
Managing Backlash: The Evolving Investment Treaty Arbitrator?

Malcolm Langford* and Daniel Behn**

Abstract

Have investment treaty arbitrators responded to the so-called 'legitimacy crisis' that has beleaguered the international investment regime in the past decade? There are strong rational choice and discursive-based reasons for thinking that these adjudicators would be responsive to the prevailing 'stakeholder mood'. However, a competing set of legalistic and attitudinal factors may prevent arbitrators from bending towards the arc of enhanced sociological legitimation. This article draws upon a newly created investment treaty arbitration database to analyse the extent and causes of a shift in treaty-based arbitration outcomes. The evidence is suggestive that arbitrators are conditionally reflexive – sensitive to both negative and positive signals from states, especially influential, developed and vocal states.

1 Introduction

The development of the modern investment treaty regime (regime) represents one of the most remarkable extensions of international law in the post-World War II period. Largely built on a network of more than 3,000 signed bilateral investment treaties (BITs), regional free trade agreements (FTAs)¹ and a handful of plurilateral investment

* Professor, Faculty of Law, University of Oslo; Co-Director, Centre on Law and Social Transformation, University of Bergen and Chr. Michelsen Institute, Norway. Email: malcolm.langford@jus.uio.no. The authors would like to thank the reviewers and Daniel Naurin, Anne Van Aaken, Andreas Føllesdal, Geir Ulfstein, Alec Stone Sweet, Gus van Harten and Stavros Brekoulakis for comments on an earlier version and participants at various workshops for comments, particularly Jose Alvarez, Joseph Weiler, Benedict Kingsbury, David Caron, Larry Helfer, Laurence Boisson de Chazournes, Ole Kristian Fauchald and Christina Voigt. The original version of this article won the Young Scholar Prize at the 2015 European Society of International Law conference.

** Lecturer in Law, Liverpool Law School, University of Liverpool, United Kingdom; Senior Research Fellow, PluriCourts Centre of Excellence, Faculty of Law, University of Oslo, Norway. Email: d.f.behn@jus.uio.no.

¹ The United Nations Conference on Trade and Development provides an extensive database on international investment agreements (IIAs), available at <http://investmentpolicyhub.unctad.org/IIA>.

treaties,² foreign investors are granted beneficiary rights primarily aimed at the protection of their investments. While each international investment agreement (IIA) is a stand-alone agreement with considerable diversity, agreements typically contain similar standards of protection,³ including, most importantly, investor–state dispute settlement (ISDS) provisions. Combined, it has been claimed that ‘no other category of private individuals’ is ‘given such expansive rights in international law as are private actors investing across borders’.⁴

For a number of reasons, the development of this regime has precipitated a backlash from some states, various civil society actors⁵ and scholars.⁶ Commonly referred to as a ‘legitimacy crisis’,⁷ even some prominent insiders and expected supporters in the media have expressed disquiet.⁸ Primarily, this phenomenon is not about the expansiveness of the substantive rights granted to foreign investors under IIAs but, rather, the combination of such rights with the robustness of the ISDS mechanisms embedded within them. The result has been an explosion of litigation. With over 878 known investment treaty arbitrations (ITAs) initiated through 1 August 2017 (almost all coming in the last 15 years)⁹ as well as a significant number of instances in which the threat of treaty arbitration has been used as a bargaining tool, states hosting foreign investors are increasingly finding themselves having to defend their laws and policies before and in the shadow of international arbitral tribunals. Many of the concerns about the regime are tied specifically to outcomes of this litigation; they include claims

² See eg Energy Charter Treaty 1994, 2080 UNTS 95; North American Free Trade Agreement (NAFTA) 1992, 32 ILM 289, 605 (1993); Association of South-East Asian Nations Comprehensive Investment Agreement, Central American-Dominican Republic Free Trade Agreement (DR-CAFTA) 2004, 43 ILM 514 (2004) (draft) as well as recently concluded or late-round negotiated treaties: Trans-Pacific Partnership Agreement between the Government of Australia and the Governments of Chile, Japan, Malaysia, Mexico, New Zealand, Peru, Singapore, United States of America and Vietnam (TPP) 2016, [2006] ATNIF; Regional Comprehensive Economic Partnership 2017 (currently under negotiation).

³ IIAs typically include: prohibitions against expropriation without adequate compensation, full protection and security, fair and equitable treatment, most-favoured nation treatment and national treatment.

⁴ Simmons, ‘Bargaining over BITS, Arbitrating Awards: The Regime for Protection and Promotion of International Investment’, 66 *World Politics* (2014) 12, at 42.

⁵ See, e.g., P. Eberhardt and C. Olivet, ‘Profiting from Injustice: How Law Firms, Arbitrators, and Financiers Are Fuelling an Investment Arbitration Boom’, *Corporate Europe Observatory* (November 2012).

⁶ Langford, Behn and Fauchald, ‘Backlash and State Strategies in International Investment Law’, in T. Gammeltoft-Hansen and T. Aalberts (eds), *The Changing Practices of International Law: Sovereignty, Law and Politics in a Globalising World* (2018) 70.

⁷ For an overview, see Behn, ‘Legitimacy, Evolution, and Growth in Investment Treaty Arbitration: Empirically Evaluating the State-of-the-Art’, 46(2) *Georgetown Journal of International Law* (2015) 363; Behn, ‘Performance of Investment Treaty Arbitration’, in T. Squatrito *et al.* (eds), *The Performance of International Courts and Tribunals* (2018) 77.

⁸ See, e.g., Douglas, ‘The MFN Clause in Investment Arbitration: Treaty Interpretation off the Rails’, 2 *Journal of International Dispute Settlement (JIDS)* (2011) 97; Kahale, ‘Is Investor-State Arbitration Broken?’, 7 *Transnational Dispute Management* (2012), available at www.transnational-dispute-management.com; ‘The Arbitration Game: Governments Are Souring on Treaties to Protect Foreign Investors’, *Economist* (11 October 2014).

⁹ PluriCourts Investment Treaty Arbitration Database (PITAD), through 1 August 2017, available at www.jus.uio.no/pluricourts/english/topics/investment/research-projects/database.html.

that ITA is pro-investor or anti-developing state; that the jurisprudence is incoherent and riddled with contested interpretations and that the levels of monetary compensation are too high.¹⁰

While critical perspectives grab the headlines, the regime also has its supporters. Some claim that individual arbitral decisions are not as expansive or pro-investor as imagined;¹¹ that arbitral tribunals provide a relatively predictable legal framework;¹² that investment treaty arbitrators are not insensitive to other branches of international law¹³ and that ITA assists in ‘depoliticizing’ international disputes by providing a significant alternative to the pre-war era of gunboat diplomacy.¹⁴ Thus, any efforts to ‘re-statify’ (or re-balance) international investment dispute resolution should be resisted.¹⁵

Between these critics and supporters, one finds an alternative evolutionary position that holds that the legitimacy crisis in investment treaty arbitration is merely a *crise de croissance* – ‘growing pains’ – and that, as the system matures, it will evolve and adapt into a more legitimate, consistent and effective form of international adjudication.¹⁶ A central part of this expected evolution will come from the arbitrators themselves. Indeed, certain empirical theories suggest that adjudicators are responsive to material and symbolic signals from other actors.¹⁷ However, the existing empirical scholarship on the legitimacy crisis in ITA has focused primarily on proving the existence or non-existence of bias.¹⁸ Only a nascent and doctrinal literature has examined whether arbitrators might be leading a shift in response to the legitimacy crisis.¹⁹

¹⁰ See, e.g., Van Harten, ‘Arbitrator Behaviour in Asymmetrical Adjudication: An Empirical Study of Investment Treaty Arbitration’, 50 *Osgoode Hall Law Journal* (2012) 211, at 251; Langford, ‘Cosmopolitan Competition: The Case of International Investment’, in C. Bailliet and K. Aas (eds), *Cosmopolitanism Justice and Its Discontents* (2011) 178.

¹¹ Collier, ‘Book Review: *Constitutionalizing Economic Globalization: Investment Rules and Democracy’s Promise* by D. Schneiderman’, 68 *Cambridge Law Journal* (2009) 231.

¹² Fauchald, ‘The Legal Reasoning of ICSID Tribunals: An Empirical Analysis’, 19 *European Journal of International Law (EJIL)* (2008) 301.

¹³ Fry, ‘International Human Rights Law in Investment Arbitration: Evidence of International Law’s Unity’, 18 *Duke Journal of International and Comparative Law* (2007) 77.

¹⁴ Behn, ‘Legitimacy’, *supra* note 7; Behn, ‘Performance’, *supra* note 7.

¹⁵ Brower and Blanchard, ‘From “Dealing in Virtue” to “Profiting from Injustice:” The Case against “Re-Statification” of Investment Dispute Settlement’, 55(1) *Harvard International Law Journal Online* (2014) 45.

¹⁶ Bjorklund, ‘Report of the Rapporteur Second Columbia International Investment Conference: What’s Next in International Investment Law and Policy?’, in J. Alvarez *et al.* (eds), *The Evolving International Investment Regime: Expectations, Realities, Options* (2011) 219.

¹⁷ See section 3 below.

¹⁸ Van Harten, *supra* note 10; Franck, ‘Conflating Politics and Development: Examining Investment Treaty Outcomes’, 55 *Virginia Journal of International Law* (2014) 1; Schultz and Dupont, ‘Investment Arbitration: Promoting the Rule of Law or Over-Empowering Investors? A Quantitative Empirical Study’, 25 *EJIL* (2014) 1147; J. Donaubaauer, E. Neumayer and P. Nunnenkamp, ‘Winning or Losing in Investor-to-State Dispute Resolution: The Role of Arbitrator Bias and Experience’, Kiel Working Paper no. 2074 (2017).

¹⁹ Schneidermann, ‘Legitimacy and Reflexivity in International Investment Arbitration: A New Self-Restraint’, 2 *JIDS* (2011) 471.

This backdrop provides an opportunity to explore how international adjudicators (in this case, investment treaty arbitrators) respond to systemic critiques and how such responses might be measured. In this article, we ask whether there is (or has been) a reflexive and evolutionary self-correction by arbitrators. Do arbitrators seek to build both the normative and sociological legitimacy of the regime by adopting more state-friendly approaches to the resolution of substantive and procedural questions? Or do we find that the behaviour of arbitrators is largely indifferent to the storm outside? We begin the article by mapping the trajectory of the legitimacy crisis and shift(s) in stakeholder mood (section 2) and then theorize as to how, why and when investment treaty arbitrators might be sensitive to the legitimacy crisis (section 3). After briefly examining the doctrinal literature, we use our PluriCourts Investment Treaty Arbitration Database (PITAD) to examine whether there has been an aggregate shift in the outcomes of ITAs and whether it can be explained by arbitrator behaviour (section 4).

2 Charting the Legitimacy Crisis

In order to understand the potential reaction of arbitrators to the legitimacy crisis in investment treaty arbitration, we need to chart the crisis's trajectory. We are particularly interested in who communicates 'signals' of crisis, what sort of signals might be communicated to arbitrators and when. As to who, states are of particular interest; they constitute the primary principals in the international investment treaty regime and may be thus particularly influential in affecting arbitrator behaviour, intentionally or otherwise.²⁰ As to what, state signals might include: exit actions such as the denouncement of the ICSID Convention²¹ and the termination of IIAs; voice actions such as the adoption of more sovereignty-sensitive model IIAs or mixed actions such as moratoriums on the signing of new IIAs, demands for renegotiations of IIAs or increasingly aggressive litigation tactics in defending ITA claims. However, states are not alone. Other stakeholders may also signal displeasure with the regime. Civil society actors might submit *amicus curiae* briefs and publicly mobilize against arbitration proceedings or new treaty negotiations; academics may criticize awards or issue collective statements and national court judges may decline to enforce awards at the domestic level. These actors often mobilize in the context of large-scale or controversial 'public interest cases' cases, whereby foreign investors' claims are pitted against a state's general regulatory measures designed to protect the environment or public health. Identifying these diverse and overlapping signals is challenging, but we sketch a brief history that is divided into three broad periods.

