

Contributions and Limitations of Empirical Research on Independence and Impartiality in International Investment Arbitration

GUS VAN HARTEN*

Abstract

The use of investment treaty arbitration to decide public law raises concerns about judicial independence and impartiality. These concerns arise from the absence of institutional safeguards of independence that are otherwise present in public law adjudication at the domestic or international level. In this article, opportunities to use empirical methods to study possible bias in investment arbitration are surveyed. The discussion includes a brief consideration of qualitative methods and a critique of two quantitative studies on outcomes in investment arbitration. The discussion then turns to the methodology of an ongoing project involving legal content analysis of decisions by investment treaty tribunals. The main conclusion reached in the paper is that empirical research can make important contributions to scholarly understanding of investment arbitration. On the other hand, empirical research has important limitations in its ability to demonstrate the presence or absence of actual bias, even at a systemic level, thus reinforcing the need for institutional safeguards.

Key words

Investment arbitration, adjudication, bias, independence, fairness, empirical legal research.

* Gus Van Harten, Associate Professor, Osgoode Hall Law School of York University, gvanharten@osgoode.yorku.ca. This paper is part of an ongoing project on empirical research and investment arbitration. The author is grateful for the comments on this paper or related work provided by José Costa, Susan Franck, Kevin Gallagher, Mark Kantor, Andrew Lang, Sol Picciotto, Lauge Poulsen, Mariana Mota Prado, Greg Shaffer, two anonymous peer reviewers of a related article, and participants in the Oñati International Institute for the Study of Law's Workshop on Socio-Legal Perspectives on the Adjudication of International Economic Disputes, the LSE Transnational Law Project's Workshop on Paradigms, Methods, and Analogies in Investment Treaty Law, and the Osgoode-Jindal Workshop on Transnational Governance. Kelly Goldthorpe, Rana Arbabian, and Kirsten Mikadze provided research assistance. Any errors or omissions are my own.

Table of contents

Contributions and Limitations of Empirical Research on Independence and Impartiality in International Investment Arbitration	1
Abstract	1
Key words	1
Table of contents	2
Introduction	3
1. Conceptual background	3
2. Qualitative research on investment arbitration	6
3. Quantitative research on investment arbitration	8
A. The general issue of relevance	9
B. Review of Franck (2009)	9
C. Review of McArthur and Ormachea (2009)	13
4. Methodology of a legal content analysis project.....	15
A. Overall aims	16
B. Identification of issues and description of issue resolutions.....	17
C. Data sources and coding process	19
D. Some other methodological concerns	23
E. Some anticipated limitations	23
5. Conclusion	24
References	26
Appendix A.....	28

Introduction

The importance of investment arbitration as a decision-making process warrants scrutiny of its design and performance. Empirical methods, alongside deductive reasoning and doctrinal analysis, may provide useful tools in this inquiry. Such methods can be used to initiate, reformulate, deflect, or clarify theories. They enable observations and insights that would not be possible without the systematic collection and analysis of information (Heise; Hall and Wright, p. 65-6). That said, empirical methods should be used with caution and modesty (Baldwin and Davis, p. 880-1; Merton, p. 103; Tyree). Empirical research has important limitations in its ability to provide definitive answers on whether inappropriate factors have influenced or will influence specific decisions (Sisk and Heise, p. 746 and 794).

There are numerous options for socio-legal research on potential bias in adjudication. In this paper, the discussion begins with a conceptualization of bias in investment arbitration and an outline of hypotheses arising from the conceptualization. The potentially constructive role of qualitative methods is then highlighted. Next, two recent studies are discussed that sought to use quantitative or statistical methods to test for bias in investment arbitration. Finally, the methodology of an ongoing project based on a legal content analysis of awards is outlined as an alternative or supplemental approach to researching possible bias.

Used with care, empirical methods may reveal new information and insights and thus advance our understanding of social phenomena. Yet none of the approaches to empirical research as discussed in this paper offer clear answers to concerns about fairness and integrity in investment arbitration arising from its institutional structure. Indeed, the limitations of empirical research in the study of adjudicative bias point to the importance of establishing an institutional structure of adjudication that accounts for reasonably perceived bias. Fairness and integrity in adjudication must be ensured ultimately at the institutional level due to the incomplete knowledge of the parties and the public about the mindsets of individual decision-makers. An important concern in investment arbitration, therefore, is whether its structure allays reasonable concerns about bias so as to support public confidence in the process (e.g. *R v Valente*, para. 22).

1. Conceptual background

In this section of the paper, the concept of bias in investment arbitration is elaborated with a view to highlighting the potential relevance of different approaches to empirical research (Berg, p. 25). The focus here is on theories that provide rationales for expectations of bias in investment treaty arbitration, following from its characterization as a unique form of public law adjudication constituted at the international level (see Van Harten (2007)).

The use of arbitration to decide public law on a final basis is unique to investment treaty arbitration. This raises special concerns about independence and fairness in the process. Perceptions of bias may follow especially from the lack of institutional safeguards of judicial independence and the absence of a robust process of judicial review. These concerns are germane to investment treaty arbitration because it combines several characteristics, as follows.

Public law dynamic of claims against the state: In investment treaty arbitration, unlike in most fields of international adjudication, states have given their general and prospective authorization for claims to be brought by private parties (typically, owners of assets of substantial value) against a state acting in its sovereign capacity, without the possibility of reciprocal claims by the relevant state against those private parties.

Autonomy from domestic courts and tribunals: Under many investment treaties, unlike any other international adjudicative arrangement that allows for claims by private parties' against states in their sovereign capacity, the customary duty of a person to resort to domestic remedies – before an international claim can be brought on the person's behalf – is removed or heavily restricted.

Remedy of state liability: Unlike other regimes of judicial review, especially at the domestic level, the primary remedy in the system is that of a damages award against the state following a finding of unlawful sovereign conduct.

International enforceability: Unlike other decisions and awards in public law, investment treaty awards are enforceable in numerous countries with limited or no opportunity for judicial review.

Due to this combination of features, arbitrators acting under investment treaties are authorized to resolve questions of public law without the prospect of judicial review by a court, whether domestic or international. This raises concerns about independence and impartiality in the adjudicative process because it permits the final resolution of public law outside of any court, whether domestic or international (Van Harten (2008)). Institutional safeguards of independence that are otherwise present in the judicial resolution of public law are not present. These safeguards include, for example, judicial security of tenure, bars on outside remuneration by the judge, and an objective method of assignment of judges to individual cases. Without such safeguards, a perception of bias may arise that decision-making by investment treaty arbitrators may be influenced by inappropriate factors.

These inappropriate factors include, in particular, the financial and career interests of arbitrators who aspire to re-appointment and who may have an interest to promote the arbitration industry as an alternative to other forms of decision-making. Such factors point to the apparent dependencies of arbitrators on at least three sets of actors: (a) those who act as senior gatekeepers or otherwise wield influence in the arbitration industry; (b) those who exercise power in the arbitration centres that operate as appointing authorities under the treaties;¹ and (c) those (here, investors) who in the asymmetrical context of public law adjudication are able to initiate claims. Each of these groups of actors has influence over the financial or career success of arbitrators by triggering the use of the system, by appointing or green-lighting arbitrators, by deciding conflict of interest claims against arbitrators, by employing arbitrators in their wider professional contexts, or by otherwise shaping on a case-by-case basis the demand for the services of investment arbitrators and the arbitration industry.

¹ The primary arbitration centres are the International Centre for the Settlement of Investment Disputes in Washington, the International Chamber of Commerce in Paris, the Stockholm Chamber of Commerce, and the Permanent Court of Arbitration in the Hague.

