of respondent status, using winsorized data, were .01. The effect size for the World Bank status of presiding arbitrators using raw data was .04. See *supra* notes 144, 156 and accompanying text.

168. See *supra* notes 144, 156–58 and accompanying text (identifying various effect sizes related to development status variables for both raw and winsorized data).

R

R

R

R

effects in this research will disappear in a larger sample with greater variance and, hopefully, a larger pool of respondents and presiding arbitrators.

D. *Discussion of the Overall Results*

The overall results of the various statistical analyses demonstrate that, irrespective of the definition of development status, there was no statistically significant relationship among the development status of the respondent, the development status of the presiding arbitrator, and outcome. This held true for both (1) winning or losing an investment treaty arbitration, and (2) the amounts tribunals awarded.

These results suggest that there is some cautiously optimistic news about investment treaty arbitration. Based upon the significance tests, there is no obvious evidence of substantial dysfunction related to development, suggesting that arbitration is not per se biased in favor of either the developed or the developing world. First, presiding arbitrators came from a variety of different development backgrounds.¹⁶⁹ Second, the development status of a presiding arbitrator's home jurisdiction does not appear on the whole to be associated with arbitration outcome.¹⁷⁰ Given that the data did not suggest a reliable link between development status and outcome, the evidence begins to suggest that the investment treaty arbitration system appears to be functioning reasonably well at the macro level. Such results suggest that radical overhaul or complete rejection of the international arbitration paradigm is, at present, unwarranted.

The presence of two statistically significant simple effects in the amounts awarded, however, suggests tempering this "good news," as certain permutations merit further reflection. The finding that tribunals with presiding arbitrators from Middle Income countries sometimes rendered lower awards against High Income countries is noteworthy. While some might suggest that presiding arbitrators from the developing world may treat respondents from the developing world unfairly, this could overstate the point.

There were differences in amounts awarded in a few narrow circumstances. Those circumstances further rest primarily upon two awards, which may not be representative of the larger population. This suggests two key

169. This diversity addresses, in part, concerns that investment treaty arbitration is a tool of economic oppression whereby arbitrators from the developed world control a process that has implications for both developed and developing countries. Given the small numbers of arbitrators from Middle Income countries and the lack of Low Income presiding arbitrators, however, there is room for improvement. See discussion *supra* Part V.B.1.

170. It is prudent to observe that, in terms of sheer numbers, there are far more presiding arbitrators from the High Income and OECD jurisdictions than from others. The absence of any presiding arbitrators from Low Income countries and the low cell counts of Upper-Middle and Lower-Middle Income presiding arbitrators requires reflection. It suggests that while the arbitration processes may be balanced in theory, there are areas of concern in the system's application. This may be part of the reason that during his tenure at ICSID Roberto Dañino "made it a point to diversify the pool of ICSID arbitrators to include a greater number of women and developing country nationals." *Roberto Dañino Leaves ICSID*, NEWS FROM ICSID (ICSID (W. Bank), Washington, D.C.), Winter 2005 (Vol. 22, No. 22), at 1.

points. First and foremost, further research is necessary to replicate these findings to assess whether they are real population parameters or the result of statistical chance.

Second, these narrow findings, which may not be replicable, may nevertheless contribute to perceived concerns about arbitration's procedural fairness. Particularized structural reforms that are targeted to eliminate perceived concerns, redress real issues, or prevent the growth of potential problems may actually have tangible value. If there is a demonstrable and replicable issue of unfair outcomes that casts doubt on arbitration's procedural integrity, a targeted reform might usefully redress a specific problem. If, however, the results do not recur, there is still value in enacting narrow measures that might redress perceived imbalances in arbitration. Taking the concerns of stakeholders seriously and enacting reforms to address those concerns responsibly can enhance procedural justice.¹⁷¹ Stakeholders can experience process control and obtain acknowledgement of their concerns. Further, they may have greater confidence that arbitrators will resolve disputes fairly.¹⁷² In this manner, the perceived legitimacy of arbitration can be enhanced.¹⁷³

The overall outlook is reasonably positive. While there may be areas of concern for future investigation, preventative measures may avert degradation of arbitration's integrity and enhance its longevity. Put another way, if investment arbitration is experiencing "growing pains," the statistical diagnosis provided here suggests that radical treatment is unnecessary. Instead, minor structural adjustments might usefully address perceived problems. Future monitoring should assess whether more invasive treatment is required or whether these structural adjustments will address stakeholders' concerns.

E. Understanding the Limitations of the Analysis

It is important to understand these results within their proper context, since they pertain to arbitration awards and not other forms of dispute resolution. First, there may be limitations to the strength of the inferences, as they may not reflect population parameters. For the two statistically significant simple effects, it is necessary to draw inferences with caution, given the small cell counts and lack of certainty about whether the subsets are repre-

171. See TOM R. TYLER ET AL., *SOCIAL JUSTICE IN A DIVERSE SOCIETY* 78 (1997) ("[P]eople choose the procedures that they would like to use to resolve their disputes in large part through assessments of procedural fairness . . . people do not simply choose the procedure they think will allow them to win. They are actually interested in finding a procedure that they think will be fair.").

172. See *id.* at 87–90 (referring to the benefits of process control and its effect upon party assessment of outcomes and the quality of the dispute resolution process).