A Pre-Crisis (1990–2001) and Building Crisis (2002–2004)

Initially, investment treaty arbitration was an obscure and largely unknown specialization that attracted little attention. While the first modern BIT was signed in 1959,²² it was

²⁰ Langford, Behn and Fauchald, *supra* note 6.

²¹ Convention on the Settlement of Investment Disputes between States and Nationals of Other States (ICSID Convention) 1965, 575 UNTS 159.

²² Agreement between the Islamic Republic of Pakistan and the Federal Republic of Germany (1959).

not until 1968 that the first BIT providing for ISDS was signed,²³ and it took a further 19 years until the first treaty-based arbitration was submitted.²⁴ After the first award in 1990,²⁵ there was only a slight trickle of cases throughout the 1990s, which we can describe as pre-crisis, and the field was largely overshadowed by contract-based investment and commercial arbitrations.

In the early 2000s, the first high-profile ITA cases occurred. These were raised under the North American Free Trade Agreement (NAFTA) against developed states, the most prominent being the *Loewen* case.²⁶ Although dismissed on jurisdiction, this case revealed that the justice system of the USA had embarrassing shortcomings that might be challenged under international law and that arbitrators might face significant political pressure when tasked with resolving these types of disputes. Together with other NAFTA cases against the USA, Canada and Mexico, these early arbitrations also highlighted a perceived threat to sovereignty and the regulatory autonomy of states. Significantly, they catalysed the production of a corrective interpretive note by the NAFTA Free Trade Commission in 2001 (with a more minimalist approach to the fair and equitable treatment (FET) standard)²⁷ and a new US model BIT in 2004 (which was more deferential to state interests).²⁸

Beyond the NAFTA, several other cases in the early 2000s raised significant and specific concerns regarding the relationship between IIA standards and environmental or human rights-based policy measures.²⁹ This included the *Aguas del Tunari* case against Bolivia, which grew out of the infamous ‘water wars of Cochabamba’ and prompted a global civil society campaign against the arbitration, resulting in the first-ever submission of an *amicus curiae* brief in an ITA case.³⁰ This period also saw some controversial examples of inconsistent case law, particularly exemplified by the *SGS* cases – two different tribunals arrived at contradictory interpretations of umbrella clauses³¹ –

²³ Agreement between the Government of the Kingdom of the Netherlands and the Government of the Republic of Indonesia on Promotion and Protection of Investment (1968).

²⁴ The first case under the ICSID Convention was filed in 1972, but this was a claim for a contractual breach. ICSID, *Holiday Inn and Others v. Morocco*, ICSID Case no. ARB/72/1, settled.

²⁵ ICSID, *Asian Agricultural Products v. Sri Lanka – Final Award*, 27 June 1990, ICSID Case no. ARB/87/3.

²⁶ ICSID, *Loewen Group and Raymond Loewen v. United States – Final Award*, 26 June 2003, ICSID Case no. ARB(AF)/98/3. See also A. DePalma, ‘NAFTA’s Powerful Little Secret; Obscure Tribunals Settle Disputes, But Go Too Far, Critics Say’, *New York Times* (11 March 2001). NAFTA, *supra* note 2.

²⁷ NAFTA Free Trade Commission, Notes of Interpretation of Certain Chapter 11 Provisions, 20 July 2001. See Kaufmann-Kohler, ‘Interpretive Powers of the Free Trade Commission and the Rule of Law’, in F. Bachand (ed.), *Fifteen Years of NAFTA Chapter 11 Arbitration* (2011) 175, at 181–185.

²⁸ See US Department of State, Treaty between the Government of the United States of America and the Government of [Country] Concerning the Encouragement and Reciprocal Protection of Investment (US Model BIT) (2004), and particularly the qualifications of the expropriation standard.

²⁹ See Fauchald, ‘International Investment Law and Environmental Protection’, in O.K. Fauchald and D. Hunter (eds), *17 Yearbook of International Environmental Law* (2006) 3, at 11–25.

³⁰ ICSID, *Aguas del Tunari v. Bolivia*, ICSID Case no. ARB/02/3, settled.

³¹ ICSID, *SGS Société Générale de Surveillance S.A. v. Islamic Republic of Pakistan – Decision of the Tribunal on Objections to Jurisdiction*, 6 August 2003, ICSID Case no. ARB/01/13; ICSID, *SGS Société Générale de Surveillance S.A. v. Republic of the Philippines – Decision of the Tribunal on Objections to Jurisdiction*, 29 January 2004, ICSID Case no. ARB/02/6.

and the *Lauder* and *CME* cases – in which two tribunals issued two different awards on essentially the same subject matter.³² We can thus see in this period a ‘building crisis’, exemplified by the regime’s transition from obscurity to partial prominence, rumblings of discontent with select cases and an emerging critical scholarship and general commentary. Notably, this discontent included pushback by the USA, one of the regime’s most dominant norm-setters.³³

B *Legitimacy Crisis (2005–2010)*

In 2004 and 2005, the phrase ‘legitimacy crisis’ emerged in the academic scholarship for the first time, and the crisis discourse extended clearly beyond its NAFTA origins. The numerous ITA awards rendered as a result of the Argentinian economic crisis of 2001–2002 were important in this regard. They engendered a large volume of commentary, particularly on a state’s regulatory autonomy in times of national emergency.³⁴ The legitimacy crisis discourse was further fuelled by a large number of cases filed against Venezuela, Bolivia and Ecuador following the passage of various nationalization laws as well as the *Foresti* case against South Africa.³⁵ The result was not only expressions of displeasure but also high-profile announcements of exit strategies. Bolivia (2007), Venezuela (2009) and Ecuador (2012) denounced the ICSID Convention; Ecuador and Bolivia terminated many of its BITs and South Africa placed a moratorium on the signing of new IIAs pending an extensive policy review.³⁶

By the end of the first decade of the new millennium, the legitimacy crisis discourse and the practice of ITA began to reach maturity. One prominent example of such literature is the publication of the first book with the word ‘backlash’ in the title.³⁷ This period is also marked by significant signals from academia, primarily exemplified by the Public Statement on the International Investment Regime. It states in part:

We have a shared concern for the harm done to the public welfare by the international investment regime, as currently structured, especially its hampering of the ability of governments to act for their people in response to the concerns of human development and environmental sustainability ... There is a strong moral as well as policy case for governments to withdraw from investment treaties and to oppose investor-state arbitration, including by refusal to pay arbitration awards against them where an award for compensation has followed from a good faith measure that was introduced for a legitimate purpose.³⁸

³² UNCITRAL, *Ronald S. Lauder v. Czech Republic – Final Award*, 3 September 2001; UNCITRAL, *CME Czech Republic B.V. v. Czech Republic – Final Award*, 14 March 2003.

³³ See Alvarez, ‘The Return of the State’, 20 *Minnesota Journal of International Law* (2011) 223, at 235.

³⁴ See, e.g., Alvarez and Khamsi, ‘The Argentine Crisis and Foreign Investors: A Glimpse into the Heart of the Investment Regime’, in K. Sauvant (ed.), *Yearbook on International Investment Law and Policy 2008–2009* (2009) 379.

³⁵ ICSID, *Piero Foresti and Others v. South Africa*, ICSID Case no. ARB(AF)/07/01, discontinued.

³⁶ South Africa Department of Trade and Industry, *Bilateral Investment Treaty Policy Framework Review*, June 2009.

³⁷ M. Waibel et al. (eds), *The Backlash against Investment Arbitration: Perceptions and Reality* (2010).

³⁸ Public Statement on the International Investment Regime, 31 August 2010, preamble, available at www.osgoode.yorku.ca/public-statement/documents/Public%20Statement%20%28%20June%20201129.pdf. It was signed by 76 academics.

C Diverging Discourses: Late Crisis and Counter-Crisis (2011–2017)

In the last six to seven years, the narrative of crisis became entrenched among a broader set of stakeholders, but countervailing narratives also emerged. The period initially saw a dramatic increase in new publications mentioning the legitimacy crisis and globally prominent cases that further stoked the fires of controversy. The *Phillip Morris* tobacco regulation cases against Australia³⁹ and Uruguay,⁴⁰ the energy utility *Vattenfall* cases against Germany⁴¹ and Chevron's US \$18 billion denial of justice case against Ecuador are notable.⁴² This litigation in turn triggered new partial exit strategies. In the wake of the *Phillip Morris* litigation, the government under Julia Gillard in Australia announced that no future IIA with Australia would include ISDS provisions.⁴³ Similarly, after being subject to a wave of cases, the Czech Republic initiated an internal policy review, mutually terminated some IIAs and renegotiated many others.⁴⁴

In 2014 and 2015, the discourse on the legitimacy of investment treaty arbitration moved into the public sphere for the first time,⁴⁵ and a number of high profile awards were rendered against, *inter alia*, Venezuela,⁴⁶ Zimbabwe,⁴⁷ Canada⁴⁸ and Russia.⁴⁹ The number of new cases grew and stabilized at an average of approximately 50 cases per annum, including more than 45 claims (in 2013–2016) against European Union

³⁹ UNCITRAL, *Phillip Morris Asia v. Australia – Award on Jurisdiction and Admissibility*, 17 December 2015, PCA Case no. 2012-12.

⁴⁰ ICSID, *Phillip Morris Brands Sàrl, Phillip Morris Products S.A. and Abal Hermanos S.A. v. Uruguay – Award*, 8 July 2016, ICSID Case no. ARB/10/7.

⁴¹ ICSID, *Vattenfall AB, Vattenfall Europe AG, Vattenfall Europe Generation AG v. Germany*, ICSID Case no. ARB/09/6, settled; ICSID, *Vattenfall AB, Vattenfall Europe AG, Vattenfall Europe Generation AG v. Germany*, ICSID Case no. ARB/12/12, pending.

⁴² UNCITRAL, *Chevron Corporation and Texaco Petroleum Corporation v. Ecuador*, PCA Case no. 2009-23, pending.

⁴³ Australian Department of Foreign Affairs and Trade, Gillard Government Trade Policy Statement: Trading Our Way to More Jobs and Prosperity, November 2011.

⁴⁴ K. Gordon and J. Pohl, 'Investment Treaties over Time: Treaty Practice and Interpretation in a Changing World', OECD Working Papers on International Investment (2015).

⁴⁵ Partly as a result of the Transatlantic Trade and Investment Partnership (TTIP) (draft dated 12 November 2015). See, e.g., 'The Arbitration Game', *supra* note 8; 'Trade Agreement Troubles', *New Yorker* (22 June 2015); 'TTIP Will Not be Approved unless ISDS Is Dropped', *Financial Times* (27 October 2014).

⁴⁶ See, e.g., ICSID, *ConocoPhillips Petrozuata B.V., ConocoPhillips Hamaca B.V. and ConocoPhillips Gulf of Paria B.V. v. Venezuela – Decision on Jurisdiction and Merits*, 3 September 2013, ICSID Case no. ARB/07/30; ICSID, *Gold Reserve v. Venezuela – Award*, 22 September 2014, ICSID Case no. ARB(AF)/09/1; ICSID, *Venezuela Holdings, B.V. et al. v. Venezuela – Award of the Tribunal*, 9 October 2014, ICSID Case no. ARB/07.