This brief outline of factors offers a set of rationales for empirical study of possible bias in investment arbitration. The perceptions follow from the unique institutional design of the process and especially from the absence of safeguards of judicial independence and robust judicial review. In this respect, the critique offers a basis for further research on expectations of bias and the institutional context. As examples, one might consider these hypotheses:

If arbitrators have an interest to encourage claims, then one is unlikely to see large numbers of claims dismissed at the jurisdictional or merits stage.

If arbitrators are sympathetic to the interests of business actors, then one is unlikely to see widespread state-friendly interpretations of ambiguous language in investment treaties.

If arbitrators are inclined to please influential actors in the arbitration institutions, then one is less likely to see claims dismissed when brought by investors of powerful states that wield influence within these institutions and more likely to see claims dismissed when brought against such states.

If arbitrators are worried that powerful states may reject use of the system, one is likely to see arbitrators adopt relatively sympathetic approaches to treaty constraints on states in cases against powerful states and stricter approaches in cases brought by investors of those states.

These are some expectations that follow from the theoretical framework of possible bias arising from the institutional structure of investment treaty arbitration. They could be tested at a systemic level by the use of empirical methods, even if only tentatively and with various limitations that would preclude any strong conclusions about actual bias in the system (let alone in particular cases) or about the degree to which there is a basis for perceived bias arising from institutional factors.

Importantly, one should not mistake an evaluation of the institutional structure of an adjudicative system for a comprehensive theory of the behaviour of adjudicators in that adjudicative system. A range of factors and complex interactions of factors will be present in the thought processes of adjudicators and the deliberations of tribunals. One may hope and trust that the predominant factors under consideration by arbitrators are concerns for fairness and integrity. However, the difficulties of establishing what factors have guided the actual decisions in specific cases make it vital – for the sake of the public, the parties, and the adjudicators themselves – to incorporate institutional safeguards of independence. Otherwise, and regardless of actual outcomes, reasonable doubts may linger that a decision was unfair. Fundamentally, the limitations of our ability to know the mindset of a decision-maker in any specific case, including by the use of empirical methods, warrants the incorporation of institutional safeguards.

2. Qualitative research on investment arbitration

Discussions of empirical research often focus on the use of quantitative and statistical methods (e.g. Schuck). However, qualitative methods provide a useful means by which to examine decision-making by uncovering unquantifiable facts about people and phenomena (Berg, p. 7). Like quantitative methods, their use involves testing specific hypotheses of how theoretical expectations will manifest themselves in observable phenomena. For example, in the present context, it might be useful to conduct interviews with participants in investment arbitration in order to gather their views about how the system operates and about the implications of the absence of safeguards of judicial independence in the system. Interview and survey-based studies have been conducted on the impacts of investment treaties for decision-making about foreign investment (e.g. Yackee (2010)) and there is room for further qualitative study of the arbitration process.

On the other hand, the issue of actual bias – whether arbitrators individually or as a group are influenced by inappropriate factors in the resolution of legal issues or disputes – would be challenging and perhaps impossible to test reliably using interviews or surveys (Wagner and Petherbridge, p. 1125-6). Even the unworkable option of questioning adjudicators under oath about whether they will be, are, or were influenced by inappropriate factors in their decisions might not provide sufficient reassurance where there is no way to verify independently the adjudicator's own reports of his or her state of mind. This problem is recognized in legal doctrine stressing the importance of institutional safeguards to counter perceptions of bias (e.g. *Locabail*, p. 471-2).

Sophisticated qualitative work has not been conducted on the adjudicative process in investment treaty arbitration. However, sophisticated work was carried out in the field of international commercial arbitration and this work demonstrated how qualitative methods can be used to gather information that elaborates on the conceptualization of decision-making in international arbitration. The seminal study is *Dealing in Virtue* by Dezalay and Garth (1996). In this study, Dezalay and Garth examined international commercial arbitration as an autonomous legal field in a global marketplace by interviewing about 300 participants in the field. The study is relevant to the issue of possible bias in investment arbitration mainly for the context that it provides about the world of international arbitration as a business activity. There is a lot of overlap between the arbitrators, lawyers, law firms, arbitration rules, arbitration centres, and treaty arrangements that operate or apply in international commercial arbitration and investment treaty arbitration.

To illustrate, the following observations by Dezalay and Garth in their study of international arbitration assist in elaborating rationales for possible bias based on the absence of institutional safeguards in investment arbitration, as outlined above. These observations are clearly not fully explanative of the state of affairs in investment arbitration. That said, they do pinpoint and elaborate some reasons to suspect or perceive bias in situations where international arbitration is tasked with the final resolution of public law:

The operation of the market in the selection of arbitrators... provides a key to understanding the justice that emerges from the decisions of arbitrators (Dezalay and Garth, p. 9).

The new generation of technocrats... emphasizes their ability to satisfy the consumers in order to gain repeat business (p. 194).

For the lawyers and their justice, the question is how to affirm the autonomy necessary for legitimacy while at the same time manifesting sufficient fidelity to the economic powers who must in the end find these services worth purchasing and deploying (p. 70).

It is good arbitration politics to thank business lawyers or other acquaintances who bring nice arbitration matters by letting them have limited access to the arbitration market. This system of exchange of favors is essential to success in arbitration, a career dependent on personal relations (p. 124).

The growth of the market in arbitration is also evident in the competition that can be seen among different national approaches and centres (p. 7).

They [the newcomer arbitrators of the 1980s and 1990s] present themselves... as international arbitration professionals, and also as entrepreneurs selling their services to business practitioners... (p. 36).

The ICC [International Chamber of Commerce] has... become one of the principal places where the "politics" of arbitration is elaborated and expressed. There are innumerable committees and multiple networks of influence that gravitate around this institution. The [ICC International Court of Arbitration], for example, which is really an oversight committee that reviews arbitration appointments and decisions, appears to be particularly sensitive to the business clientele... (p. 45).

The multinational companies are in this way investing in the construction of these legal services that serve them (p. 93).

These observations convey that arbitrators operate in a marketplace in which each is a supplier of symbolic capital arising from his or her reputation. As operators in the market, each has an interest to further his or her position and that of the industry overall (Dezalay and Garth, p. 8 and 18). With the passing of the old generation of "gentleman" arbitrators in the 1980s and 1990s, a new generation of technocrat arbitrators was found to be more likely to seek to promote their industry or preferred arbitration forum in relation to its competitors (Dezalay and Garth, p. 50).

Naturally, some arbitrators would have a greater connection than others to the enterprise of the arbitration industry or a particular arbitration centre. In light of Dezalay and Garth's observations, one could attempt to measure the degree of an arbitrator's connection to the industry by gauging the frequency of the arbitrator's appointment in the relevant field of arbitration. The assumption would be that arbitrators appointed more often have a greater career stake in the relevant field than those appointed less often. In turn, one would expect frequently-appointed arbitrators to take greater care to respond to the interests of powerful actors, as laid out above. These are examples of how one might draw on the observations of Dezalay and Garth in order to refine and test theoretical expectations about decision-making in investment arbitration.

Another observation of Dezalay and Garth was that international arbitration consists of networks of cross-connected players who affiliate around prominent centres of arbitration, such as the International Chamber of Commerce. One could seek to examine this point in the context of investment arbitration by identifying the prominent appointing authorities. In order of priority, based on the frequency of claims in different forums, these appear to be ICSID (in Washington), the Permanent Court of Arbitration (the Hague), the International Chamber of Commerce (Paris), and the Stockholm Chamber of Commerce. Alternatively, one could seek to elaborate the professional backgrounds of arbitrators (see Costa), perhaps focusing on those who are appointed most frequently. This would allow further testing of Dezalay and Garth's observations that arbitrators tend to come from a discrete professional community and social milieu.