173. See Yuen J. Huo et al., *Superordinate Identification, Subgroup Identification, and Justice Concerns: Is Separatism the Problem; Is Assimilation the Answer?*, 7 *PSYCHOL. SCI.* 40 (1996) (arguing that even when people are dissatisfied with outcomes, they can believe that a process is legitimate because authorities demonstrate respect for stakeholders during conflict management).

sentative. Replication with expanded data is necessary to avoid establishing a population parameter that may be due to chance alone.

Second, there may be issues about the validity of the statistical conclusions. Effect sizes suggest that the power of the research is relatively low.¹⁷⁴ It would be prudent to establish a broader pool of data, based on a priori power analysis, to confirm, clarify, contradict, or supplement these findings.¹⁷⁵

Given the nascent nature of investment treaty arbitration and related empirical research, the statistical models used were blunt and the variables were necessarily 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.¹⁷⁶ More complex designs such as multiple group ANOVAs analyzing other variables and/or combinations of variables, factorial analysis, multi-variate regressions, control variables, or matching variables might help reflect the complexities of reality and simultaneously decrease the likelihood of statistical error.¹⁷⁷

Third, 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 analyses was coded from awards that were publicly available as of June 1, 2006. Since there is now nearly three additional years' worth of data to gather and analyze, future research should replicate the analysis. Also, there are issues related to case selection bias.¹⁷⁸ While there have been anecdotal comments suggesting that the data is representative,¹⁷⁹ this has not been empirically confirmed. It is possible

174. Even when retaining the null hypothesis that there is no difference between the different conditions of an independent variable (i.e., countries from the developed versus developing world) there is a possibility that this conclusion is wrong. This is a Type II error, in which the researcher incorrectly determines that there is no statistically significant effect. Effect sizes can be used to estimate the likelihood of a Type II error and the potential strength of an independent variable's effect. See PAUL CONNOLLY, *QUANTITATIVE DATA ANALYSIS IN EDUCATION* 206 (2007).

175. This may require a sample as large as 1,562, which is the largest sample necessary to identify the smallest effects. This is the .11 effect size of the main effect of a respondent state's World Bank development status, using raw data. See *supra* note 132. If, however, one used the largest sample necessary to isolate the smallest effect size that came from winsorized data, the sample would need to be in the order of 1,372 to identify the interaction related to the World Bank development status of the presiding arbitrator and respondent state. See *supra* note 161.

176. Additional control variables could minimize the risk of statistical confounds. For example, given the limited and missing data in the present database, it was not possible to control for variables such as differences between amounts claimed versus amounts awarded. Future research might have enough data to usefully add this factor and consider other variables such as the number of arbitrators, the gender of arbitrators, the institutions administering the arbitrations, and the identity of lawyers representing the parties.

177. See Franck, *Evaluating Claims*, *supra* note 16.

178. *Id.* at 19–20.

179. See, e.g., GUS VAN HARTEN, *INVESTMENT TREATY ARBITRATION AND PUBLIC LAW* viii (2007) ("Many (probably the great majority) of investment treaty awards . . . are publicly available in some cases

R

R

R

that these analyses may be limited and/or systematically biased, particularly as the author knows of assertions that there is a recent trend for investors to avoid ICSID, which may mean that fewer cases are finding their way into the public domain. Further research should consider the effect of this possible case selection bias.¹⁸⁰

Fourth, we should monitor the results carefully, as there is not complete agreement in the direction of potential trends in the winsorized and raw data.

As a result of these cautionary considerations, more research is required to create the power, stability, statistical control, and enhanced validity necessary to reach truly compelling conclusions. Given that UNCTAD's research suggests that there are currently 290 known investment treaty arbitrations,¹⁸¹ it may take several years before a sufficient pool of awards is available to run the requisite analysis. This makes future research challenging but does not diminish the importance of replication and convergence. As the data pool expands, analysis will be possible, but it is important to recognize the potential limitations of its statistical power.

VI. DESIGNING DISPUTE RESOLUTION SYSTEMS

The appropriate design of dispute resolution systems for managing investment treaty-related conflict is an area of emerging interest.¹⁸² Dispute systems design is the process of analyzing existing patterns of disputing, creating new processes, and implementing and testing the new designs to ensure resolution of disputes in accordance with stakeholder needs.¹⁸³ This quantitative research, which aids the diagnosis of the existing system, supports the initial steps of designing more effective conflict management systems.¹⁸⁴ Given the initial results and related inferences from the statistical analyses, it is important to ascertain the normative implications.

Statistical research has a role in designing dispute resolution systems. The theme from the present research suggests creating targeted adjustments when there is a valid and reliable diagnostic demonstrating the value of such modifications, while recognizing the possible limitations when implement-

only via the internet"); Franck, *Evaluating Claims*, *supra* note 16, at 17–18 (describing the sample of publicly available awards that came primarily from cases freely available via the internet). R

180. See Franck, *Evaluating Claims*, *supra* note 16, at 17–21. R

181. U.N. Conference on Trade and Dev. [UNCTAD], *Latest Developments in Investor-State Dispute Settlement*, 1, U.N. Doc. UNCTAD/WEB/ITE/IIA/2008/3 (2008).