⁴⁷ ICSID, *Bernhard von Pezold and Others v. Zimbabwe – Award*, 28 July 2015, ICSID Case no. ARB/10/15; ICSID, *Border Timbers Limited, Border Timbers International Limited, and Hangani Development Co. Limited v. Zimbabwe – Award*, 28 July 2015, ICSID Case no. ARB/10/25.

⁴⁸ UNCITRAL, *William Ralph Clayton, William Richard Clayton, Douglas Clayton, Daniel Clayton and Bilcon of Delaware v. Canada – Award on Jurisdiction and Liability*, 17 March 2015, PCA Case no. 2009-04.

⁴⁹ The three landmark cases collectively granting US \$50 billion to Yukos shareholders: UNCITRAL, *Yukos Universal v. Russian Federation – Final Award*, 18 July 2014, PCA Case no. AA 227; UNCITRAL, *Hulley Enterprises v. Russian Federation – Final Award*, 18 July 2014, PCA Case no. AA 226; UNCITRAL, *Veteran Petroleum v. Russian Federation – Final Award*, 18 July 2014, PCA Case no. AA 228.

(EU) member states in relation to subsidization schemes for the promotion of solar energy.⁵⁰ Certain states continued to terminate and/or renegotiate their IIAs as a response to defending against treaty-based arbitration, including the Czech Republic, Romania, Indonesia, India and Poland.⁵¹

However, the last five years also have produced contradictory shifts in sovereign state policy, reflecting a possible countervailing mood or tendency. Negotiations on new regional mega-agreements including ISDS provisions burst into life; the USA and the other NAFTA states formally joined (and largely took over) the negotiations for the Trans-Pacific Partnership (TPP).⁵² Moreover, negotiations for the Regional Comprehensive Economic Partnership (RCEP) among almost all south and east Asian states were launched, and efforts to develop a Transatlantic Trade and Investment Partnership (TTIP) between the EU and the USA were invigorated after the release of a high-level expert report on 11 February 2013.⁵³ In terms of bilateral treaties, the EU emerged as an IIA negotiator with third states following the entry into force of the Lisbon Treaty⁵⁴ and has sought to negotiate and sign new FTAs (including with Brazil, Canada, China, India, Indonesia, Japan, Mexico, Singapore, and Vietnam). China continued to renegotiate IIAs with stronger protections for foreign investors (as have Germany and the Netherlands), the Norwegian government (long absent in new investment treaty developments) released a new model BIT in 2015, Brazil started signing new IIAs (with Angola, Chile, Colombia, Malawi, Mexico and Mozambique) after famously refusing to ratify any of their previously signed agreements from the 1990s⁵⁵ and Australia reversed its anti-ISDS policy and signed the TPP in February 2016. To be sure, this mood change should not be over-emphasized. The EU has sought to promote more balanced investment agreements, and, in January 2017, the incoming US president Donald Trump abruptly cancelled participation in the TPP and the TTIP. Nonetheless, the most recent period is certainly marked by more contradictory tendencies than the former.

3 Theorizing Arbitrator Reflexivity

We now turn from this depiction of the legitimacy crisis to interrogate its effects and, specifically, whether it has impacted arbitrator behaviour in investment treaty

⁵⁰ See Behn and Fauchald, 'Governments under Cross-Fire: Renewable Energy and International Economic Tribunals', 12(2) *Manchester Journal of International Economic Law* (2015) 117.

⁵¹ See overview in Langford, Behn and Fauchald, *supra* note 6; T. Jones, 'Poland Threatens to Cancel BITs', *Global Arbitration Review* (26 February 2016).

⁵² TPP, *supra* note 2.

⁵³ Mention of the TTIP, *supra* note 45, was included in the US president's state of the union address the next day, and an announcement of new talks by the European Commission president came the day after that.

⁵⁴ The Lisbon Treaty conferred competence to the European Union (EU) in the area of foreign direct investment for the first time: Treaty of Lisbon Amending the Treaty on European Union and the Treaty Establishing the European Community 2007, OJ 2007 C 306.

⁵⁵ P. Martini, 'Brazil's New Investment Treaties: Outside Looking ... Out?', *Kluwer Arbitration Blog* (15 June 2015).

disputes. In other words, do arbitrators seek to manage, consciously or unconsciously, the legitimacy of arbitration? The techniques available for managing legitimacy are common to all international courts and arbitral bodies and perhaps even more so to ITA where a formal doctrine of precedent is absent.⁵⁶ Such methods might include tightening jurisdictional criteria, exhibiting greater deference to respondent states on the merits, reducing the number of claims upon which a claimant/investor wins or the amount of compensation awarded, shifting legal costs on to claimant/investors more frequently or a combination of all of these. However, we first need to establish theoretically why arbitrators would turn to such techniques. A useful way of distinguishing two competing sets of hypotheses is to employ the analytical framework or heuristic of delegation theory. The extent to which investment treaty arbitrators are reflexive arguably comes down to the extent to which they act as ‘trustees’ or ‘agents’.⁵⁷

A ‘Trustee’ Null Hypothesis: Status Quo

Some scholars argue that adjudicators on international courts and arbitral bodies can be characterized as ‘trustees’. They adjudicate through delegated authority and according to their own professional judgments on behalf of states and other beneficiaries.⁵⁸ This conception suggests that an external legitimacy crisis would exert little influence on adjudicative decision-making. On its face, the argument that investment treaty arbitration lies at the trustee end of the delegation continuum is attractive. For a start, many of the typical characteristics of an agency relationship are not present as arbitrators wield significant discretionary powers with minimal accountability. Arbitral jurisdiction is made compulsory in most IIAs; arbitral appointments in a particular case are largely beyond challenge; awards can only be overridden on very narrow technical and procedural grounds;⁵⁹ it is difficult for states to amend treaty provisions in order to avoid any precedential effects that an award may have on future cases with a similarly placed investor⁶⁰ and there is no formal channel by which states

⁵⁶ See, e.g., Madsen, ‘The Legitimization Strategies of International Courts: The Case of the European Court of Human Rights’, in M. Bobek (ed.), *Selecting Europe’s Judges* (2015) 260; Weiler, ‘Journey to an Unknown Destination: A Retrospective and Prospective of the European Court of Justice in the Arena of Political Integration’, 31(4) *Journal of Common Market Studies* (1993) 417; S. Dothan, ‘Why Granting States a Margin of Appreciation Supports the Formation of a Genuine European Consensus’, iCourts Working Paper Series no. 22 (2015).

⁵⁷ See, e.g., Alter, ‘Agents or Trustees? International Courts in their Political Context’, 14(1) *European Journal of International Relations* (2008) 33. Delegation is a fairly open framework such that principal-agent approaches are more a ‘highly flexible family of models, rather than an overarching set of assumptions and results’. Gailmard, ‘Accountability and Principal-Agent Models’, in M. Bovens, T. Schillemans and R.E. Goodin (eds), *Oxford Handbook of Public Accountability* (2014) 90. In this respect, see the response of M. Pollack, ‘Principal-Agent Analysis and International Delegation: Red Herrings, Theoretical Clarifications, and Empirical Disputes’, Bruges Political Research Papers no. 2 (2007) to claims that the fiduciary model of beneficiary-trustee is an alternative to the principal-agent model.

⁵⁸ Alter, *supra* note 57.

⁵⁹ There are very limited grounds for appeal – either through annulment procedures (under the ICSID Convention) or domestic court set-aside proceedings (non-ICSID cases).

⁶⁰ Gordon and Pohl, *supra* note 44.

can express their discontent with arbitral awards rendered against them. As in trusteeship theory, arbitrators are also usually selected on the basis of their ‘personal and professional reputation’,⁶¹ and international adjudicators might be relatively insulated from the signals of states as they ‘serve publics with diverse and often conflicting preferences’.⁶² Thus, the expectation that investment treaty arbitrators will act strategically and reflexively overrates their ability or need to appreciate the existence, extent and nature of any legitimacy crisis.

If the trustee model captures the space in which investment treaty arbitrators operate, we would therefore expect the underlying values of the regime and/or arbitrators to largely guide decision-making. As Karen Alter puts it, the result will be a ‘rhetorical politics’ in which the appointing actors will appeal to the trustee’s ‘mandate’ and ‘philosophies’.⁶³ A likely candidate for such underlying values would be legal positivism. We might therefore expect that arbitrators, according to their professional judgment, would seek to apply IIA provisions in good faith to the specific facts of the case. Indeed, the fact that arbitrators regularly find for respondent states as much as claimant/investors may suggest a certain even-handedness.⁶⁴ Accordingly, any change in arbitral behaviour could only be explained by legal shifts in the regime’s substantive rules or a significant shift in the average set of factual circumstances. Yet, it is hard to say that there has been a major change in the former. Some recent and revised IIAs include general, but vague, clauses concerning the right to regulate or greater exceptions for domestic environmental and labour policies, but their significance is not yet clear.⁶⁵ An empirical study based on computational text analysis suggests that renegotiations of IIAs tend to result in less room for state regulatory powers and more investor-protective ISDS provisions.⁶⁶ Therefore, we might expect that most changes would occur due to a change in the nature of case facts.

Paradoxically, an attitudinalist perspective of adjudicative behaviour would suggest the same static and trustee-based hypothesis. Here, adjudicators make decisions according to their sincere ideological attitudes and values (according to their ‘personal judgment’)⁶⁷ because they are relatively unconstrained by other actors, including states.⁶⁸ As investment treaty arbitrators represent an elitist and largely Western-based epistemic community, the commitment to promoting and protecting foreign investment may be particularly strong.

⁶¹ Y. Dezalay and B. Garth, *Dealing in Virtue: International Commercial Arbitration and the Construction of a Transnational Legal Order* (1998).

⁶² Lupu, ‘International Judicial Legitimacy: Lessons from National Courts’, 14(2) *Theoretical Inquiries in Law* (2013) 437, at 438. Some scholars claim that the problem even extends to domestic courts. National judges will only have ‘vague notions’, for example, about parliamentary preferences and the risk of legislative override. See Segal, ‘Separation-of-Powers Games in the Positive Theory of Congress and Courts’, 91 *American Political Science Review* (1997) 1, at 31.

⁶³ Alter, *supra* note 57.

⁶⁴ Behn, ‘Legitimacy’, *supra* note 7; Behn, ‘Performance’, *supra* note 7; Franck, *supra* note 18.

⁶⁵ Broude *et al.*, ‘Who Cares about Regulatory Space in BITs? A Comparative International Approach’, in A. Roberts *et al.* (eds), *Comparative International Law* (2017) 527.

⁶⁶ Alschner, ‘The Impact of Investment Arbitration on Investment Treaty Design: Myth versus Reality’, 42(1) *Yale Journal of International Law* (2017) 1.

⁶⁷ See generally J. Segal and H. Spaeth, *The Supreme Court and the Attitudinal Model* (1993).

⁶⁸ Segal, *supra* note 62, at 28.

Arbitrators from Western Europe and North America make up a total of 70 per cent of all appointees to investment treaty arbitrations⁶⁹ and 84 per cent of the top 25 arbitrators that have dominated a large percentage of all arbitral appointments.⁷⁰ Such differences can matter. In the context of the International Court of Justice, Eric Posner and Miguel Figueredo report that permanent judges are significantly more likely to vote for a disputing state that shares a similar level of economic development and democracy with their home state.⁷¹ In the context of ITA, there has been a slight uptick in the appointment of arbitrators hailing from lesser developed states, but many of them tend to come from a similar 'epistemic community',⁷² and some suggest that aspiring arbitrators need to adhere to the 'rules of the club' in order to gain appointments.⁷³

B 'Agent' Reflexivity Hypothesis: The Evolving Arbitrator

The alternative to these predictions of stable arbitral behaviour is to suggest that investment treaty arbitrators do follow the mood shifts of states and other actors – as agents rather than trustees. The principal prism through which to understand and model such behaviour is rational choice. Adjudicators: (i) may hold diverse preferences that extend beyond political ideology or good lawyering; (ii) 'take into account the preferences and likely actions of other relevant actors, including their colleagues, elected officials, and the public' and (iii) operate in a 'complex institutional environment' that structures this interaction.⁷⁴ The idea that adjudicators are sensitive to these actors is not hard to find. Evidence from various domestic jurisdictions suggests that judges are strategically sensitive to signals from the executive and legislature,⁷⁵ although the scholarship is divided on the extent of this shift.⁷⁶ As to public opinion, there is consensus that it has an indirect influence on judgments though judicial appointments but is divided over whether it exerts a direct influence on judges.⁷⁷ At

⁶⁹ PITAD through 1 August 2017.