Another observation of Dezalay and Garth highlighted how prominent arbitrators may name each other for appointments and exclude those who are not part of the industry. To what degree, if at all, does this happen in investment arbitration? One way to test this would be to collect information about the identity of legal counsel and of arbitrators in specific cases and to look for correlations between individuals. If one saw that X was appointed where Y was counsel and that Y was appointed where X was counsel, this would support the finding of cross-connections based on Dezalay and Garth's work. In turn, it might raise questions about the degree to which arbitrator decision-making is influenced by the identity of counsel appearing before the arbitrator or by the influence of prominent counsel over the re-appointment of arbitrators.

These are a few examples of how the qualitative work of Dezalay and Garth, informed by the conceptualization of independence and impartiality in public law that was laid out in a previous section of this paper, could be used to refine expectations about possible bias in investment arbitration. This indicates how qualitative research can enrich understanding of decision-making even if it cannot prove or dis-prove actual bias in specific cases or affirm or dispel reasons for perceived bias. There are, of course, many opportunities for further research. For example, it appears that persons from outside the realm of commercial arbitration have been appointed increasingly to decide investor-state disputes. Do these persons come from different professional networks and do they approach their role differently from Dezalay and Garth's technocrat arbitrators?

3. Quantitative research on investment arbitration

There have been a handful of studies that seek to test hypotheses of possible bias in investment arbitration. In this section, after elaborating on the issue of relevance in empirical research, I review briefly two examples: Franck (2009) and McArthur and Ormachea (2009).² Both focus on hypotheses of actual bias at a systemic level rather than of actual bias in specific cases or perceived bias arising from institutional factors. Each provides an interesting elaboration of empirical research, although the findings in the studies are mixed and inconclusive, especially due to the lack of data. Both also have important limitations and in some cases flaws, including a tendency – especially in Franck's study – to mis-state or over-state conclusions (see Garth, p. 103-4). This latter point highlights the need for caution in the use of empirical methods and for particular scrutiny of empirical research

² A more extensive review of both studies, which as I understand may include responses from the authors of the two studies, is forthcoming in the *Yearbook on International Investment Law and Policy*.

where it purports to connect statistical findings to strong statements of findings and conclusions.

A. The general issue of relevance

Dezalay and Garth's study of international arbitration was discussed in the previous section of this paper. It was indicated that some of the conclusions reached in that study had a degree of relevance to investment arbitration, but that the primary focus of the study was on other research topics, primarily relating to commercial arbitration. One can relate the findings of Dezalay and Garth to questions of bias in investment treaty arbitration only with caution. In doing so, one should focus on observations that follow from the original research assumptions and that are relevant to distinct research questions concerning possible bias in investment arbitration.

The same points about relevance apply to quantitative research. Indeed, it is particularly important to connect the design of a quantitative project to a specific research question and hypothesis, given the demands of statistical analysis and the need to track assumptions underlying the reduction of qualitative phenomena to numerical representations (Aitken and Taroni, p. 203). If one is not careful about how the numbers are presented, they may be given undue weight by readers who do not understand the assumptions and limitations behind the researcher's quantitative findings.

B. Review of Franck (2009)

Franck examined hypotheses arising from certain factors that could generate bias in investment arbitration. These factors involved possible prejudices of arbitrators linked to their nationality and/ or the identity of the respondent state. Both factors were grouped according to the "development status" of arbitrators and countries and then compared to the outcomes in cases (Franck (2009), p. 438). The researcher's hypothesis was that development status would not affect outcome and "that arbitrators can make decisions neutrally on the basis of the facts and law" (Franck (2009), p. 454). As designed, the study tested the first element of this hypothesis, and led to further conclusions on the second element.

To test the hypothesis, the researcher analyzed outcomes in 52 treaty cases with publicly-available awards. Two measures were then applied to classify the development status of presiding arbitrators and respondent states in those cases. The first measure was OECD membership; presiding arbitrators and respondent countries were treated as "developed" if they (or their countries of nationality) were members of the OECD and as "developing" if they were not OECD members. Second, presiding arbitrators and respondent countries were classified based on the World Bank income classification system, which divides countries into low income, lower-middle income, upper-middle income, and high income categories.³ Based on

³ In classifying outcomes, the researcher characterized as a win for the respondent state an award of no damages against the state; and treated as a loss an award of damages against the state. Second, the researcher evaluated outcome in terms of the amounts of money awarded in cases where the claimant was successful. Due to limitations in the available data, the researcher's sample was reduced to 49 cases

this analysis, the researcher did not find reliable evidence of variations between case outcomes linked to the development status of presiding arbitrators and respondent countries. In turn, Franck reached various conclusions about the integrity and fairness of investment arbitration.

The outline of the methodology used in this study is commendable for its clarity and transparency. However, the study has important limitations and in some cases flaws. Most important is the extent to which the researcher over-stated or mis-stated certain findings and conclusions. The following are examples.⁴

In the study's abstract, it is stated:

"The results demonstrate that, at the macro level, development status does not have a statistically significant relationship with outcome. This suggests that the investment treaty arbitration system, as a whole, functions fairly and that the eradication or radical overhaul of the arbitration process is unnecessary." (Franck (2009), p. 435).

In the text of the study, regarding the results on development status and win/ loss rates, it is stated that the consistency of her results "offers a powerful narrative that there is procedural integrity in investment arbitration" (p. 464).

In the text of the study, regarding the results on development status and amount of damages awarded, it is stated that the "lack of a main effect for a respondent's development status stands in sharp contrast to the assertions that investment treaty arbitration unfairly privileges the developed world or improperly harms the developing world" (p. 470).

In the study's conclusion, it is stated:

"The notion that outcome is not associated with arbitrator or respondent development status should be a basis for cautious optimism. It provides evidence about the integrity of arbitration and casts doubt on the assumption that arbitrators from developed states show a bias in terms of arbitration outcomes or that the development status of respondent states affects such outcomes. It suggests that major structural overhaul may not be necessary because it is not clear that arbitration is inherently predisposed towards particular outcomes." (p. 487).

These statements were misplaced for various reasons, summarized below.

The study suffered from a gross lack of data which led to a 40 to 80% risk of error across most of the results.⁵ The researcher did not report this high risk of error –

for the OECD measure and 47 or 49 cases (depending on whether a win/ loss or total damages outcome was being measured) for the World Bank measure.

⁴ A companion study by Franck (2007) was also reviewed in the research for this paper.

⁵ To elaborate, a claim of statistical significance about a hypothesized connection (or lack of connection) between variables requires sufficient data to remove significant risks that apparent relationships are explained by chance. The researcher calculated the number of awards needed to generate reliable findings (on the chosen standard, findings that carried a 20% chance of error). Depending on the

and the corresponding finding that there was insufficient data to test reliably the hypotheses of the study – alongside the above statements of conclusions about the fairness and integrity of investment arbitration. The lack of data precluded the drawing of conclusions to support or refute the hypothesis that development status would not affect outcome.

In the study, the high risk of error was identified in a series of footnotes (Franck (2009), p. 461-70 (various footnotes)). However, the main conclusion drawn by the researcher from this risk of error was that further research should be conducted. In fact, the risk of error should have precluded the study's main conclusions, as quoted above. Problematically, Franck relied on the lack of reliable evidence of a hypothesized link between development status and outcome to infer that there was no link between development status and outcome (Franck (2009), p. 460-2). This inference does not follow because it is possible – indeed, it is to be expected in light of the limited data – that there would not be reliable evidence of either possibility (i.e. of either a link or the absence of a link).