182. See generally Franck, *Dispute Systems Design*, *supra* note 26; 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 (2005); Barton Legum, *The Difficulties of Conciliation in Investment Treaty Cases: A Comment on Professor Jack C. [sic] Coe's 'Toward a Complementary Use of Conciliation in Investor-State Disputes—A Preliminary Sketch'*, MEALEY'S INT'L ARB. REP., Apr. 2006, at 23; Noah Rubins, *Comments to Professor Jack C. [sic] Coe's Article on Conciliation*, MEALEY'S INT'L ARB. REP. (Apr. 2006), at 21. R

183. See Franck, *Dispute Systems Design*, *supra* note 26, at 167.

184. See *id.* at 178, 184–85, 187–90. R

ing policy changes. Here, the initial statistical diagnostic suggests two things. First, the investment treaty arbitration system is functioning relatively well at the macro level, which suggests that major structural reform or the eradication of arbitration is unnecessary. Second, the two statistically significant simple effects suggest there may be specific areas of potential concern that require particularized solutions.¹⁸⁵ Even if replication does not reveal the same effects, implementing targeted reforms to address particularized concerns can begin to address the perceived (if not actual) concerns of stakeholders, enhance procedural justice by improving structural safeguards, and promote the legitimacy of the arbitration process.

A. *General Issues*

As a general matter, if the system does not appear per se biased in favor of either the developed or the developing world, it suggests that the system is working reasonably well. The overall lack of a statistically significant relationship among the development status of the presiding arbitrator, the development status of the respondent state, and outcome suggests that the system is not unfairly balanced per se. Initial results suggest that outcome may be due less to development status or a north-south divide than to the merits of individual claims or other variables.¹⁸⁶ Promoting neutral, merits-based adjudication can encourage the rule of law and prevent dispute resolution from becoming a tool for economic oppression based upon spurious variables related to development. Given that the data suggest that the system appears healthy at the macro level, there may not be a need for major structural overhaul of arbitration mechanics or wholesale abandonment of arbitration. In other words, if statistics do not reflect a systemic problem, there is evidence that substantial modification or re-design may not be necessary.

Despite this general cautiously good news, however, there are areas for improvement and prudence. Section B therefore explores various opportunities stakeholders may take to improve investment treaty dispute resolution, and section C suggests areas of caution so that counsel and arbitrators can take even greater care during the adjudicative process to promote procedural justice.

185. Initial diagnoses or empirical evaluations are not the only influences on the design of dispute resolution processes. Rather, they are the first steps in a process facilitating systematic change on the basis of mutually agreed principles, which in turn creates practical solutions for improving conflict resolution's efficiency. *Id.* at 178, 184–85.

186. Given the general lack of relationship between outcome and development, one might suggest that analyzing development should not be a core focus. If, upon replication of the results, there continues to be no reliable link between the development status and outcome, this would support the conception that development is not a critical aspect of the variance related to arbitration outcome and that other variables may be more critical.

B. Opportunities for Improving Investment Treaty Dispute Resolution

Despite the cautiously good news about the integrity of the dispute resolution process at the macro level, there are issues concerning the investment treaty arbitration system's operation at the micro level. The results suggest that it is necessary to temper this cautious optimism properly. Empirical diagnosis found two simple effects where presiding arbitrators from the developing world made larger awards against developing countries and smaller awards against developed countries. Also, while not statistically significant and potentially affected by statistical outliers, there was a phenomenon whereby tribunals with arbitrators from developed countries rendered higher awards against respondents from the developed world, whereas tribunals with presiding arbitrators from the developing world rendered higher awards against countries from the developing world. While they bear watching and are worthy of replication, the initial results suggest that there may be areas that require—or would simply benefit from—targeted interventions to improve the design of future dispute resolution systems. Such an approach, even if future research does not replicate these results, could make the system more responsive to stakeholders' concerns and enhance arbitration's procedural integrity.

Even at this stage, it is worth considering what should be done to address the potential bias of arbitrators from the developing world in favor of the developed world. Follow-up research could involve structured interviews with presiding arbitrators to isolate variables that affect their decision making and case outcomes. It might also involve further quantitative research that controls for other variables that could potentially influence outcome to obtain greater statistical control.¹⁸⁷

One might also theorize, on a preliminary basis, about what could usefully be done to address both the phenomenon and the potential causes, in order to make targeted recommendations for how to address the potential problem. Such normative fixes might occur on various levels.¹⁸⁸ For exam-

187. True causal analysis, which might indicate that development variables *cause* outcomes, is not yet available. Causal interpretability generally comes from experimental models with random assignment to conditions and experimental procedures. See NOREEN L. CHANNELS, *SOCIAL SCIENCE METHODS IN THE LEGAL PROCESS* 64–68 (1985); D.A. SPROTT, *STATISTICAL INFERENCE IN SCIENCE* 198–99 (2000); URDAN, *supra* note 15, at 80; Lorne Hulbert, *Experimental Research Designs*, in *DOING SOCIAL PSYCHOLOGY RESEARCH: RESEARCH 10*, 39–40 (Glynis M. Breakwell ed., 2004). Investment treaty arbitration is a natural group design in which parties self-select into categories. In the absence of experimental control, causal interpretability is limited. Statistical control might serve as a proxy for experimental control. See HERBERT B. ASHER, *CAUSAL MODELING* 11–13 (2d ed. 1983); William N. Dunn, *Campbell's Experimenting Society: Prospect and Retrospect*, in *THE EXPERIMENTING SOCIETY: ESSAYS IN HONOR OF DONALD T. CAMPBELL* 1, 19–20 (William N. Dunn ed., 1998); *see also* Epstein & King, *supra* note 15, at 37 (discussing limitations of causal research). R