⁷⁰ Langford, Behn and Lie, 'The Revolving Door in International Investment Arbitration', 20(2) *Journal of International Economic Law (JIEL)* (2017) 301. See also Puig, 'Social Capital in the Arbitration Market', 25 *European Journal of International Law* (2014) 387.

⁷¹ Posner and de Figueiredo, 'Is the International Court of Justice Biased?', 34 *Legal Studies* (2005) 599. Together, these correlations explained a remarkable 60% to 70% of variance among individual judicial votes.

⁷² Langford, Behn and Lie, *supra* note 70.

⁷³ Dezalay and Garth, *supra* note 61.

⁷⁴ Epstein and Knight, 'Reconsidering Judicial Preferences', 16 *Annual Review of Political Science* (2013) 11, at 11. On diverse goals, see, in particular, L. Baum, *Judges and Their Audiences: A Perspective on Judicial Behavior* (2008). On strategies, see Alex Schwartz, *An Agent-Based Model of Judicial Review*, ICON-S Annual Conference, Hong Kong, 25-27 June 2018.

⁷⁵ See, e.g., Rodriguez-Rada, 'Strategic Deference in the Colombian Constitutional Court, 1992–2006', in G. Helmke and J. Rios-Figueroa (eds), *Courts in Latin America* (2011) 81; Epstein and Knight, *supra* note 74; Kapiszewski, 'Tactical Balancing: High Court Decision Making on Politically Crucial Cases', 45 *Law and Society Review* (2011) 471.

⁷⁶ Cf. Bergara, Richman and Spiller, 'Modeling Supreme Court Strategic Decision Making: The Congressional Constraint', 28(2) *Legislative Studies Quarterly* (2003) 247; Segal, *supra* note 62.

⁷⁷ Flemming and Wood, 'The Public and the Supreme Court: Individual Justice Responsiveness to American Policy Moods', 41 *American Journal of Political Science* (1997) 468, at 480. See also B. Friedman, *The Will of the People: How Public Opinion Has Influenced the Supreme Court and Shaped the Meaning of the Constitution* (2009); Epstein and Martin, 'Does Public Opinion Influence the Supreme Court? Possibly Yes (But We're Not Sure Why)', 13 *University of Pennsylvania Journal of Constitutional Law* (2010) 263, at 270; Unah et al., 'U.S. Supreme Court Justices and Public Mood', 30 *Journal of Law and Politics* (2015) 293.

the international level, empirical and doctrinal scholarship suggests that the Court of Justice of the European Union⁷⁸ and the World Trade Organization (WTO) dispute settlement body⁷⁹ are sensitive to the balance and composition of member state opinion within institutional constraints.

Turning to investment treaty arbitrators, a strategic account would imply that a behavioural correction in response to legitimacy critiques could forestall certain material and reputational ‘costs’. In practice, arbitrators may lack trusteeship freedoms and are reduced to agents engaged in ‘contractual politics’ with their principals. There might be three grounds for thinking so. First, arbitrators may be concerned collectively about backlash by principals as it may increase the risk of non-compliance by respondent states, encourage greater exits from the regime, reduce the rate of new treaties being entered into or, for those with pro-investor proclivities, result in weaker future and/or revised IIAs. Such state behaviour would inhibit the ability of arbitrators to impose their political preferences (comparable to the concern with ‘overrides’ in the judicial context)⁸⁰ and maintain their general reputational standing. Second, investment treaty arbitrators may be concerned about their own individual reputation and material chances of future appointment.⁸¹ If they experience a reversal through annulment procedures,⁸² set-asides in domestic courts or criticism by their colleagues or scholars, their behaviour may adjust. Arbitrators interested in the role of the chair or the respondent wing arbitrator may be particularly sensitive, given the common role of states in these appointments.

Could arbitrators be so strategic and consequentialist? Well, arbitrators themselves have acknowledged that the notion is not far-fetched. In a recent survey, 262 international arbitrators, which included a subset of 67 with experience in ITA,⁸³ were asked whether they considered future re-appointment when deciding cases.⁸⁴ Remarkably, 42 per cent agreed or were ambivalent. Given the sensitive nature of the question, it is arguable that this figure is understated.⁸⁵ These strategic predictions may be also enhanced by sociological forces.⁸⁶ According to the theory of discursive

⁷⁸ Larsson and Naurin, ‘Judicial Independence and Political Uncertainty: How the Risk of Override Affects the Court of Justice of the EU’, 70(2) *International Organization* (2016) 377; M. Pollack, *The Engines of European Integration: Delegation, Agency, and Agenda Setting in the EU* (2003).

⁷⁹ C. Creamer, ‘Between the Letter of the Law and the Demands of Politics: The Judicial Balancing of Trade Authority within the WTO’, Working Paper (2015).

⁸⁰ Larsson and Naurin, *supra* note 78.

⁸¹ Studies of domestic judges that are subject to reappointment processes reveal higher levels of strategic behavior among this group. See I. Lifshitz and S.A. Lindquist, ‘The Judicial Behavior of State Supreme Court Judges’, APSA 2011 Annual Meeting Paper (2011).

⁸² Van Aaken, ‘Control Mechanisms in International Investment Law’, in Z. Douglas *et al.* (eds), *The Foundations of International Investment Law: Bringing Theory into Practice* (2014) 409.

⁸³ S. Franck *et al.*, ‘International Arbitration: Demographics, Precision and Justice’, ICCA Congress Series no. 18 – Legitimacy: Myths, Realities, Challenges (2015) 33.

⁸⁴ *Ibid.*, at 91.

⁸⁵ *Ibid.*

⁸⁶ On this empirical conundrum, see Gilles, ‘Reputational Concerns and the Emergence of Oil Sector Transparency as an International Norm’, 54 *International Studies Quarterly* (2010) 103.

institutionalism, discourse (such as the legitimacy crisis) is not simply a static, internalistic and slow-moving phenomenon but also an independent, dynamic and liminal phenomenon. Shifts to stakeholder discourse may shape the 'background ideational abilities' of judicial agents.⁸⁷ Or as Benjamin Cardozo puts it, 'the great tides and currents which engulf the rest of men do not turn aside in their course and pass the judges by'.⁸⁸ Arbitrators may shift their background preferences as they become acquainted or engaged in the legitimacy debate. The crisis may also affect their 'foreground discursive abilities' and the space in which they 'communicate critically about those institutions, to change (or maintain) them'.⁸⁹ Arbitrators may simply adapt to a different palette of legitimate reasons that can be foregrounded in their decision-making. Thus, changes in arbitrator behaviour may not only be strategic. It may also be a process of rapid adjustment to a new social norm that affects arbitrator preferences and speech acts.

These specific predictions of dynamic arbitral behaviour suggest a number of positive hypotheses of reflexivity, which can be divided into three categories: (i) 'stakeholders'; (ii) 'case-based' and (iii) 'role-based.' As set out below, the stakeholder hypotheses foreground the mood of different external actors (states and others) in shaping the calculus of all arbitrators; the case-based hypotheses suggest that the features of the case at hand determine the degree of arbitral reflexivity; while the institutional, role-based hypothesis suggests that external influence is only mediated through arbitrators in particular roles. Each is explained in turn.

1 Stakeholder Hypotheses

In light of domestic research, we might expect only states to exert any really influence on arbitrator behaviour. This is because it is only states that can impose material costs. As principals (treaty parties), states are essential for the institutional survival or development of the regime, and, as litigants (respondents), states participate in the appointment of arbitrators⁹⁰ and decide on whether to comply with an adverse arbitral award. We therefore suggest Hypothesis 1 on state signals: arbitrators will respond strongly to the preferences of states compared to other actors.

Alternatively, we can hypothesize that investment treaty arbitrators would respond to the general stakeholder mood.⁹¹ The multiplicity of 'micro-publics' (investors, counsel, arbitral institutions, academics, civil society and national judges) can all

⁸⁷ Schmidt, 'Discursive Institutionalism: The Explanatory Power of Ideas and Discourse', 11 *Annual Review of Political Science* (2008) 303, at 304.

⁸⁸ B. Cardozo, *The Nature of the Judicial Process* (1921), at 168.

⁸⁹ Schmidt, *supra* note 87, at 304.

⁹⁰ In most cases, states solely appoint one of three arbitrators and, in many cases, jointly appoint, along with the claimant/investor, the presiding arbitrator.

⁹¹ Yates and colleagues find that these judges are sensitive partly to shifts in changes in opinion in their home state, parties to which they are ideologically aligned, and a longer period of residence in the more liberal Washington, DC. See "'For the Times They Are A-Changin':" Explaining Voting Patterns of U.S. Supreme Court Justices through Identification of Micro-Publics', 28 *Brigham Young University Journal of Public Law* (2013) 117.

affect regime reputation and discourse, and arbitrators may be conscious that these stakeholders can also affect state positions in the medium to long run. We therefore suggest Hypothesis 2 on non-state signals: arbitrators will respond strongly to the general stakeholder mood.

More narrowly, we might expect that investment treaty arbitrators are particularly responsive to the views of certain audiences – namely, large, powerful or particularly influential states.⁹² Displaying such sensitivity may be strategic for reputational reasons, and it possibly enhances the prospect of more arbitrations entering into the pipeline (particularly due to the large capital exports of these states' foreign investors). Moreover, the views of these states are more likely to be publicized in various communication channels.⁹³ Thus, we suggest Hypothesis 3 on influential state signals: arbitrators will respond strongly to the preferences of large, powerful or particularly influential states.

2 Case-Based Hypotheses

Investment treaty arbitrators may be only reflexive to certain types of cases. It is conceivable that deferentialism to states may be a particularly safe strategy in ITA cases that are more thematically controversial because they court public or scholarly scrutiny that resonate with the underlying values of regulatory autonomy expressed in the legitimacy critiques. Many of these 'public interest cases' also raise questions of coherence with other branches of international law (for example, international environmental and human rights law) in a way that shapes the surrounding discourse and the relevant institutional-legal environment. Thus, we suggest Hypothesis 4 on public interest cases: arbitrators will act more deferentially towards states in high-profile cases that court public or scholarly scrutiny.

Greater arbitral deference may be shown to certain categories of respondent states that have experienced a stronger pattern of losing in ITA – that is, certain lesser-developed states. With Tarald Berge, we have demonstrated that claimant/investors are less likely to win the higher the respondent state's gross national income (GNI) per capita and largely regardless of the level of a state's democratic governance.⁹⁴ Given this phenomenon, we might expect that over time, investment treaty arbitrators would moderate asymmetries in outcomes between these different classifications of states. We thus suggest Hypothesis 5 on development status asymmetries: arbitrators will narrow asymmetries in arbitral outcomes between states that are more developed and those that are lesser developed.

⁹² By large and powerful, we specifically include influential states actors participating actively in the regime: the USA, the EU (including its member states) and China. We note that Larsson and Naurin, *supra* note 78, found that influential states had a greater influence on the Court of Justice of the European Union, although they theorized that this occurred through greater voting weights in potential overrides of judgments in the Council of Ministers.