Likewise, due to the lack of reliability of the results, the study did not establish that “development status does not have a statistically significant relationship with outcome”, as was concluded by the researcher (Franck (2009), p. 435). And, the high risk of error should have precluded the researcher from reporting that “outcome is not associated with arbitrator or respondent development status” (Franck (2009), p. 487). Although not highlighted by the researcher, the clearest conclusion to be drawn from the study was simply that there was insufficient data to test the hypothesis with an acceptable degree of reliability.

In the study, developed-country status was equated with OECD membership. However, some OECD countries are reasonably regarded as developing or transition countries. Following a review by the present author of the study's results,⁶ it emerged that 44% of the 18 cases classified in the study as against developed countries were cases against Mexico and 17% were cases against the Czech Republic or the Slovak Republic. If one were to treat these countries as developing or transition countries, then 61% of the cases classified as being against developed countries (and 22% of the total cases) would need to be re-classified. The significance of these cases as a proportion of the cases classified as being against developed countries, and the lack of disclosure about these coding outcomes in the study itself, raises a concern that the OECD measure might have been chosen in order to even out the cases against developed versus developing countries in a limited dataset. Although use of the OECD measure in the study was balanced by the use of the World Bank measure, discussed below, the lack of transparency about the implications and limitations of the OECD measure was problematic.

The World Bank income classification system was also in the study as a measure of development status. This was a more sophisticated measure than the OECD measure and it was commendable for the researcher to use both of the measures. However, the use of the World Bank measure was also frustrated by the lack of

measure, between 382 and 781 awards were required (Franck (2009), p. 461-70 (various footnotes)). However, the available sample sizes in the study were between 47 and 49 awards. This generated a 40 to 80% chance of error (so-called “Type II error”, which refers to the risk of a false negative) across most of the study's results.

⁶ Following requests to the researcher and an attempt to replicate an aspect of the study, the data for Franck (2009) was provided to the author. This enabled the review of the data relating to the OECD measure as summarized in this paper.

data. There were no cases decided by presiding arbitrators from low income countries and there was only one case decided by a presiding arbitrator from a lower-middle income country. As a result, all but one of the cases classified in the study as being decided by a “developed country” presiding arbitrator involved an arbitrator from an upper-middle income country (Franck (2009), p. 459). The researcher was transparent about this limitation of the study. However, the lack of data nevertheless undermined the relative sophistication of the World Bank measure and was a further reason to moderate the study’s conclusions.

The study also used winsorized data, meaning that the available data was truncated to remove “outliers” (Franck (2009), p. 456). However, winsorizing works best with data that is symmetric and creates biases with data that is non-symmetric. The monetary awards reviewed in the study were not symmetric, which justified discarding the winsorized method. Further, when winsorizing data, it is typical to truncate 20 to 25% of the data. However, the tables for winsorized and non-winsorized data have the same number of cases (49) and it is not clear how other data was substituted for the “outliers” and whether this might have led to the loss of important information about the actual dataset (Franck (2009), p. 465-8).⁷ Finally, the overall lack of data in could mean that all of the data should be treated as outliers.

Leaving aside these important limitations of the study, Franck’s statements of conclusions focused on a limited and selective set of explanations for the results – especially the possibility that the system functions fairly – while neglecting others. One of many alternative explanations was the possibility that arbitrators do not make decisions based on nationality but rather on their membership in a common culture or industry of arbitrators. Another was the possibility that the facts of cases involving developing countries are more or less favourable to investors than the facts of cases involving developed states, and that these differences should lead us to expect variations in outcomes. That is, even if it did find reliable evidence of either a link or no link between development status and outcomes, the study would have benefited from greater clarity about the fact and range of alternative explanations for the results.

The study examined only narrow aspects of questions about bias in investment arbitration. There was insufficient data to answer reliably the hypotheses of the study. Regardless of its results, then, the study could not provide a “powerful narrative” for or against the procedural integrity of the system (Franck (2009), p. 464). Likewise, the statement that the study “suggests that the investment treaty arbitration system, as a whole, functions fairly and that the eradication or radical overhaul of the arbitration process is unnecessary” (Franck (2009), p. 435) was over-stated in light of the limitations of the study. Far more information involving a wide range of factors, examined in numerous studies, would be required to contemplate these sort of conclusions.

To be clear, this critique of the study does not lead to the conclusion that it offered any reliable evidence of bias or of a lack of fairness or integrity in investment arbitration. The key lesson is that there is too little information available to draw reliable conclusions either way. More broadly, these problems with the study are not flaws with empirical methods in general but with the way in which those

⁷ I am grateful to an anonymous peer reviewer of a related article and to Lauge Poulsen for pointing out these limitations of the winsorizing method.

methods were employed in one study. An earlier study by Franck provided useful descriptive data on investment arbitration and demonstrated how empirical research can advance understanding of what is happening in the system (Franck (2007)). However, this review of the present study indicates that a researcher must present findings and conclusions accurately and with care. Otherwise, there is a danger that readers, including policy makers, will take up a study for positions that the research does not support, as appears to have happened in the case of Franck's work.⁸

C. *Review of McArthur and Ormachea (2009)*

Another study that sought to test arbitrator performance, including in relation to possible bias, was carried out by McArthur and Ormachea. Their approach was to review ICSID decisions on jurisdiction (up to February 2007), arising both from investment contracts and from investment treaties, to assess whether jurisdiction was accepted or declined by the tribunal. This aspect of case outcomes was then evaluated against a series of variables, including the legal basis for the claim (i.e. contract or treaty), the economic status of the claimant's state of nationality and of the respondent state, and the institutional quality of the respondent state.

The researchers reported a number of findings and highlighted three as the most provocative.⁹ The first was that ICSID tribunals rarely dismissed claims on jurisdictional grounds. The second was that tribunals were more likely to dismiss claims on jurisdictional grounds where the claim was brought under an investment contract rather than an investment treaty. The third was that tribunals were less likely to dismiss claims by investors from the wealthiest countries.

Relative to the study by Franck that was reviewed above, McArthur and Ormachea were modest and cautious in presenting their conclusions. The study was framed as a "first step toward the development of a comprehensive empirical literature" on ICSID arbitration (McArthur and Ormachea, p. 562). Its main concluding statement was as follows (McArthur and Ormachea, p. 583):

The integrity and neutrality of the ICSID system are thus of the utmost importance. Our study on the jurisdictional jurisprudence represents an important first step towards developing a literature that facilitates understanding and awareness of these issues.

This statement does not make strong claims about the reliability of findings or overstate conclusions. In large part, the researchers took the approach of laying out their results without engaging in a stretched commentary on their significance (McArthur and Ormachea, p. 563 and 582-3). Even so, there are aspects of the

⁸ See the separate statement by a member of the U.S. Model BIT Review committee, Mark Kantor, who cited Franck's conclusion on development status and outcome in support of his position against reform of the U.S. Model BIT. United States, *Report of the Advisory Committee on International Economic Policy Regarding the Model Bilateral Investment Treaty* (30 September 2009), Annex B (Particular Viewpoints of Subcommittee Members), Statement of Mark Kantor, p. 28-9.

⁹ The researchers did not offer much explanation for why these findings were thought to be provocative other than to report that they contradicted the researchers' expectations (McArthur and Ormachea, p. 574).

study, especially involving the conclusions drawn, that are problematic. These are summarized below.

In the first place, the researchers mentioned issues of interest to them in the conduct of the study, including possible bias in investment arbitration (McArthur and Ormachea, p. 564, 576, and 583). However, they did not identify any *a priori* hypothesis of the study.