188. Changes might occur in various manners. They might include legislative reforms in drafting treaties, barrier building to restrict access to arbitration, rejection of the arbitration process, or the creation of incremental structural safeguards to modify the existing arbitration process. Susan D. Franck, *The Legitimacy Crisis in Investment Treaty Arbitration: Privatizing Public International Law Through Inconsistent Decisions*, 73 *FORD. L. REV.* 1521, 1587–1610 (2005) [hereinafter Franck, *Legitimacy Crisis*]. R

ple, one might consider making changes to the system's application rather than requiring revision of treaties, institutional rules, and national laws. It would also be prudent to reconsider the normative design of the system.

1. *Improving the Application of the Existing Arbitration Process*

Having statistically significant simple effects from a small sample of presiding arbitrators from Middle Income countries has implications. First, it means that there are reasonable concerns about the reliability of the statistical conclusions—namely that tribunals with presiding arbitrators from Middle Income countries treat countries differently on the basis of their development background. Second, it means that having a broader pool of presiding arbitrators from the developing world would aid replication of these findings. Additional information can facilitate making more informed, and presumably better, policy choices.

Third, having a broader pool of arbitrators may enhance the perceived legitimacy of arbitration. There is value to having a broader cross-section of decision makers and minimizing a perceived democracy deficit between arbitrators and those affected by arbitration.¹⁸⁹ Beyond this, one might wonder whether arbitrators from the developing world feel a need to penalize the developing world or privilege the developed world in an effort to be part of a “club” that plays by an unspoken set of international rules. Dezalay and Garth explain, in the context of international commercial arbitration, that sometimes “an arbitrator from the third world must find a way to gain access to [] credibility.”¹⁹⁰ It may be that arbitrators from the developing world (particularly those seeking repeat appointments) believe that rulings in favor of the developed world are the price of admission to the “club.”

While the current research provides limited support for the “price of admission” narrative,¹⁹¹ the overall results provide greater support for the theory that an arbitrator's development background does not affect outcome.¹⁹² One might imagine that professional integrity and the increase in professional prestige associated with adjudicative independence provides a counterweight to the power of the “price of admission” narrative. Moreover, with a critical mass of arbitrators from the developing world, the perceived phenomenon could disappear. Theoretically, there may be less pressure to

189. Increasing adjudicative diversity may be associated with enhanced legitimacy, partly because of the breadth of representation available in the decision-making process. See Richard Devlin et al., *Reducing the Democratic Deficit: Representation, Diversity and the Canadian Judiciary, or Towards a 'Triple P' Judiciary*, 38 ALTA. L. REV. 734, 753 (2000); *Final Report and Recommendations of the Eighth Circuit Gender Fairness Task Force*, 31 CREIGHTON L. REV. 9, 12 (1997); Sherrilyn A. Ifill, *Racial Diversity on the Bench: Role Models and Public Confidence*, 57 WASH. & LEE L. REV. 405, 436 (2000); Sherrilyn A. Ifill, *Judging the Judges: Racial Diversity, Impartiality and Representation on State Trial Courts*, 39 B.C. L. REV. 95, 101 (1997); Sally J. Kenney, *Breaking the Silence: Gender Mainstreaming and the Composition of the European Court of Justice*, 10 FEMINIST LEGAL STUD. 257 (2004).

190. DEZALAY & GARTH, *supra* note 44, at 25.

191. See *supra* notes 87–94 and accompanying text.

192. See *supra* notes 95–99 and accompanying text.

decide cases in particular ways if there is sufficient variability, or decreased tokenism, in the pool of presiding arbitrators. A critical aspect of re-tooling the current arbitration process—without revising the law—therefore involves expanding the pool of arbitrators from the developing world and introducing variability into the population. Given that parties and institutions control appointments, the question becomes how to expand the pool and promote broader appointment choices.

Part of the solution involves addressing the pipeline problem and ensuring that there is a broader cross-section of arbitrators with the requisite skills available from the developing world. Creating backstops and structural safeguards to build capacity and know-how aids the availability of a broader spectrum of arbitrators to resolve disputes. The arbitration community might take various steps in this respect.

International organizations like UNCTAD may continue and even expand their capacity-building efforts. UNCTAD has various training courses and online materials.¹⁹³ The London Court of International Arbitration's Young International Arbitrators Group ("YIAG") is making efforts to build capacity by providing training opportunities to those in the developing world.¹⁹⁴ Basic education that encourages understanding of the legal framework and practice of investment treaty arbitration provides the foundation for a new generation of sophisticated arbitration practitioners from the developing world.

Beyond the educational baseline, arbitrators need a degree of sophistication and expertise. Parties and institutions must have confidence in the credentials, background, and gravitas of an arbitrator before trusting him or her to adjudicate a dispute involving governments, issues relating to the exercise of sovereignty that affect the public, and substantial claims for damages. As the creation of such a reputation takes years, there are benefits to acting early in order to develop the requisite qualities.