⁹³ However, the actions of small Latin American states in partially exiting the international investment regime have also been communicated widely.

⁹⁴ Behn, Berge and Langford, 'Poor States or Poor Governance: Explaining Outcomes in Investment Treaty Arbitration', 38(2) *Northwestern Journal of International Law and Business* (2018).

3 Role-Based Hypotheses

Unlike other international courts, investment treaty arbitrators are quite limited in their ability to communicate and act in a collective fashion. The polycentric and ad hoc nature of ITA may prevent arbitrators from acting in a systemic manner, even if they wish to do so. Unlike a centralized court, an individual arbitral tribunal may feel it can make little contribution to signalling a systemic shift – it is one of many. The incentive to take extra *inter partes* action is thus minimal.⁹⁵ Moreover, arbitrators may be doctrinally constrained in considering general concerns, and one line of investment treaty jurisprudence suggests that individual arbitrators should not systemically reflect and act as they are constituted as specialist adjudicative bodies.⁹⁶

However, it is not clear that this constraint applies equally to all investment treaty arbitrators. Repeat arbitrators (especially repeat tribunal chairs) are likely to be more sensitive to systemic threats and opportunities in comparison to one-shot arbitrators. They might constitute ‘the guardians of the regime’, engaged in wider discussions over investment treaty law practice, development and legitimacy. Since tribunal chairs exert tremendous influence over the arbitral process, we propose Hypothesis 6 on prominent arbitrators: repeat tribunal chairs will respond to signals from states and/or other stakeholders.

4 Measuring Arbitrator Reflexivity

How can we determine if investment treaty arbitrators adjust their behaviour in response to the legitimacy crisis, and without asking arbitrators themselves to disclose their approaches? The first approach is doctrinal. Recent jurisprudential scholarship in investment treaty arbitration suggests a potential reflex on a number of critical areas, whether it is ITA cases involving an environmental component⁹⁷ or how investment treaty arbitral tribunals analyse particular IIA standards such as the criteria for a breach of the (indirect) expropriation standard,⁹⁸ the FET standard,⁹⁹ the full protection and security standard,¹⁰⁰ most-favoured nation provisions¹⁰¹ or the jurisdictional

⁹⁵ On this challenge at the domestic level in civil law courts, see Young and Lemaitre, ‘The Comparative Fortunes of the Right to Health: Two Tales of Justiciability in Colombia and South Africa’, 26 *Harvard Human Rights Journal* (2013) 179.

⁹⁶ M. Reisman, ‘Case Specific Mandates versus Systemic Implications: How Should Investment Tribunals Decide? The Freshfields Arbitration Lecture’, 29 *Arbitration International* (2013) 131.

⁹⁷ Behn and Langford, ‘Trumping the Environment? An Empirical Perspective on the Legitimacy of Investment Treaty Arbitration’, 18 *Journal of World Investment and Trade* (2017) 14; Viñuales, ‘Foreign Investment and the Environment in International Law: The Current State of Play’, in K. Miles (ed.), *Research Handbook on Environment and Investment Law* (forthcoming).

⁹⁸ Henckels, ‘Indirect Expropriation and the Right to Regulate: Revisiting Proportionality Analysis and the Standard of Review in Investor-State Arbitration’, 15(1) *JIEL* (2012) 223.

⁹⁹ Dolzer, ‘Fair and Equitable Treatment: Today’s Contours’, 12 *Santa Clara Journal of International Law* (2014) 7.

¹⁰⁰ Alexandrov, ‘The Evolution of the Full Protection and Security Standard’, in M. Kinnear *et al.* (eds), *Building International Investment Law: The First 50 Years of ICSID* (2015) 319.

¹⁰¹ J. Maupin, ‘MFN-Based Jurisdiction in Investor-State Arbitration: Is There Any Hope for a Consistent Approach?’, 14(1) *JIEL* (2011) 157.

requirements relating to the definition of a ‘foreign investor’.¹⁰² While this research often points to an evolution of the jurisprudence – such as a move towards proportionality analysis in indirect expropriation cases¹⁰³ (which recognizes more clearly a state’s regulatory autonomy) – the development is partial. Some arbitral tribunals criticize or ignore these doctrinal advances.¹⁰⁴

The advantage of such a doctrinal approach is that it provides a fine-grained perspective on the legal mechanics of change and permits a swift focus on those areas that have attracted the most criticism. It is also a field that can be developed, for example, through longitudinal doctrinal studies of repeat arbitrator decisions on the same topic. However, the disadvantage of a doctrinal lens is that one may be tracking unwittingly a subterfuge of verbiage; arbitrators may simply craft and tweak their foregrounded discourse without visiting any material consequences upon the actual decision-making. Tracking the ongoing interaction between doctrine and factual and political contexts requires also a broader aggregative perspective. Our approach is therefore outcome based. It is decidedly more quantitative in orientation and provides an analysis of patterns in arbitral tribunal decision-making over time. Its prime advantage is its focus on the concrete nature of decisions and remedies, which cannot be obscured by written reasoning or oral speech.

A Raw Data

Using a range of output variables, we first ask whether outcomes of ITA cases change across time. The measured outcomes are win/loss ratios for finally resolved cases, jurisdictional decisions and liability/merits decisions, together with compensation ratios. The data is obtained from a new and first-of-its-kind database (PITAD) that codes all ITA cases since their inception. We include only cases whose legal claim is treaty based, meaning that claims based exclusively on a contract or a host state’s foreign investment law are excluded. Discontinued or settled cases are also omitted. With these conditions in place, and, as of 1 August 2017, the dataset includes 389 finally resolved cases, where the claimant/investor wins on the merits or loses on jurisdiction or the merits. Cases can also be sliced another way, and we can separate 748 discrete decisions, which are made up of 453 jurisdiction decisions¹⁰⁵ and 291 liability/merits decisions.¹⁰⁶

Both types of decision categorization is useful in analysing reflexivity. ‘Finally resolved’ cases may capture diachronic strategic planning across a case, whereby

¹⁰² Van Harten, *supra* note 10, at 251.

¹⁰³ See, e.g., ICSID, *Técnicas Medioambientales Tecmed, S.A. v. Mexico – Award*, 29 May 2003, ICSID Case no. ARB (AF)/00/2, para. 122.

¹⁰⁴ Waincymer, ‘Balancing Property Rights and Human Rights in Expropriation’, in P.-M. Dupuy *et al.* (eds), *Human Rights in International Investment Law and Arbitration* (2009) 275; Henckels, *supra* note 98, at 237.

¹⁰⁵ The jurisdiction decisions include bifurcated and non-bifurcated cases. For a non-bifurcated case, a decision where the claimant/investor ultimately loses on the merits will be coded as two decisions: one jurisdiction decision counted as a win for the claimant/investor and one merits decisions counted as a loss.

¹⁰⁶ These liability/merits decisions do not count quantum awards. In other words, a liability award in favour of a claimant/investor is counted in the same way as a merits award where damages are included.

arbitrators may allow a claimant/investor to win at the jurisdiction stage but not the liability/merits stage. 'Discrete decisions' may better capture synchronic signals from actors at a particular point in time. However, one issue in coding outcomes in ITA is the measurement of how and to what degree a claimant/investor can be said to win at the jurisdiction or liability/merits stage of the dispute. Our database provides some nuance and makes a distinction between full wins and partial wins. This results in two different outcome indicators. The first indicator is 'any win' (at least a partial win) and the second is 'full win' (only full wins counted).¹⁰⁷ In this article, we conduct analysis for both outcome indicators, although only results for the former are reported here (with full win results reported online).¹⁰⁸ The any win indicator is the most reliable measure as distinguishing partial wins from full wins is not an exact science.¹⁰⁹ It provides also a strong analytical measure; failing to award anything to a claimant/investor represents a strong signal to both stakeholders and states about the posture of an arbitral tribunal. Nonetheless, the complementary use of the full win indicator may help us discern reflexivity. A move towards partial wins – so-called 'splitting the baby' – may demonstrate an attempt at greater even-handedness by arbitrators.

Figure 1 shows the any win success ratios across time for the claimant/investor at the jurisdiction stage and the liability/merits stage of the ITA dispute. It also tracks the any win success ratios for finally resolved cases. Eye-balling the trends, it is relatively clear that claimant/investors did well in the first decade of litigation. In the period 1990 to 2001, investors rarely lost at the jurisdiction stage (94 per cent success rate in 32 decisions), and they won in approximately 72 per cent of finally resolved cases (25 cases) and 78 per cent of liability/merits awards (23 decisions). From 2002, an observable drop in claimant/investor success occurs in finally resolved cases and liability/merits awards. The trend downwards appears to begin in 2002 and bottoms out a few years later. From 2002 to 1 August 2017, the success rates in finally resolved cases was 44 per cent for claimant/investor. For liability/merits awards, the trends are slightly different. Overall, the success rates drop to 59 per cent for this period (2002–1 August 2017), but there is a drift upwards in claimant/investor liability/merits awards successes from 2012 followed by a downward correction from about 2015.

Jurisdictional decisions reveal a partially inverse pattern. There is a shift downwards to an average of 82 per cent success for investors in the period from 2002 to 2010 (from 94 per cent in the period from 1990 to 2001), but a further drop downwards to about 69 per cent from 2011 onwards. These divergent patterns in recent years help explain why the success ratio for claimant/investors in finally resolved cases remains fairly steady at about 44 per cent throughout the period from 2002 to 1 August

¹⁰⁷ For any win (full and partials wins are coded as (1) and losses as (0)) and for a full win (full win coded as (1) and partial wins and losses as (0)).

¹⁰⁸ See PITAD, *supra* note 9.

¹⁰⁹ At the liability/merits stage, a full and partial win are not categorized according to the ratio of amount claimed and awarded or the number of successful claims. Rather, the distinction between a full win and a partial win is based on whether the claimant/investor – in a holistic assessment of the case – was made whole by the arbitral tribunal. At the jurisdiction stage, a full win is scored when no jurisdictional objections are sustained and a partial win is scored where the jurisdiction of the tribunal is restricted in scope.