In some cases, the researchers appeared not to be familiar with theoretical rationales for expected bias in investment arbitration. For example, they did not connect the descriptive finding that arbitrators were less likely to accept jurisdiction in cases under investment contracts than investment treaties (McArthur and Ormachea, p. 574) to expectations of pro-investor bias in investment treaty arbitration. The expectations of perceived bias outlined earlier in this paper suggest that arbitrators acting under investment treaties (as opposed to investment contracts) might have an incentive to decide issues in favour claimants so as to encourage future claims. This is because, under the treaties, investment arbitration is asymmetrical in that only investors can initiate a case against states; in contrast, in contract arbitration both states and investors can typically bring claims against each other.¹⁰

In some respects, the researchers were not rigorous and transparent in describing their methodology.¹¹ For example, they stated that they each coded independently a small sample of each other's cases (McArthur and Ormachea, p. 564) but did not indicate the number of cases subjected to this double-coding and the level of consistency (i.e. inter-coder reliability) obtained. Also, while the researchers discussed some coding challenges, such as the problem of how to code cases with multiple claimants, they did not identify or discuss other apparent issues.¹²

It is difficult to determine from the study precisely how many cases were coded and the relevant breakdown for each of the researchers' analyses, making it difficult to evaluate the significance of the results. However, the researchers listed in an appendix to the study all of the cases reviewed with a summary of the coding results. This was helpful because it allows other researchers to replicate the process of the study, including simply by coding the same cases for the same variables.

The researchers did not attempt to calculate the statistical significance or effect sizes of the results. That said, unlike in Franck (2009), the researchers did not state any findings in terms of statistical significance. Generally speaking, the study was

¹⁰ This is not to suggest that this descriptive information on jurisdictional outcomes under investment contracts should be interpreted to support expectations of pro-claimant bias. This interpretation would derive from only one of many possible explanations for the finding, which is based on very limited data.

¹¹ This was partly because the researchers relied (reasonably) on measures of economic status and institutional quality that were drawn from outside sources. Also, without wishing to downplay the useful contributions of the study, it should be noted that the researchers were legal practitioners, not academics, and thus may have had less background on methodology or less time to commit to developing a more rigorous methodology.

¹² For example, the researchers reported that they coded cases for the existence of a "claim based on a breach of [a] contract" with a government entity (McArthur and Ormachea, p. 566). However, it was not clear whether and how this was coded in any situations where a claim was brought under an investment treaty but based on an argument of a breach of a contract (e.g. in the case of an umbrella clause in the treaty). One assumes that this issue was resolved in a reasonable way but it would have been helpful to see the issue identified and explained.

presented appropriately as simply a tentative outline of descriptive information about ICSID arbitration.

The study was based on limited data (74 cases were classified in the analysis of ICSID jurisdictional outcomes) and, in some instances, McArthur and Ormachea did not moderate their conclusions in light of the limited data. For example, it was reported that tribunals were less likely to accept jurisdiction when the respondent state had a low score of institutional quality (McArthur and Ormachea, p. 576). However, the data on this variable was extraordinarily limited; for countries with high institutional quality, 1 of 8 cases was dismissed on jurisdictional grounds, while for countries with low institutional quality, 2 of 7 cases were dismissed. In light of this limited data, the researchers described the finding of a possible connection between high institutional quality and the acceptance of jurisdiction as tentative (McArthur and Ormachea, p. 576). Yet they returned to this finding in the study's conclusion, as discussed below.

In the study's conclusion, it was not appropriate to claim that the data demonstrated "that investors are always better off relying on a BIT to establish jurisdiction" (McArthur and Ormachea, p. 582) and to state that "claims are most likely to fail against those very same countries where an investor had the best reasons to demand international arbitration protections..." (McArthur and Ormachea, p. 583). These predictive claims were not supported by the study.

Some statements in the conclusion suggest that the researchers may have approached their project with some pre-set conclusions. For example, in the text of the study the researchers identified as "striking" and a "remarkable finding" the descriptive information that U.S. investors were 40% less likely to have their claims rejected on jurisdictional grounds (McArthur and Ormachea, p. 581). However, the researchers did not mention this finding in their concluding statements about the possibility of pro-investor bias. By itself this would be fine, but the researchers opted instead to highlight that investors were more likely to see their claims rejected where the host state had poor institutional quality and, in turn, to cite this as something that "undermines the contention that the entire international arbitration system evinces an undue preference for investor interests over state interests" (McArthur and Ormachea, p. 583). There are various problems with this statement, including the exceptionally limited data on institutional quality and the lack of explanation for why the results should be taken to reflect a pro- or anti-investor bias. It would have been more appropriate to convey simply that the study's results presented mixed and inconclusive evidence on possible bias, and perhaps to highlight the relevance and reliability of the highly tentative descriptive results in either direction. Relative to Franck, however, the researchers were more cautious about the presentation of results and conclusions.

4. Methodology of a legal content analysis project

The two studies discussed above focused on the classification of investment arbitration decisions based on their outcomes. This is a valid approach to gathering and analyzing information about adjudication. But there are also important limitations of an outcome-based approach (Hall and Wright, p. 85-7; Wagner and Petherbridge, p. 1127-8). For example, data on outcomes as a measure of actual performance is open to a range of alternative explanations such as variations in the strength of parties' claims, diversity of fact situations, possible inflation by claimants of amounts claimed, procedural variations among forums, varying

experience levels and incentives among arbitrators, and varying political influences of states and private actors. It is also very difficult, if not impossible, to identify the “appropriate” or “fair” spread of outcomes against which actual outcomes are to be measured. Related to this is the problem that data on outcomes at one or another stage of a case do not capture aspects of tribunal decisions – such interpretations of the law – that may reflect bias independently of outcome. Also, cumulative data on case outcomes does not explain whether some element of the decision-making in any particular case was influenced by inappropriate factors. As such, there will always remain a prospect for perceived pro-investor or pro-state bias in any case based on inadequacies of the institutional structure. This is a concern that institutional safeguards of judicial independence are meant to address.

Alternative approaches to empirical research may be useful to supplement outcome-based studies. In this section of the paper, the methodology of an ongoing project is outlined, especially as it relates to the collection and coding of data. The project relies on an analysis of the legal content of investment treaty awards rather than of outcomes (Hall and Wright; Neuendorf; Krippendorf; Stemler). The purpose of this outline of the project is not to proselytize for legal content analysis as a form of empirical legal research (Hall and Wright, p. 64) or to augment criticisms conveyed in the previous section. It is simply to introduce an ongoing project and its methodology. Because the project is ongoing, results and findings are not reported. When findings are reported, they should be scrutinized carefully and criticized for any mis-statements or over-statements of findings and conclusions or for any other flaws.

A. Overall aims

The purpose of the project was to collect descriptive information on awards (and other decisions) in investment treaty arbitration. Of particular interest was the degree to which different awards may reflect different approaches to the resolution of legal issues. A more specific objective was to test hypotheses that followed from the absence of institutional safeguards of independence in investment arbitration.¹³ Methodologically, the project was conceived as an experiment in legal content analysis, motivated by an aspiration to produce research that combined interdisciplinary research tools in order to avoid the methodological weaknesses of each tool operating in isolation. The project blended the legal expertise of a researcher in the analysis of legal content of investment treaty awards with the statistical expertise of a research collaborator.¹⁴ Legal expertise was employed primarily to pinpoint hypotheses and to refine the data collection process in order to provide a sophisticated base of data for the statistical analysis (Yackee (2008)). The statistical analysis was in turn designed and implemented primarily by an expert in statistics (Hall and Wright). This, it was hoped, would enhance the role of legal expertise as a component of empirical legal research and avoid limitations of

¹³ The three hypotheses under examination were that (a) arbitrators would tend to adopt expansive approaches to legal issues more frequently than restrictive approaches; (b) the tendency toward expansive approaches would be accentuated for claims by investors of the U.S., U.K., Germany, and France as major capital-exporters, powers within the appointing authorities, and historical treaty-makers in the system; and (c) the tendency would shift toward restrictive approaches for claims against the U.S., U.K., Germany, and France. Data analysis to test these hypotheses is still underway.