Part of the solution involves providing structured opportunities for sophisticated education and professional development. This might mean, for example, that arbitrators could take on clerks (or secretaries, for tribunals) from the developing world. Having the opportunity to see how tribunals function and potentially be involved in writing complicated awards provides a fertile ground for intellectual and professional development of future arbi-

193. See, e.g., U.N. Conference on Trade and Dev. [UNCTAD], *Dispute Settlement: General Topics*, U.N. Doc UNCTAD/EDM/Misc.232/Add.30 (2003) (edited by Erik Chrispeels); U.N. Conference on Trade and Dev. [UNCTAD], Project on Dispute Settlement in International Trade, Investment and Intellectual Property, <http://www.unctad.org/Templates/WebFlyer.asp?intItemID=4403&lang=1> (last visited Mar. 12, 2009); James X. Zhan, Chief of Int'l Arrangements, UNCTAD, Opening Remarks, Symposium on Making the Most of International Investment Agreements: A Common Agenda (Dec. 12, 2005), available at <http://www.oecd.org/dataoecd/5/9/36053809.pdf>.

194. *Second YIAG Training Seminar—Vilnius*, YIAG E-LETTER (Young Int'l Arb. Group/London Ct. Int'l Arb., London, U.K.), Dec. 2007, at 2, http://www.lcia.org/YIAG_folder/documents/YIAGE-LetterDecember07FINAL.pdf.

trators. Similarly, it might mean that arbitral institutions, such as ICSID, the Stockholm Chamber of Commerce, or other local arbitral institutions, could make more concerted efforts to have staff members from developing countries who could later translate those professional experiences into serving as arbitrators. Even NGOs and UNCTAD could consider soliciting interns from the developing world in order to build the reputation and capacity of arbitrators from the developing world. One step in the right direction would be, for example, the creation of foreign investment moot court competitions and establishment of awards to identify legal talent from the developing world.¹⁹⁵

Beyond this, professional organizations have a role in helping potential arbitrators from the developing world raise their profiles. Groups such as the American Society for International Law have provided financial aid to certain academics and practitioners from the developing world to attend conferences with the Society. Likewise, the newly established Society for International Economic Law may wish to consider creating opportunities for networking and professional support for potential investment arbitrators in training. The creation of professional opportunities to speak, network, and meet with one's peers is critical. It offers the chance for intellectual exchange in order to develop a sophisticated and nuanced understanding of the subject matter. It also provides an occasion to begin building one's professional reputation and receiving invitations to speak at conferences. Further, it can create opportunities for discussion that might result in arbitrator appointments.

However, capacity and reputation building alone may be insufficient. Parties and arbitral institutions need to know of the existence of arbitrators before they can appoint them. Professor Catherine Rogers has conducted interviews with those connected to investment treaty arbitration and found that there appears to be a sense that developing nations are at a loss when selecting arbitrators:

Anna Joubin-Bret of the United Nations Conference on Trade and Development ("UNCTAD") reports that representatives from developing nations, who are often unfamiliar with the field and unable to obtain reliable information about arbitrator candidates, have been known to resort to relatively random selection criteria, such as nominating an academic they happened to encounter at a conference on investment arbitration. Other experts in the field confirm that, without the aid and guidance of one of the leading

195. For example, the Frankfurt Investment Arbitration Moot Court Competition gives an award for the "Best Team from a Non-OECD Member Country." Frankfurt Investment Arbitration Moot Court, Outline of Rules: Rule 3.4, available at http://www.merton-zentrum.uni-frankfurt.de/Startseite/FIAC_International_Student_Moot_Court/Rules/index.html (last visited Mar. 15, 2009). One challenge, however, is making sure that teams from the developing world have sufficient funding to enable them to participate in the competitions.

investment arbitration law firms, developing countries flounder in their efforts to obtain reliable information about arbitrator candidates.¹⁹⁶

This has a variety of implications. First and foremost, it suggests that it may be useful to create a publicly available database of arbitrators.¹⁹⁷ Such a database might offer a list of those party-appointed and presiding arbitrators who have served on cases in the past and various identifying information such as nationality; gender; age; educational background; languages spoken; occupational background, such as work with NGOs, IOs, government entities, and practice in the private sector; and the number of cases in which an arbitrator has been involved as counsel, expert, or arbitrator. It might also include information about the treaty cases in which the arbitrators were involved. Likewise, it could offer basic information about those who wish to be arbitrators in the future. Such a publicly available database, perhaps maintained by ICSID or UNCTAD, may eliminate or minimize problems of information access.

Overall, taking steps to expand the pool of potential arbitrators and reducing the pipeline problem might eliminate the simple effects observed in the current research without changing the legal framework of arbitration. Areas worthy of exploration include building capacity, developing opportunities for professional advancement, and providing for public dissemination of information about arbitrators. These actions could provide a possible remedy for the issues raised by the current research. Even if the results are not replicable, expanding the pool of qualified arbitrators enhances party choice related to arbitrator selection in order to promote party control and arbitrators' perceived legitimacy.

2. *Redesigning the Normative Structure of Investment Treaty Dispute Resolution*

It is important to consider normative changes that could offer targeted solutions for specific problems. This section explores those possibilities. First, it suggests that the results warrant systematic consideration of additional dispute resolution options. Second, it discusses the creation of targeted procedural safeguards. Third, it suggests that policymakers should consider textual revisions to model investment treaties and existing agreements.