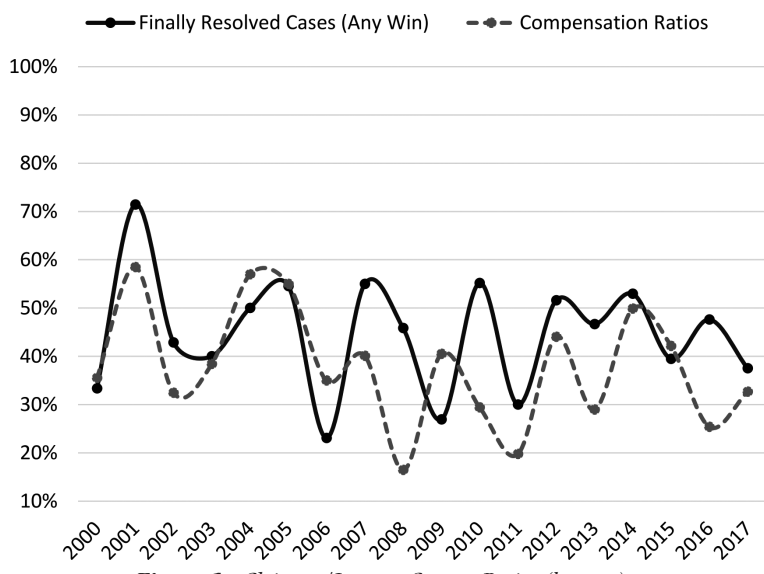
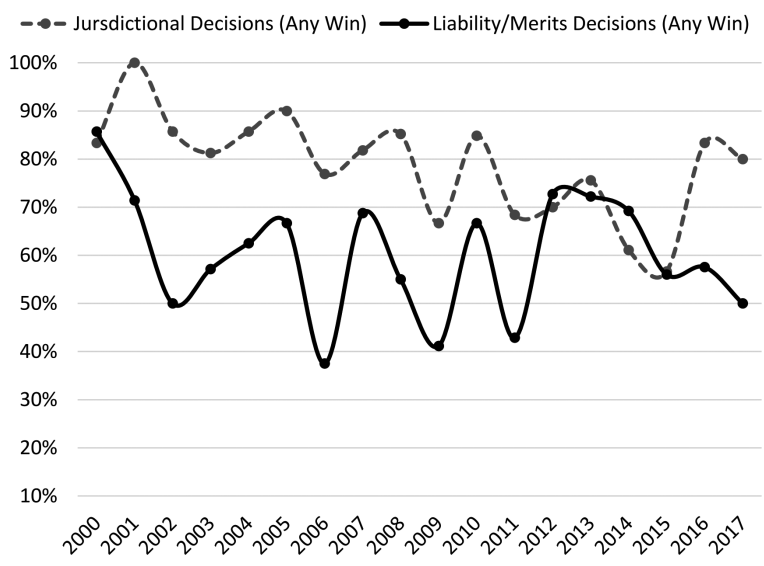


Figure 1: Claimant/Investor Success Ratios (by year)

2017. In other words, claimant/investors are currently more likely to be stopped at the jurisdictional stage, but, if they move through, they are more likely to succeed at the liability/merits stage. However, only so much can be read into this raw data as the outcomes are not controlled for structural features (an issue we address below).

In addition, we created a compensation ratio in cases in which the claimant/investor won on the merits and was awarded compensation. The ratio is the amount awarded divided by the amount claimed. However, it could only be calculated for a subset of 148 cases (out of 178 cases where the investor won on the merits) since information

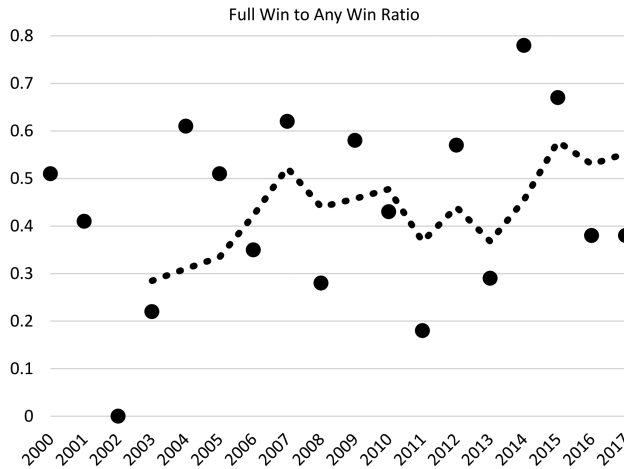


Figure 2: *Splitting the Baby? Full Win to Any Win Ratio in Finally Resolved Cases*

on both the amount of compensation claimed and awarded was not always known. The ratio has a large amount of annual variation but a surprising amount of stability over time. Between 1990 and 2004, the ratio was 44 per cent; which fell to 36 per cent for the period from 2005 to 2010 and stayed at 36 per cent between 2011 and 1 August 2017. The overall rate across all periods is 39 per cent.

The above figures only relate to any wins up to 1 August 2017. Yet, as noted earlier, strategic arbitrator behaviour may involve switching from full wins to partial wins. Figure 2 shows the ratio of the full win indicator to the total number of cases where the claimant/investor was successful at least partially. As is apparent from the four-year moving average (dotted line in Figure 2), there is a decline in the ratio in the early 2000s (after a period in which claimant/investors were almost always completely successful) with a slight drift upwards since 2012. On its face, the data are suggestive of a shift towards splitting the baby.

We now turn to ask whether these downward shifts in outcomes in ITA can be explained by arbitrator reflexivity, in accordance with the six hypotheses.

B Operationalization and Results

1 General and State Mood

In seeking to test the reflexivity expectations, we have operationalized the first two stakeholder hypotheses (state signals, non-state actor signals) into three different models. Each model tests the effects of a mood indicator with a lag of one year. First, Hypothesis 1 on state signals is operationalized by two separate indicators that we have constructed to measure state mood in relation to the international investment regime. The State Mood 1 indicator for treaty exits records a unilateral withdrawal by one state party to an IIA, including the ICSID Convention. As Figure 3 shows, this phenomenon begins in 2007, peaks in 2008 (with 19 treaty exits) and has remained at a steady annual average of about six treaty exits. An alternative version of this indicator weights the three ICSID Convention withdrawals by Latin American states by a factor of 10 on the basis that they received tremendous media and academic coverage.

The State Mood 2 indicators operationalize a positive state signal and records the number of new treaties (IIAs) signed by year. Importantly, this indicator is weighted for

remaining available treaties that could be signed. However, the effects of this weighting are not hugely significant as the number of possible IIAs that could be signed is still colossal.¹¹⁰ As Figure 4 shows, the number trends steadily downwards throughout the 2000s.

Hypothesis 2 on non-state signals is tested with an indicator that measures the annual number of references to the legitimacy crisis in investment treaty arbitration in the scholarship. It records a Google Scholar search of the legitimacy crisis discourse in academic publications.¹¹¹ As Figure 5 shows, this discourse commenced

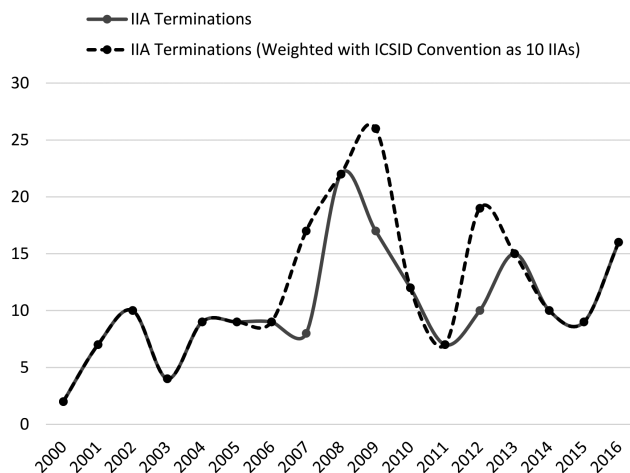


Figure 3: State Mood 1: Number of Unilateral Treaty Exits (by year)

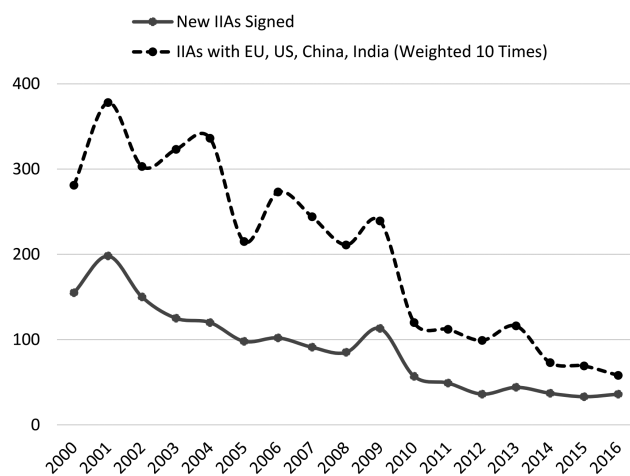


Figure 4: State Mood 2: Number of New Treaties Signed (by year)

¹¹⁰ There are currently 3,329 IIAs (mostly bilateral) that have been signed globally (through 1 August 2017). However, to receive the same coverage as the World Trade Organization (WTO) agreements, for example, states would need to sign the equivalent of 13,041 bilateral investment treaties.

¹¹¹ We began by entering the Boolean search terms of 'legitimacy crisis', 'investment treaty', and 'arbitration' in with a custom range for each year we identified. The 823 publications that appeared in the search through the end of 2016 were studied and further publications identified.

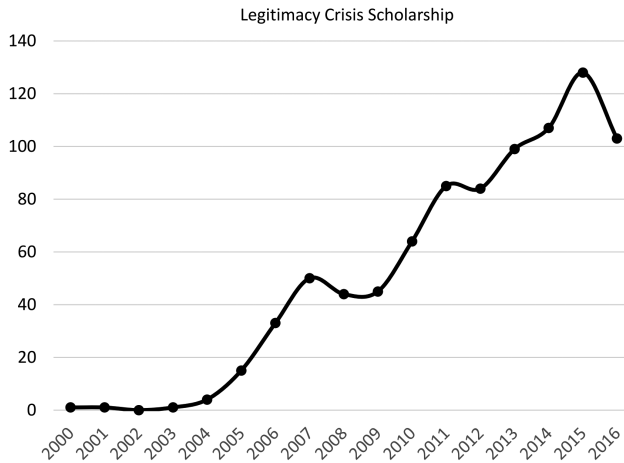


Figure 5: Stakeholder Mood: Legitimacy Crisis Scholarship (by year)

in 2004, spiked in 2009 and 2010 and has continued steadily upwards ever since. Obviously, this heavily weights academic scholarship but, given that academics tend to research broader trends, it may be representative of a broader non-state discourse.

In order to avoid potentially misleading bivariate results for the correlation between these three indicators and investment treaty arbitration outcomes, we include a set of controls for each model. The basic attributes are summarized in Table 1 alongside the independent variables. First, we include a dummy variable for treaty-based arbitration type, specifically NAFTA-based cases and cases administered by the International Centre for Settlement of Investment Disputes (ICSID).¹¹² Second, we apply an extractive industry cases dummy that measures whether the investment leading to a claim is in the extractive industries economic sector. These cases often involve varying degrees of nationalization with the dispute centring on levels of compensation, not liability (and, thus, claimant/investors will be more likely to win). Third, we add a measure of law firm advantage to control for the effect of the quality (or at least the expense) of legal counsel as measured by whether claimant/investors and respondent states retained counsel from a Global 100 law firm.¹¹³ Fourth, we include a dummy variable for state learning to control for the effect of previous exposure to investment treaty arbitration.¹¹⁴ Fifth, to control for situations where specific events or circumstances create an artificially large caseload against a respondent

¹¹² We include this dummy because NAFTA-based arbitrations matured earlier, while ICSID-administered arbitrations are based on a specific treaty (the ICSID Convention) with some specific structural features. ICSID-administered cases constitute 59% (523 of 878 ITA cases) of all known treaty-based arbitrations registered through 1 August 2017.

¹¹³ See *American Lawyer*, available at www.law.com/americanlawyer/sites/americanlawyer/2017/09/25/the-2017-global-100/. The dummy takes the value of (1) if only the claimant/investor counsel is from a Global 100 law firm; (-1) if only the respondent state retains a Global 100 law firm or (0) if both the claimant/investor and the respondent state both have the same type of law firm representing them.

¹¹⁴ We assume the marginal effect of state learning to diminish over time and code how many cases any given respondent state has had filed against it at the time of case registration up until the tenth case.