¹⁴ The collaborator was a law student, now graduated and working in immigration law, who was previously a Senior Analyst with Statistics Canada. A second statistician has been engaged to develop statistical models to test and analyze the data.

outcome-based research that is carried out by legal researchers who are not statisticians.

B. Identification of issues and description of issue resolutions

Awards were coded for the presence and resolution of any of seven *a priori* issues of law. All of the issues related primarily to the topics of jurisdiction or admissibility, although in some cases they also engaged other types of legal issues. The focus on issues of jurisdiction and admissibility had no special attraction. It was chosen to provide an initial basis by which to organize the classification system and coding process, leaving options for further research that focused on substantive standards or procedural aspects of investment law.

The seven issues, and different interpretive approaches to each, were identified and described in advance based on a review of existing awards and secondary literature. Where an issue was found to have arisen in a case, each arbitrator's resolution of the issue was classified as expansive, restrictive, or "non-classifiable".¹⁵ The expansive and restrictive approaches to each issue were defined in advance. They reflected interpretive approaches that tended either to enlarge or reduce the compensatory promise of investment treaty arbitration for prospective claimants and, in turn, the liability risk for states. "Non-classifiable" issue resolutions encompassed situations where the issue resolution, as explained in a tribunal's reasons, did not fall reasonably within the scope of the expansive or restrictive category for the issue.¹⁶

The expansive and restrictive approaches reflected interpretive positions on legal issues of uncertainty or ambiguity in investment treaties. Due to divergent interests of investors and states, however, it is not necessarily the case that the expansive approach is "pro-investor" and the restrictive approach is "pro-state". For example, in considering whether expansive interpretations of the treaties are pro-investor, would we examine the specific claimant, future claimants, or investors as a whole? If expansive interpretations were to lead states to withdraw from the system or if they undermine the system's legitimacy then they might be said to harm investors as a whole while benefiting specific claimants. Likewise, a restrictive position might be considered anti-state if it was adopted because arbitrators who were motivated by a desire to protect the interests of a few powerful states. For these reasons, the language of pro-investor and pro-state was consciously avoided.

An issues template laid out the interpretive approaches that would qualify as expansive or restrictive for the seven issues.¹⁷ Coding was intended to cover a reasonable sample of divergent approaches to issues that related to the topics of jurisdiction or admissibility. The template was developed based on a review of

¹⁵ An arbitrator was deemed to have resolved an issue where the arbitrator put his or her name to the reasons for an award or decision that explained the issue resolution in question. An arbitrator who gave separate reasons for a decision would be coded differently than other members of the tribunal where his or her separate reasons resolved the issue differently from other members.

¹⁶ These non-classifiable situations included those in which (1) the issue was resolved specifically and expressly by the terms of the relevant treaty, (2) the claim or argument was withdrawn by a party, or (3) the tribunal reasoned that it was unnecessary to resolve the issue that had arisen.

¹⁷ A copy of the issues template, laying out the expansive and restrictive approaches to each of the issues, is attached as Appendix A.

existing awards and secondary literature as well as consultations with experts in the field (Hall and Wright, p. 107). This approach was adopted to reduce the discretion of the primary legal researcher (Van Harten) in the identification of issues and delineation of issue resolutions, although significant discretion undoubtedly remained. The coding was not intended to encompass all aspects of the possible legal interpretation of coded issues. Rather, for each issue, one or more descriptions of an expansive or restrictive approach was identified and issue resolutions were then coded either as expansive or restrictive based on whether the tribunal's interpretation fell reasonably within any one of the descriptions.

The following issues were identified for coding:

Corporate person investor – whether a claim would be permissible where ownership of the investment extends through a chain of companies running from the host to the home state via a third state.

Natural person investor – whether a claim would be permissible where brought by a natural person against the only state of which the person is a citizen or against a state of which the person is a citizen without confirmation of dominant and effective nationality.

Investment – whether the *Fedax* criteria would be applied to the concept of investment under the *ICSID Convention* or, regardless of whether under the *ICSID Convention*, whether there would be a requirement for an actual transfer of capital into the host state as a feature of an investment or an extension of the concept of investment to non-traditional categories of ownership

Minority shareholder interest – whether a claim by a minority shareholder would be permitted where the treaty did not allow claims by minority shareholders (such as where the treaty did not include “shares” in the definition of investment) or without limiting the claim to the shareholder's interest in the value and disposition of the shares (as opposed to interests of the domestic firm itself).

Permissibility of investment – whether there would be an evident onus placed on the investor or the state to show that an investment was or was not affirmatively approved or was or was not based on corrupt practices.

Parallel claims – whether a claim would be allowed in the face of a treaty-based duty to resort to local remedies that clearly had not been satisfied by the claimant; in the face of a contractually-agreed dispute settlement clause relating to the same factual dispute; in the face of an actual claim, arising from the same factual dispute, via the relevant path of a treaty-based fork-in-road clause; or in the face of an actual claim, arising from the same factual dispute, via another treaty that could lead to a damages award in favour of the investor.

Scope of most-favoured-nation treatment – whether the concept of most-favoured-nation treatment would be extended to non-substantive provisions of other treaties (such as dispute settlement provisions).

This sample of issues was intended to capture legal questions that had arisen in existing cases and that were likely to have led to different approaches by different arbitrators. The focus was legal issues of general significance on the basis that such issues can operate as signals for investors and for states of the likelihood of successful claims and of state liability (Trujillo, p. 364-5), in this case at the jurisdictional and admissibility stage. Legal issues that were highly specific to a single treaty or case, on the other hand, were not examined. Factual determinations were also excluded from the study on the basis that there is greater variation among facts of cases than among general legal standards. That said, the dividing line between law and outcome was not always clear. In some cases, based on the approaches adopted in existing awards, aspects of the resolution of legal issues were linked to aspects of the substantive outcome in a case (e.g. whether a particular resolution of a legal issue was accompanied by a decision to allow the claim to proceed). However, an issue resolution was not coded for whether it determined the outcome of an award. Thus, distinctions between the *ratio decidendi* and *obiter dicta* of a decision were avoided on the basis that they could be difficult to maintain and that an issue resolution, even if only *obiter*, could still give signals to future claimants and respondents based on its symbolic meaning (Krippendorff, p. 22).

The issues template was developed in the spring and summer of 2009. It was recognized that over time divergences in the law may evolve or be resolved. Particular issues and interpretive approaches may fade in importance. Recognizing this, a degree of residual flexibility was maintained for exceptional cases in which an interpretive approach appeared to have evolved from that described in the issues template in ways that were nevertheless consistent with the underlying rationale for the classification as expansive or restrictive approaches to the resolution of an issue. This flexibility was exercised rarely, based on the general guideline that the *a priori* characterization of the categories needed to be maintained strictly but not slavishly. The flexibility was limited to cases in which the interpretive approach appeared clearly to build on, rather than to take in new directions, the *a priori* expansive or restrictive approach.

The unit of analysis for the study was issues rather than cases. An issue-by-issue approach allowed an examination of the reasoning of arbitrators on a variety of issues that may affect a decision. However, a limitation of this approach was that it gave greater weight to cases in which multiple issues arose and were resolved. This raised the possibility of an in-built bias in favour of either expansiveness or restrictiveness in cases where multiple issues arose. In particular, it was thought that an in-built bias in favour of an expansive approach to issue resolutions might arise because of the need, in theory, for tribunals to reject all objections to jurisdiction and admissibility where a claim was allowed to proceed, and to accept only one such objection where a claim was rejected.