196. Rogers, *Have-Nots*, *supra* note 31, at 358; *see also* Kahale, *supra* note 88, at 5 (noting that "states . . . often lack the experience to make an appropriate selection for [the] party-appointed arbitrator"). *But see* Poppi Hagan & Zachary Lomo, *International Law and the Developing World: A Millennial Analysis*, 41 HARV. INT'L L.J. 595, 606 (2000) (noting that "bias in the arbitration system results from the common training, intellectual background, and shared principles of the arbitrators—most notably, a shared idea of to what extent the public sphere can impede on the private").

197. Franck, *Evaluating Claims*, *supra* note 16, at 51; Catherine A. Rogers, *The Vocation of the International Arbitrator*, 20 AM U. INT'L L. REV. 957, 1009–10 (2005).

a. Considering Additional Dispute Resolution Mechanisms

The initial results provide a basis for the systematic evaluation of the existing dispute resolution mechanisms and consideration of the value of dispute resolution processes other than arbitration. Given concerns related to disparate awards granted by tribunals with certain kinds of arbitrators, there is utility in exploring the merits of other dispute resolution mechanisms that (1) take arbitrators out of the picture and provide opportunities for interest-based, creative solutions that meet the mutual needs of investors and states, (2) may not disparately favor the developed world over developing countries, and (3) may not require the same cash outlays or other resource allocation, which deny access to justice for investors with limited means or countries that need to focus on their core obligations of providing services to their citizens.¹⁹⁸

There is a financial cost to arbitration beyond the potential social and political costs. Irrespective of a host state's development classification, arbitrators must be paid, and lawyers generally require compensation for their legal services. In one case that is arguably a statistical outlier, *EnCana v. Ecuador*,¹⁹⁹ a winning state had to pay for the legal services of the losing party.²⁰⁰ This means that even if a state has not violated an investment treaty, it may nevertheless incur financial expenses connected with the arbitration beyond its own fiscal expenses. Given the potential costs of arbitration, it would be worthwhile to assess the value of alternative forms of dispute resolution. In other words, from a business perspective, the initial results help create the case for reconsidering the value of different types of dispute resolution mechanisms to permit parties to gain the most value and minimize their costs when resolving investment treaty conflicts. Some research has already begun to assess this area,²⁰¹ and institutional stakeholders have begun to explore the benefits of supporting dispute resolution processes other than arbitration.²⁰²

198. See, e.g., Franck, *Dispute System Design*, *supra* note 26 (outlining the utility of dispute resolution mechanisms such as interest-based negotiation, mediation, and ombuds, and related opportunities for creative problem solving).

199. *EnCana Corp. (Canada) v. Republic of Ecuador*, LCIA Case No. UN3481 (Feb. 3, 2006), available at <http://ita.law.uvic.ca/documents/EncanaAwardEnglish.pdf>.

200. Franck, *Evaluating Claims*, *supra* note 16, at 67–70.

201. See *supra* notes 182–84 and accompanying text; see also Susan D. Franck, *Challenges Facing Investment Disputes: Reconsidering Dispute Resolution Options in International Investment Agreements*, in *APPEALS MECHANISM IN INTERNATIONAL INVESTMENT DISPUTES* 143 (Karl P. Sauvant ed., 2008); Ucheora Onwuamaegbu, *The Role of ADR in Investor-State Dispute Settlement: The ICSID Experience*, NEWS FROM ICSID (ICSID (W. Bank), Washington, D.C.), Winter 2005, at 12.

202. See, e.g., Conference Schedule, U.N. Conference on Trade and Dev. and Ministry of Justice of the Ukraine, Conference on Inv. Treaties and Alternative Methods of Inv. State Dispute Resol., June 2–3, 2008, available at http://www.unctad.org/sections/dite_pcbb/docs/dite_pcbb_ias0062_en.pdf; see also U.N. Conference on Trade and Dev. [UNCTAD], *The Development Dimension of International Investment Agreements*, U.N. Doc. TD/B/CJI/MEM.3/2 (Dec. 2, 2008) (observing that ADR is “hardly ever referred to in IIAs, although they are available under international instruments” and suggesting that it “would

R

R

R

Should the results not prove replicable, reconsidering dispute resolution options in response to a reasoned critique enhances the procedural integrity of arbitration. To the extent that the simple effects reflect perceived rather than real concerns, it creates an opportunity to put more dynamic dispute resolution systems in place that minimize the likelihood that problems might arise later. It also suggests a capacity to consider and adapt to the concerns of stakeholders. Such responsiveness enhances the legitimacy of investment treaty dispute resolution.

b. The Role of Structural Safeguards

Beyond considering alternative dispute resolution processes, there are other ways to improve investment treaty arbitration. This section focuses on creating targeted procedural safeguards and argues for the creation of organizations like a legal assistance center or an appellate body.

As regards a legal assistance center, commentators have raised concerns about the “capacity of the poorest developing countries to defend” investment treaty claims, particularly given the process’s costs and uncertainties.²⁰³ Others suggest that there are concerns about how developing nations select their arbitrators.²⁰⁴ A legal assistance center for developing countries could provide strategic advice to enhance the quality of arbitration and eliminate disparities in outcome related to development status.