Table 1: Summary Statistics of Fully Resolved Cases

	Mean	Standard deviation	Min	Max	Observations
Outcome variables					
Any win	0.45	0.50	0	1	389
Full win	0.22	0.42	0	1	389
Independent variables					
State Mood 1 (Treaty Exits)	13.01	6.01	2	26	389
State Mood 2 (New Treaties)	74.89	43.80	33	198	389
General mood (Google Scholar)	66.55	41.4	0	128	389
Public interest case (Section 5C)	0.22	0.41	0	1	389
Prominent arbitrator (Section 5C)	0.72	0.49	0	1	389
Controls					
NAFTA-based case	0.09	0.29	0	1	389
ICSID-administered case	0.62	0.48	0	1	389
Extractive industry case	0.17	0.38	0	1	389
Law firm advantage	-0.10	0.56	-1	1	389
State learning	5.84	4.09	1	10	389
Case cluster	0.11	0.31	0	1	389
GNI per capita (logged)	8.71	1.09	5.37	11.01	388

state in a short space of time, we use a case cluster dummy.¹¹⁵ Sixth, we include a GNI per capita (logged) for respondent states as a control, particularly since states with lower GNI per capita are more likely to lose.¹¹⁶ Finally, we have included a cubic year trend variable in all models.¹¹⁷

Before turning to the results, it is important to note the limits of the explanatory model. Our approach is ‘X’ focused; we are testing whether specific mood variables could explain variation in outcomes in investment treaty arbitration, subject to a set of controls. We are not conducting a larger ‘Y’-based and fully explanatory model that seeks to capture all reasons for claimant/investor success rates. This potentially limits the casual findings in two ways. The first is that it is important to distinguish between arbitrator and systemic reflexivity. As principals, states may adopt strategies that directly affect the underlying legal framework (that is, the IIAs themselves) in which investment treaty arbitrators operate. While we are developing measurements to capture this variable, we doubt, however, that the legal framework governing foreign investment has shifted by such a degree to solely explain the variance in arbitration outcomes. This is principally because almost all of the decisions under analysis in this article are based on IIAs that were drafted before the emergence of the legitimacy crisis. Moreover, even where there is an arbitration based on a newer generation treaty, we have found that expected outcomes are not always generated.¹¹⁸

¹¹⁵ This measure takes the value (1) if a respondent state has had five or more cases registered against it in a given year and (0) otherwise. The case clusters in the full set of cases registered are: Argentina (2002, 2003, 2004), Czech Republic (2005), Ukraine (2008), Egypt (2011), and Venezuela (2011, 2012).

¹¹⁶ Behn, Berge and Langford, *supra* note 94.

¹¹⁷ We have followed here the approach of Creamer, *supra* note 79. However, we also tested other specifications for a time trend. See discussions below.

¹¹⁸ See Behn and Langford, *supra* note 97. Moreover, Alschner, *supra* note 66, finds that newer treaties on average are more investor-friendly than state-friendly.

The second danger is that the relationship between claimant/investor success and future litigation may be partly endogenous. The growing awareness of the open legal opportunity structure¹¹⁹ of investment treaty arbitration may have prompted foreign investors to bring more dubious cases. If so, a possible consequence is a rise in the number of cases being lost, but not because of arbitrator reflexivity. We have only begun to develop a legal strength indicator for each case, but we are uncertain as to whether there has been a recent uptick in dubious cases, at least at the liability/merits stage. The likelihood of claimant/investor success dropped quite early – well before the possibility of a wave of dubious cases entering the system – and has remained quite steady across time, at least until very recently. In the case of jurisdiction decisions, this endogeneity argument may have more explanatory power. There has certainly been a recent decrease in claimant/investor success rate in jurisdictional cases; although even this might be explained by reflexivity, with arbitrators tightening jurisdictional criteria as a response to the legitimacy crisis. In any case, trying to separate out these effects is a clear task for a future research agenda on investment treaty arbitrator behaviour.

Table 2 sets out the logit regression results. The controls in Model 1 are largely as expected. Law firm advantage and extractive industry case controls are positively correlated

Table 2: Logit Regression Results for State and Stakeholder Mood (Any Win)

	Controls	Treaty exits	New treaties	Google Scholar	All
	Model 0	Model 1	Model 2	Model 3	Model 4
Independent Variables					
State Mood 1 (treaty exits)		-0.03 ^d			-0.03
State Mood 2 (new treaties)			0.01 ^d		0.01 ^a
General Mood (Google Scholar)				0.001	0.01
Controls					
NAFTA-based case	-0.44	-0.50	-0.48	-0.44	-0.53
ICSID-administered case	-0.14	-0.15	-0.16	-0.14	-0.17
Extractive industry case	0.58 ^b	0.60 ^b	0.58 ^b	0.57 ^b	0.59 ^b
Law firm advantage	0.38 ^a	0.34 ^a	0.35 ^a	0.38 ^b	0.31
State learning	0.06	0.07 ^a	0.06 ^a	0.05	0.07 ^a
Case cluster	0.70	0.63	0.70 ^a	0.69 ^a	0.62
GNI per capita (logged)	-0.33 ^b	-0.34 ^b	-0.33 ^c	-0.32 ^c	-0.35 ^c
Cubic year trend	0.00004 ^b	0.00004 ^a	0.00008	-0.00007	-0.00005
Chi ²	33.80	39.40	36.13	34.03	40.03
Observations (number of cases)	388	388	388	388	388

Notes:

^a $p < 0.10$; ^b $p < 0.05$; ^c $p < 0.01$.

^d In an earlier online workshop version (with cases up to 31 July 2016), the variable State Mood 1 (treaty exits) and State Mood 2 (new treaties) were both significant. The coefficients now lie just outside the zone of significance: $p = 10.9$ and 12.9 respectively.

¹¹⁹ Hilson, 'New Social Movements: The Role of Legal Opportunity', 9 *Journal of European Public Policy* (2002) 238.

with claimant/investor success, while respondent state development status (as measured by GNI per capita [logged]) is negatively correlated. The remaining control variables are not statistically significant although they carry the expected sign with the exception of state learning; respondent states do not appear to gain an advantage from facing repeat litigation. Turning to the State Mood 1 indicator, the coefficient is as expected, namely negative. An increase in unilateral IIA exits corresponds with a decrease in claimant/investor success. While this indicator is not significant in Model 1 and the full Model 4, it is so for the subset of liability/merits decisions.¹²⁰ The State Mood 2 indicator is positive as expected but only significant in Model 4. A rise in the number of IIAs signed correlates with investor success.

The general mood indicator is not statistically significant, and the coefficient is surprisingly positive (although very small). Testing with alternative outcome indicators (full wins, jurisdiction decisions, liability/merits decisions), we only find a significant and negative relationship at the jurisdictional stage.¹²¹ This relationship may reflect the recent decline in jurisdictional successes for claimant/investors and the ongoing rise in the amount of legitimacy crisis scholarship annually. There is some doctrinal evidence of reflexivity in jurisdictional awards, although we are doubtful that the recent drop in success rates here for claimant/investors can be fully explained by arbitral behaviour. This is particularly so when there has been an upwards drift in claimant/investor success at the liability/merits stage of the dispute (especially in the period from 2012 to 2014).

Returning to the two state mood indicators, we now look at the magnitude of the measured shift. In other words, how much work do these factors (which have been significant in some or many models) potentially do in explaining variation in outcomes? This can be graphically observed in Figure 6, which shows the predicted probabilities for five-unit differences in the treaty exits indicator (State Mood 1). Holding all other control and mood variables constant at their means, the probability of a claimant/investor win is 56 per cent when the treaty exit indicator is at zero. Yet it falls to 38 per

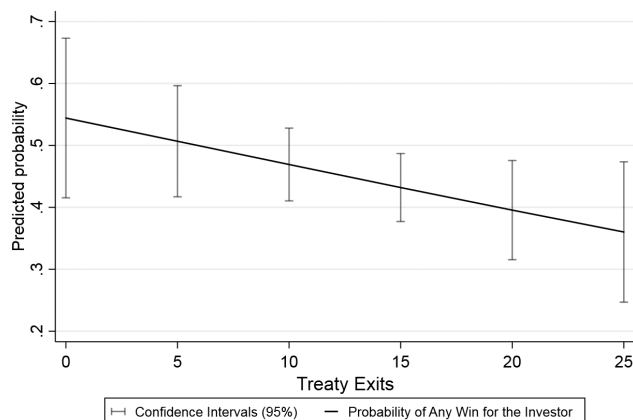


Figure 6: Predicted Outcomes for State Mood 1 (Treaty Exits)

¹²⁰ See extra tables available at www.jus.uio.no/pluricourts/english/topics/investment/research-projects/database.html.

¹²¹ *Ibid.*

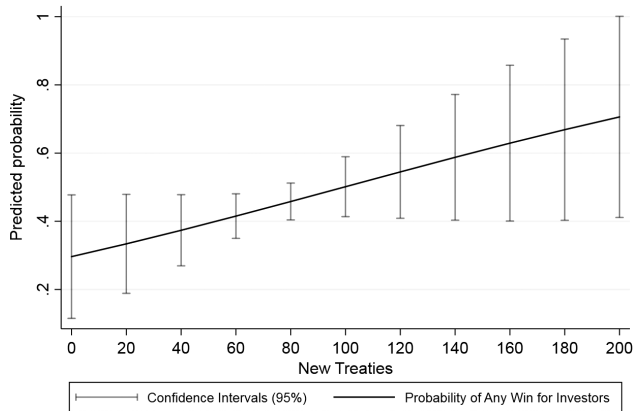


Figure 7: Predicted Outcomes for State Mood 2 (New Treaties)

cent when the number of annual IIA exits rises to 25 (which occurred in 2009). The differences are not enormous, but they are noticeable nonetheless.

In the case of the new treaties indicator (State Mood 2), the differences across the indicators' range are even more dramatic. Holding all other variables constant, claimant/investors achieved 80 per cent success rates in lagged years where there were close to 200 IIAs that were signed annually. But this drops to 30 per cent in lagged years where the number of annual IIAs signed bottoms out at 30 per year (see Figure 7). However, it is important to note that the confidence intervals at the ends of ranges for both state mood indicators are large. Thus, many claimant/investors are still able to achieve reasonable levels of success in these years, but the mean likelihood of such successes is lower. Overall, the tests on these yearly indicators suggest a weak or modest relationship between stakeholder mood and arbitral outcomes. The significance of the correlations is sensitive to changes in the model and sample period. For example, use of a walking and squared time trend variable led to a decline in the number of mood variables that were significant. Cutting off the sample period earlier in 2016 increased the number of mood variables that were significant. Variables such as development and case type (for example, extractives sector) explain significantly more of the variance in outcomes.

2 Influential States

The final stakeholder hypothesis, Hypothesis 3 on influential state signals, is measured differently. We break up outcomes according to five three-year crisis periods that follow 2001 and correspond to our analysis in section 2. This disaggregation allows us to examine possible structural breaks after interventions by a small number of large influential states (primarily the USA but also the EU) that we believe may have disproportionate signalling power. It is a cruder approach to measurement but may better reflect the nature of legal adjudication with periodic, rather than frequent, paradigm shifts.¹²² The first structural breaks relating to influential states are the pro-state signals sent by the NAFTA

¹²² Rueda, 'Legal Language and Social Change during Colombia's Economic Crisis', in J. Couso, A. Huneus and R. Sieder (eds), *Cultures of Legality: Judicialization and Political Activism in Latin America* (2010) 25.