C. Data sources and coding process

The primary source of data was the text of awards (i.e. all decisions by a tribunal) in known cases decided by May 10, 2010 and publicly-available by June 1, 2010. An award was deemed "publicly-available" where it was posted on the Investment Treaty Arbitration website maintained by Andrew Newcombe of the University of Victoria. A case was deemed "known" where it was listed on that website with an indication that the case was brought under an investment treaty or where it was

listed as a treaty-based case on any of a series of official websites.¹⁸ In the case of a few fields, data was derived from sources other than the text of awards. In NAFTA Chapter 11 cases, for example, information on the date of claims was taken from materials filed by the parties rather than simply the text of awards. Also, the nationality of arbitrators was derived from a google search as a supplement to the text of awards.

The coding process had two stages. One involved the collection of basic descriptive information on cases. The other involved a content analysis of awards. The process of the first stage was follows. Coding was based on a codebook that was reviewed periodically in order to address unanticipated permutations of the coding process.¹⁹ The codebook included numerous fields of relatively innocuous information, of which the following were most relevant to the legal content analysis: name of case, name and nationality of claimant, name of respondent state, date of claim, dates of known awards in a case, and identity and nationality of arbitrators associated with each award in a case. Beginning in the summer of 2008, extensive descriptive data on known investment arbitrations was collected by three research assistants (all JD students) acting in sequence over a three-year period.²⁰ The second and third researchers acted with knowledge of earlier coding. Thus, the descriptive data was double-coded, in some cases triple-coded, but not blindly. At this stage of the project, coder discretion was relatively limited; the discretion involved mainly how to resolve issues not originally anticipated and dealt with in the codebook.²¹ A key concern was to ensure that coders were well-trained in the codebook – in the present case, they were afforded generous time to study the codebook, an opportunity to practice its application on a sample of cases, detailed feedback on practice coding, and an ongoing opportunity to discuss issues arising from the application of the codebook (while retaining autonomy in the coding of specific cases) – and confident of the independence of their decisions as coders.

Roughly 15% of the descriptive data was checked by the primary legal researcher. Where discrepancies were found, the data was referred back to the coder or in rare cases re-coded by the primary legal researcher (in the case of the fields listed above, only data on dates of claims was re-coded in this way). Formal inter-coder reliability was not tracked at this stage of the project, although periodic checks of the data by coders or the primary researcher revealed that errors were uncommon and, where they arose, were typically due to straightforward misunderstandings of the codebook.

A more rigorous coding process was used at the legal content analysis stage. Following development of the issues template, a sample of cases was coded as a test. Next, awards were reviewed systematically in groups according to treaty type and, in the case of Argentina (which has faced many more claims than other states), the respondent state. The coding at this stage involved reading all of the awards in known cases in order to determine whether an issue had arisen and, if

¹⁸ These included the websites of ICSID, the Permanent Court of Arbitration, and the Energy Charter Treaty Secretariat, and the government websites of Canada, Mexico and the U.S. on NAFTA Chapter 11 arbitration.

¹⁹ Further information on the codebook is available from the author: gvanharten@osgoode.yorku.ca.

²⁰ All of the coders were experienced in the use of Excel and in data collection and were able to make useful contributions to the development of the coding process.

²¹ For example, under some investment treaties, a notice of claim is filed by an investor at a date prior to the filing of the actual claim. In such situations, which date should be recorded as the date of the claim? The approach adopted was to use the earlier of the two dates as the date of the claim.

so, whether its resolution fit one of the expansive or restrictive approaches for that issue. Coding was carried out by three coders beginning in the summer of 2009. The first and main coder was a research assistant and JD student who coded independently all of the awards. The second was the primary legal researcher who coded independently all of the issues identified by the first coder as having arisen, with access to the decisions of the first coder on whether particular issues arose in a case. Discrepancies in the coding of issue resolutions were discussed by the two coders. Where the coders did not agree on the appropriate coding of the issue resolution, the issue was referred to a third coder who was a research assistant and lawyer. This third coder resolved on an independent and anonymous basis the remaining differences between the first two coders.

As originally planned, the coding process did not envision that the second coder would review the first coder's decisions that an issue had not arisen in a case. Instead, it was planned that an audit would be performed of a sample of cases under different treaties or treaty types.²² In the audit, all of the awards in the sampled cases were reviewed by the second coder in order to determine whether any issues from the template, not identified by the first coder as having arisen, had in fact arisen in the view of the second coder. The results of the audit – especially, the frequency with which issues not identified by the first coder were found to have arisen by the second coder²³ – led to a review of all cases by the second coder. Where an issue was found to have arisen that had not been identified by the first coder, this was referred back to the first coder who reviewed the second coder's decision that an issue had arisen and that it should be coded as expansive, restrictive, or non-classifiable. If there was a difference of views between the first and second coders on whether an issue had arisen and, if so, on the classification of the issue resolution, then this was referred to the third coder.

At this second stage of the project, the collection of data involving issue resolutions was based on an analysis of the text of publicly-available awards. Reflecting a limitation of content analysis, this meant that factors and reasons not outlined in the written reasons were not captured in the coding (Hall and Wright, p. 100). On this point, to maintain consistency across cases, the data source of publicly-available awards was not supplemented by analysis of other sources, such as materials filed by the parties (as are typically available under NAFTA Chapter 11) or scholarly or journalistic reports on particular cases. Further, a tribunal's reasons in an award were examined where they engaged an issue in the issues template, regardless of the stage of the proceedings at which the issue arose or the manner in which it was characterized by the tribunal. Thus, the specific question for coders was whether the issue arose and, if so, how it was resolved based on the issues template; it was not how the issue was framed by the parties or the tribunal, although this clearly had a bearing on whether an issue could be found to have arisen.

²² The cases reviewed in this audit were chosen at random, using random.org, from a list of cases under each treaty or treaty type. Where a randomly-generated case clearly was not representative of coded cases, a substitute case was chosen (this occurred in one instance, where *Loewen* was substituted for *Canfor* on the basis that the latter dealt with issues of consolidation of cases rather than issues of jurisdiction or admissibility). Also, substitute cases were chosen randomly where the award(s) in a case was not available in English or not publicly-available.

²³ In the case of NAFTA, 4 cases were audited, leading to 0 instances in which an issue not identified by the first coder was found to have arisen. For the Energy Charter Treaty, 3 cases were audited, leading to 1 instance in which an issue not identified by the first coder was found to have arisen. For BIT cases involving Argentina, 4 cases were audited, leading to 2 instances in which an issue not identified by the first coder was found to have arisen. For all other BIT cases, 19 cases were audited, leading to 9 instances in which an issue not identified by the first coder was found to have arisen. Of these 9 instances, 4 involved a separate opinion by an arbitrator in the case.

In all cases, the approach was to review cases where publicly-available awards were available that referred to matters of jurisdiction or admissibility. Where a tribunal's award was not available in English,²⁴ the case was not coded for legal content so as to maintain consistency across the three coders (who shared only English as a common language). Where there was no reference in any publicly-available award to any jurisdictional or admissibility issue, the case was coded as not public for all issues.²⁵ Thus, issues were coded only when they arose in the text of an award, meaning that, in some cases, issues arising in a case may not have been coded because they were dealt with by the tribunal in an award that was not among the publicly-available awards for the case. This was a consequence of the varying levels of transparency across cases and treaties. That said, it was very rare to see cases that had a published award which did not address jurisdictional or admissibility issues to some degree.