This research, particularly if replicated, creates a strong case for the creation of an appellate body or even a stand-alone investment court. If outcome is linked to the development status of the presiding arbitrator and there is disparate pressure to favor the developed world, having standing judges with secure tenures may enhance integrity and independence. In order to eliminate pressure to join a club or secure repeat appointments, a standing body could provide judicial oversight and create an environment that favors rule of law adjudication. Moreover, such an institution could foster the judicialization of international economic law and provide a backstop to create certainty about contested legal issues, thereby increasing the integrity of the dispute resolution system.²⁰⁵

c. The Value of “Legislative” Reforms

Beyond structural safeguards, there is also value in considering “legislative” approaches to redesigning dispute resolution. In particular, treaty

be worthwhile to consider giving a more prominent role to ADR—such as mediation and conciliation—in future IIAs”).

203. Peterson, *UK BIT Program and Sustainable Development*, *supra* note 52, at 10; *see also* Eric Gottwald, *Leveling the Playing Field: Is It Time for a Legal Assistance Center for Developing Nations Facing Investment Treaty Arbitration?*, 22 AM. U. INT’L L. REV. 237 (2007) (discussing challenges faced by less developed countries who have had to defend themselves in investment treaty and/or ICSID cases and the need for a legal assistance center and capacity building).

204. *See* Rogers, *supra* note 31, at 358.

205. *See generally* Franck, *Legitimacy Crisis*, *supra* note 188 (discussing the possible creation of an appellate body).

R

R

R

negotiators may wish to consider targeted fixes of the text of investment treaties. While the overall results suggest that radical textual reform may not be appropriate, these targeted fixes may prove useful. Given the simple effects, negotiators should determine whether they are involved in an arbitration involving one of the two country-pairs at issue (that is, High versus Upper-Middle income countries and High versus Low Income countries). If so, governments might consider creating backstops to minimize the risk of disparate outcomes.

First, they might consider providing more particularized guidelines about arbitrator appointment. This may involve giving institutions direct authority to appoint presiding arbitrators, appointing off an existing list of arbitrators, or employing a pre-approved party choice. Governments might also consider creating a narrow list of characteristics that presiding arbitrators must have, such as experience in both the governmental and commercial sectors. The objectives of these two reforms might be to select the highest quality arbitrator, gain the broadest possible diversity, or minimize the opportunity for disparate treatment by creating market forces that are more akin to those at work in the selection of neutral adjudicators.

Second, to redress the perception that outcomes are biased according to development-related variables, treaty negotiators might consider providing particularized interpretation rules. While the rules of the Vienna Convention on the Interpretation of Treaties typically apply, more specific guidance may help orient expectations about what an investment treaty will do. It may also provide guidance to arbitrators about how they should exercise their discretion and when they must cede to the parties' intentions concerning treaty interpretation. Regarding damages, for example, one might imagine governments setting guidelines in treaties for when to use discounted cash flow or another damage methodology. For expropriation damages, a treaty might also identify the appropriate date of expropriation from which to measure damages, which could be either the date of expropriation or the date of the award. In the costs context, likewise, governments could create rules about who will bear the costs of the arbitrators' fees and the possibility for shifting costs. Provision of greater interpretative precision has the potential benefit of minimizing confusion, setting expectations, and offering guidance to arbitrators about how they must or may exercise their authority. The hope would be that a more specific interpretative mandate promotes integrity in the dispute resolution process. Instead of permitting an association between outcome and development-related variables, countries can recommit themselves to deciding disputes according to neutral, mutually agreed legal principles and demonstrating their commitment to rule of law adjudication.

C. Cautionary Tales for Counsel and Arbitrators

The results should serve as a wake up call for lawyers and arbitrators. Both groups may wish to consider this information to see how it might impact their future behavior. While each treaty and dispute is unique, stakeholders may nevertheless use the data to frame their understanding of the conflict.

This analysis has certain implications for counsel involved in the strategic appointment of arbitrators. Lawyers representing High Income states may want to take steps to appoint presiding arbitrators from Middle Income countries to increase the statistical probability of having to pay the lowest award possible. It may likewise mean that governments from developing countries may do better with presiding arbitrators from High Income countries. But parties should act on these strategic decisions with great caution given previously described limitations.

There are also implications for arbitrators. First, this research underscores the need to take care not to resolve disputes on the basis of the respondent state's development status. Professor Schneiderman recounts the experience of a party-appointed arbitrator in the *Loewen v. United States* NAFTA dispute. Although it does not pertain to presiding arbitrators and it is unclear how representative the example is, Judge Mikva reportedly relayed this story:

After agreeing to serve on the Loewen tribunal by the US Department of Justice (DOJ), Mikva, a retired DC circuit court judge, met with DOJ officials prior to the panel being constituted. "You know, judge," he was told by DOJ, "if we lose this case we could lose NAFTA." "Well, if you want to put pressure on me," Mikva replied, "then that does it."²⁰⁶

Arbitrators should not succumb to pressure related to their national, developmental, or other affinities that may inappropriately affect their adjudicatory mandates. Rather, they should focus on the merits of claims.