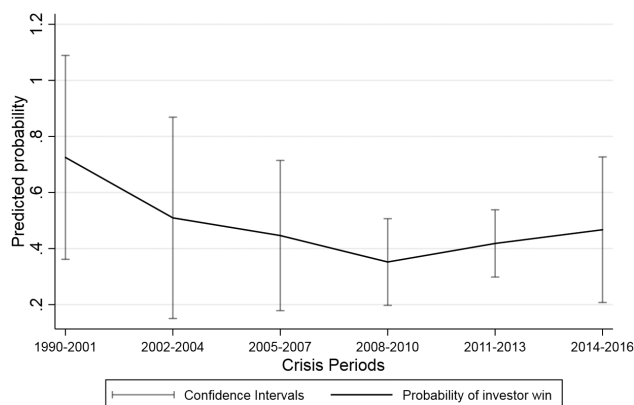


Figure 8: *Influential State Signals and Structural Breaks*

state parties after the issuance of the FTC Interpretive Note in 2001 and the release of the new US Model BIT in 2004 as well as the pro-investor signals sent by the ramping up of negotiations by the USA, the EU and China for large-scale bilateral and plurilateral trade and investment treaties (which include ISDS), particularly after February 2013.¹²³

Controlling for the same factors as above, Figure 8 shows the predicted probabilities in each period for claimant/investor success. It is notable that the probability of success does fall after the first break (after 2001) and the second break (after 2004), but the decrease in claimant/investor success is only statistically significant after the second break.¹²⁴ Turning to the last structural break (after 2013), the average success rate for claimant/investors is comparable to all of the preceding periods. However, it is notably that claimant/investor success rates are not different (no statistical significance) from the period from 1990 to 2001. While the p-scores hover close to the 10 per cent level, the large confidence interval from 2014 to 1 August 2017 reveals the fact that many claimant/investor are enjoying success that is almost comparable to the period from 1990 to 2001. While the measure is crude, the recent pattern may suggest that influential states may be exercising a renewed subtle influence on investment treaty arbitrators; a factor paradoxically enhanced by the results for the development status asymmetries shown below.

3 Public Interest Cases

For Hypothesis 4 on public interest cases, it was hypothesized that investment treaty arbitrators may display greater state deferentialism in cases raising public interest concerns or arousing public interest. These are cases that have gained notoriety during the process of being litigated because they involve very large compensation claims, challenges to human rights-related or environment-related measures, or challenges to legislative rather than administrative (executive branch) action. Extreme examples would include, *inter alia*, the *Phillip Morris*,¹²⁵ *Vattenfall*,¹²⁶ and *Yukos* shareholder cases.¹²⁷ Less extreme

¹²³ A number of these large bilateral and plurilateral negotiations were officially launched prior to 2012, but we use 2012 as the year when these negotiations ramped up significantly.

¹²⁴ See note 114 above.

¹²⁵ See *Phillip Morris*, *supra* note 40, at 40.

¹²⁶ See *Vattenfall*, *supra* note 41.

¹²⁷ See *Yukos*, *supra* note 49.

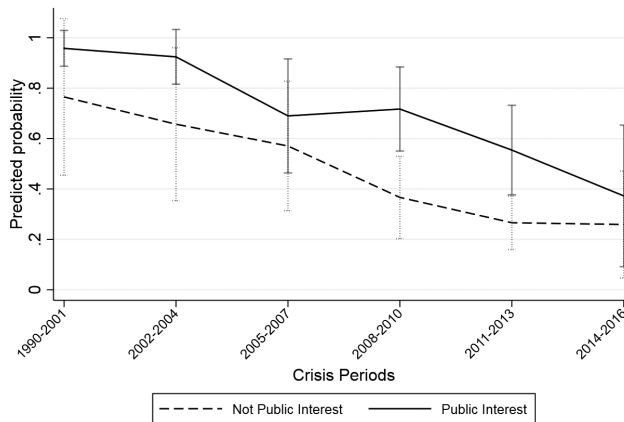


Figure 9: Public Interest Cases across Crisis Periods

examples include, *inter alia*, cases discussed largely on points of law such as the *Salini* case,¹²⁸ the Argentinian bondholder cases,¹²⁹ and some of the early NAFTA-based arbitrations like *Metalclad*¹³⁰ and *Methanex*.¹³¹ We have thus created a binary public interest case indicator for all investment treaty arbitrations that fall into these categories and in which we would expect arbitrators to be aware of the public interest dimension involved in these cases; 84 of the 389 finally resolved cases in our dataset meet this criteria.

Turning to measurement, we created an interaction term between public interest cases and the different crisis periods. This enables us to measure both the general trend in these cases and whether arbitrator reflexivity is comparably greater than in all other investment treaty arbitrations. Figure 9 sets out the results.¹³² As is clear, claimant/investors consistently do better in public interest cases as we have defined them. This is largely because such cases disproportionately occur in the extractive industry economic sector where claimant/investor success rates are consistently high, which we suspect is caused by the fact that these cases are often about the amount of compensation, not liability.

Turning to reflexivity, the rate of claimant/investor success in public interest cases has fallen in parallel with the general decline in claimant/investor success. This is notable given that we might suppose that claimant/investors will generally have consistently good chances in the subset of extractive industry cases. However, it is nonetheless difficult to discern any significant change between public interest cases and all other cases over time. Only in the periods from 2005 to 2007 and from 2014 to 2016 do we notice a contraction in the difference between these two categories. But this difference is not statistically significant.

¹²⁸ ICSID, *Salini Costruttori S.p.A. and Italstrade v. Morocco – Decision on Jurisdiction*, 31 July 2001, ICSID Case no. ARB/00/4.

¹²⁹ See ICSID, *Abaclat and Others v. Argentina*, ICSID Case no. ARB/07/5, settled; ICSID, *Ambiente Ufficio and Others v. Argentina*, ICSID Case no. ARB/08/9, discontinued; ICSID, *Giovanni Alemanni and Others v. Argentina*, ICSID Case no. ARB/07/8, discontinued.

¹³⁰ ICSID, *Metalclad v. Mexico – Award*, 30 August 2000, ICSID Case no. ARB(AF)/97/1.

¹³¹ UNCITRAL, *Methanex v. United States – Final Award*, 19 August 2005.

¹³² As based on the 84 finally resolved cases that meet the definition of a ‘public interest’ case.

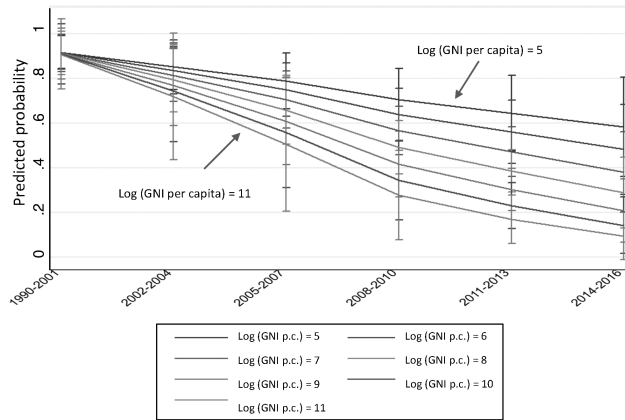


Figure 10: Development Status across Crisis Periods

4 Development Status Asymmetries

For Hypothesis 5 on development status asymmetries, we predicted that investment treaty arbitrators might be more deferential to lesser developed states after the emergence of concerns about an alleged anti-developing state bias. We measured this by breaking up the sample into the five three-year crisis periods and examining the claimant/investor success rates for various levels of development. The results were the opposite of what was expected. As Figure 10 demonstrates, states with higher levels of development enjoyed most of the decline of claimant/investor success over time. Thus, states at the high end of the scale (GNI per capita [logged] between 9 and 11) have seen claimant/investor success rates drop from 90 per cent to around 15–20 per cent over the last two decades, but the poorest states (between 5 and 7) faced claimant/investor success rates of 40–60 per cent. This suggests that some states matter more than others, and lends some further and indirect support to the influential state hypothesis.

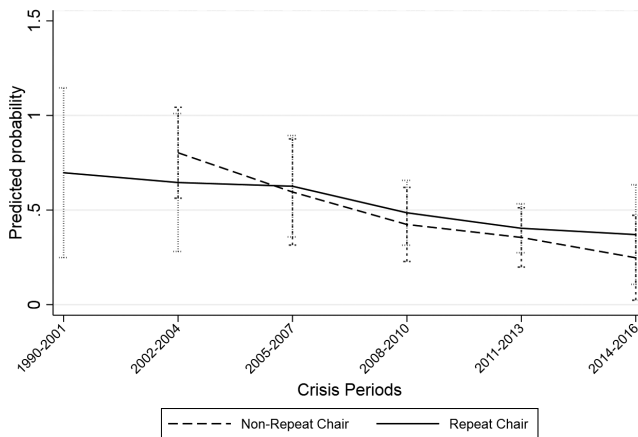


Figure 11: Prominent Arbitrators across Crisis Periods

5 Prominent Arbitrators

Finally, regarding Hypothesis 6 on prominent arbitrators, we code for the presence of a tribunal chair who has rendered five or more decisions (as a tribunal chair). Using an interaction term, we test whether the presence of a prominent tribunal chair decreased the chances of claimant/investor success in the different periods after 2001 relative to other cases. As Figure 11 shows, investment treaty arbitrations with a prominent tribunal chair were slightly more likely to award claimant/investors any success from 2005 onwards, which is the reverse of what was expected. However, the differences are not statistically significant, suggesting that prominent tribunal chairs are not acting in any significantly different way than other tribunal members.

5 Conclusion

Since the mid-2000s, the international investment regime has been subject to a 'legitimacy crisis'. While the regime has its ardent supporters, the mood of various stakeholders from a diverse group of states, scholars and global social movements has tilted towards viewing the regime as pro-investor, pro-Western and jurisprudentially incoherent. We have not tried to solve this normative debate in this article but, instead, have focused on its effects. We have asked whether investment treaty arbitrators are reflexively evolving and helping the system adapt to more legitimate and effective forms of international adjudication (by becoming more deferential to respondent states in investment treaty arbitration). The article set out various rational choice and discursive-based reasons for thinking that investment treaty arbitrators would be sensitive and adaptive. We countered these reasons with a competing set of legalistic (and attitudinal) reasons that may inhibit investment treaty arbitrators from acting in such a fashion. Drawing upon a newly developed investment treaty arbitration database (PITAD), we demonstrated that there has been a significant drop in claimant/investor success across time and found modest and suggestive evidence that investment treaty arbitrators have shifted their behaviour on some types of outcome.

Our main finding on reflexivity is that states matter. Indicators measuring general state mood (IIA exits and new IIAs signed) and the intervention of influential states were the most strongly correlated with the variations in claimant/investor success (or lack thereof), while our indicator for the general stakeholder mood tracked investment treaty arbitration outcomes poorly. Notably, these results resonate with the general doctrinal developments in international investment law and cohere with research on domestic courts, where judges are found to be sensitive to influential state actors but less responsive to broad diffuse public opinion. However, the research represents only a first take on the question of reflexivity in investment treaty arbitration. The field is ripe for further quantitative and qualitative research. We have only touched the surface and our findings remain tentative. There is room for improvement of the dependent and independent variables as well as the use of qualitative methods. A particular issue is why some of our specific hypotheses were not confirmed. Why were prominent arbitrators no more reflexive than their counterparts? Why was there not a greater drop in

claimant/investor success rates in public interest cases than in all other cases? Is it because the reflexivity effect is not so strong? Or is it because reflexive arbitral behaviour is fragmented throughout the international investment regime, such that only some arbitrators act in a strategic manner? This latter speculation would certainly cohere with the patchwork nature of doctrinal development, but it remains a question to be investigated.

Returning to the normative debates with which we began, our research presents a divergent contribution. On the one hand, we have shown that the system has adjusted considerably since its infancy. Claimant/investor success rates have fallen dramatically and hover around 40 per cent today. This is certainly much lower than the comparable figure of 90 per cent-plus for applicant states in WTO litigation. On the other hand, we have found a clear asymmetry in the distribution of the reflexive gains for states. It is developed states that are the beneficiaries of the large drop in claimant/investor success rates; less developed states have only registered marginal benefits. Moreover, claimants in investment treaty arbitration continue to do dramatically better than litigants success rates in the international human rights regime. The result is thus a mixed picture. The investment treaty arbitration system has been able to enhance effectively its respect for state sovereignty (and partly regulatory autonomy), but some states are more equal than others.