Ideally, empirical results from this study would add another layer of consciousness and encourage a renewed focus on the merits. In the long run, it is in arbitrators' rational self-interest to take these initial findings seriously and engage in behavior that enhances the actual and perceived integrity of arbitration. Recognizing the need for particular care may create incentives for behavior that enhance arbitrators' personal credibility and make arbitration viable and useful for stakeholders. There are benefits to providing a depoliticized, rule-based system that people can rely upon, predict, under-

206. David Schneiderman, *Judicial Politics and International Investment Arbitration: Seeking an Explanation for Conflicting Outcomes* (unpublished manuscript, on file with the author) (citing Audio tape: Abner Mikva, *Judiciary and Environmental Law Symposium*, held by Pace University School of Law (Dec. 6–8, 2004)).

stand, and trust. Such a system allows participants to prevent conflict from festering and proceedings from becoming controlled only by those with political and economic power. The objective should be fair, efficient, and effective rule of law adjudication rather than pure exercise of power or use of war or military means to resolve conflict. Since the early 1900s, judicialized international economic adjudication has grown. There is social utility in providing access to justice. There is further utility in administering the system in a way that is perceived to be fair and that operates on the basis of agreed principles and a shared social contract, rather than on unprincipled exercises of political will or physical power.

It is also valuable when presiding arbitrators work to promote adjudicative neutrality within their own analyses and encourage such behavior within tribunals. In both the arbitral process and the rendering of awards, it is critical to administer justice and ensure that stakeholders perceive this administration. Beyond the potential to eliminate or minimize the two simple effects observed in the initial research, promotion of adjudicative neutrality may enhance the perceived public integrity of the arbitration system. Such initiatives are welcome advances in the creation of legitimate mechanisms for the resolution of international economic disputes.

VII. Conclusion

The statistical analyses consistently showed that, at a general level, the outcome of investment treaty arbitration was not reliably associated with the development status of the respondent state, the development status of the presiding arbitrator, or some interaction between those two variables.

The notion that outcome is not associated with arbitrator or respondent development status should be a basis for cautious optimism. It provides evidence about the integrity of arbitration and casts doubt on the assumption that arbitrators from developed states show a bias in terms of arbitration outcomes or that the development status of respondent states affect such outcomes. It suggests that major structural overhaul may not be necessary because it is not clear that arbitration is inherently predisposed towards particular outcomes.

The lack of a reliable relationship between development status and outcome suggests that other variables or combinations of variables may drive arbitration results. Some of these variables may be completely disassociated from the arbitration process. Possible variables could include those traditionally associated with neutral, adjudicative forums, whether courts, claims, commissions, or arbitrations, such as the quality of expert evidence, the nature and scope of legal representation, and submissions by *amicus curiae*. Other variables affecting results may, however, be intrinsically tied to arbitration, such as the qualities and experiences of arbitrators. Future research might usefully assess the impact of these and other variables in order to gain

a more nuanced understanding of factors that are reliably associated with outcome. This could inform decisions about creating processes for managing investment treaty-related conflict.

While the general initial results are encouraging, one should contextualize them properly, given their limitations. The presence of the two statistically significant simple effects also suggests that there are areas ripe for targeted reform. For example, future research may indicate that presiding arbitrators from the developing world disproportionately grant larger awards against the developing world and smaller awards against the developed world. If this is so, either a targeted response to particular investment treaties or reform of the arbitrator appointment process may be useful. Even if the results are not replicable, there is nevertheless value in adapting the system to address perceived concerns and take steps that may prevent conflict from escalating. In any event, the critical message from the initial results is clear, namely that more empirical research is needed to examine development issues in greater detail and consider how best to enhance the integrity of the dispute resolution process.

As Freud wrote in *Civilization and Its Discontents*, “The first requisite of civilization, therefore, is that of justice . . . [and t]he final outcome should be a rule of law to which all . . . have contributed . . . and which leaves no one . . . at the mercy of brute force.”²⁰⁷ To the extent that governments continue to negotiate investment treaties that implicate both substantive and procedural rights, they are, in Freud’s words, contributing to the rule of law and a community standard of acceptable behavior. One might also hope for fair and non-arbitrary application of this community-created standard in investment treaty dispute resolution. At a minimum, one might hope for a system that, in practice, leaves neither claimants nor respondents at the mercy of brute force. After all, the judicialization of investment dispute was designed to avoid such use of force and to leave gunboat diplomacy to the annals of history.

This Article suggests that while there may be some problems with arbitration, it is not clear that a development divide affects outcomes. However, this is subject to evolution based upon new research. While we should monitor whether presiding arbitrators from Middle Income countries unduly favor particular respondents, we must acknowledge that the research may develop over time.

Ultimately, we must be vigilant, particularly in light of the modern political realities of investment treaty dispute resolution. If the cautionary tales from the two simple effects or effect sizes bear fruit, showing that development status is somehow connected reliably with arbitration outcome, then there will be implications for the entire arbitration system. It will then be

207. SIGMUND FREUD, *CIVILIZATION AND ITS DISCONTENTS* 81–82 (James Strachey ed., W.W. Norton & Co. 2005).

imperative to consider the relevant causal mechanisms in order to permit stakeholders to evaluate the fundamental fairness and legitimacy of the current arbitration process.

We have much to gain from shining an empirical and systematic light on investment treaty dispute resolution. We can assess conventional wisdom and articulate where it diverges from reality. We can ascertain the areas that require further research when considering articulated concerns about arbitration. Perhaps most importantly, rather than permitting immersion in political rhetoric, we can promote an engaged, informed dialogue in order to create systems that maximize mutual interests and promote justice. In an area with profound political, economic, and social implications, such dialogue is a critical component of the long-term well-being of a fair and sustainable global political economy.