The data thus included the full universe of publicly-available awards in known investment treaty arbitrations. As such, the coded cases reflected a sample of total cases, assuming that there are cases that have been decided under investment treaties but the existence of which is not public. There is good reason to believe that there are such cases.²⁶ Thus, if the body of known awards is different from unknown awards, then the results of the study will be subject to case-selection bias (Drahozal, p. 294). For example, in order for awards in known cases to be public, the cases must not have been settled before the stage of a first award by a tribunal. And, the characteristics of the conflicts in known cases may differ from the characteristics of conflicts that have been settled at an earlier stage. In turn, the study of known awards may reflect case-selection bias in favour of certain kinds of investor-state disputes. Even if there are significant differences between known and unknown awards in treaty arbitrations, analysis of known awards remains a feasible method by which to identify trends in arbitrator decision-making, so long as any conclusions drawn are based on appropriate inferences and limited to the studied universe (Hall and Wright, p. 105; Krippendorf, p. 25-8).

To reduce the risk of selection bias arising from the analysis of only known awards, the project focused on issues arising from decisions that must be made by arbitrators in or after an award on jurisdiction. This is likely to mitigate the problem of data availability, on the assumption that the existence of a case is less likely to remain confidential where the case has reached the stage of an award on jurisdiction. On the other hand, this does not address cases where an award is known to exist but has not been released to the public in a form that allows for legal content analysis of the award. This is further limitation and possible basis for selection bias due to the secrecy of some investment arbitrations.

²⁴ Awards were coded where available originally in English, via an official translation to English, or via an unofficial translation posted on the Investment Treaty Arbitration website.

²⁵ Thus, in *Lemire v Ukraine*, the only publicly-available decision was a record of settlement by the parties which followed a non-public award on jurisdiction. The case was coded as not public.

²⁶ Besides anecdotal reports, the International Chamber of Commerce reported during 2005-2009 an average of 69 arbitrations per year involving a state or state entity (International Court of Arbitration). However, further information about such arbitrations is confidential. It is, as such, unclear how many of these arbitrations involving states were pursuant to investment treaties, contracts, or domestic legislation.

D. Some other methodological concerns

The coding of separate opinions, whether dissenting or concurring, presented some challenges that were dealt with as follows. Where the separate opinion dealt with a code-able issue in explicit terms, then the reasons were coded separately from the tribunal's main reasons. In such circumstances, where the separate opinion identified an issue that was dealt with in the main reasons for the award, but the author of the opinion did not declare a position on the issue of interpretation, then the issue resolution for the relevant arbitrator was coded as non-classifiable. However, where the separate opinion did not raise explicitly an issue that was dealt with in the tribunal's main reasons, then the relevant arbitrator was assumed to share the view of the tribunal as a whole. In the vast majority of cases that led to separate opinions by arbitrators, it was not difficult to either distinguish the separate opinion from the tribunal's main reasons or identify it with the tribunal's main reasons.

Coder discretion was an integral part of aspects of the legal content analysis and various steps were taken to limit and guide this discretion (Hall and Wright, p. 109). These included the development of an *a priori* issues template; resort to external sources to generate the descriptions of interpretive approaches in the template; and use of a double-coding process, supplemented by a tie-breaker, to check coding decisions. At the legal content analysis stage of the project, especially, coding decisions of the first and second coders will need to be evaluated for inter-coder reliability. Even if there is a reasonable level of inter-coder reliability, however, discretionary choices will clearly have been present in various elements of this stage of the project. Ultimately, accountability for these coding decisions lies in the publication of the template and coding results, such that the results can be reviewed by other researchers for purposes of replication (Krippendorf, p. 49-50).

E. Some anticipated limitations

Analysis of the results of the study is ongoing and so the results are not reported here. However, some limitations of the study can be highlighted at this stage (see Wagner and Petherbridge, p. 1128-30). Fundamentally, the study seeks to assess tendencies in legal reasoning based on the quantification of complex processes of adjudicative decision-making. Such research has important limitations. First, it is difficult if not impossible to reduce complex qualitative and variable phenomena to the quantitative indicators that are required for statistical analysis (Aitken and Taroni, p. 203). Studies in the field deal in approximated correlations rather than firm conclusions. Second, the focus on jurisdiction and admissibility may make the results unrepresentative of other types of legal issues; this points to opportunities for further research. Third, there is no control group. In the present project, a control group might be one in which judges who enjoyed institutional safeguards of independence then decided cases under investment treaties. The interpretive approaches of these judges could then be compared to those of arbitrators in similar (or the same) cases. Naturally, there is no control group of judges, thus precluding this sort of comparison. Fourth, in investment arbitration, there is a gross lack of data relative to other adjudicative contexts (in which thousands or tens of thousands of decisions are available for analysis). Attempts at quantitative analysis thus tend to carry a high risk of error.

Finally, it is important to emphasize that the project was not intended to prove or disprove actual bias. Rather, the aim was to identify tentative trends in decision-making on matters of legal interpretation. Thus, the project seeks to identify roughly the manoeuvring space available for the interpretation of issues in existing awards and secondary literature, and the positioning of the reasons of tribunals within that space. As the jurisprudence evolves, so too may definitions of what qualifies as an expansive or restrictive approach. This is especially important where, as in investment treaty arbitration, there is no hierarchical system of adjudicative decision-making, thus creating greater fluidity and less predictability in case law. At most, the project will offer a blurry snapshot of how arbitrators have resolved jurisdictional and admissibility issues in publicly-available awards. By way of comparison to an actual photograph, it should be possible to make out the people and what they are doing, but not the details of their expressions or the background relief.

5. Conclusion

This paper has provided an eclectic discussion of the use of empirical methods to research possible bias in investment arbitration. It has pointed out benefits and limitations of such methods, especially in examples of quantitative research on the topic. An underlying point has been that empirical research is an important component of scholarly efforts to understand adjudicative decision-making. However, legal doctrine and theory also maintain important roles in delineating the appropriate standards by which to conceptualize fairness and integrity in adjudication. In this respect, the role of institutional safeguards to ensure independence and impartiality has been emphasized. It has also been suggested that an awareness of the limitations of empirical research points to further reasons why doctrines of judicial independence are concerned with perceived as well as actual bias (e.g. *Locabail*, p. 471-2).

Empirical research in this area is at a fledgling stage. There are no dedicated qualitative studies on investment arbitration. Dezalay and Garth's work on international arbitration has relevance although there are limitations to its transferability to investment arbitration. Likewise, the initial quantitative work on bias in investment arbitration has in some cases been presented in ways that overstate or mis-state conclusions about actual bias. An alternative methodology of empirical research has been presented here as an option for supplementing other methodologies. It takes the form of a legal content analysis of the reasoning in awards. Legal content analysis may avoid important limitations of outcome-based research and, as such, an opportunity to triangulate data and methods (Berg, p. 4-6; Hall and Wright, p. 82-3). A further benefit is that it appears to provide a more value-added role for legal expertise in quantitative research. On the other hand, legal content analysis is complex and discretionary, and thus reliant for its validity on the rigour and transparency of the coding process.

Ideally, a range of methods and data sources would be used in complementary ways to elaborate theoretical expectations about bias in investment arbitration and seek to triangulate data, methods, and theories. That said, it will be a very long time before we achieve a mature body of empirical work on investment arbitration. Moreover, even when mature, the empirical work will not provide definitive proof of the presence or absence of actual bias in particular cases (Sisk and Heise, p. 746 and 794). Researchers must take great care to avoid unsupported claims about actual bias and to condition statements of the limitations of their research accordingly. Empirical methods should also not be used to give false assurances

that concerns about perceived bias in investment arbitration – arising from the absence of institutional safeguards otherwise present in the final resolution of public law – have been addressed based on statistical analysis of outcomes.

The criticisms of some empirical studies, as conveyed in this paper, are not meant as a wider criticism of the use of empirical methods. Rather, the aim here is to support an appropriately modest and refined use of empirical methods in order to make observations that would be impossible without the systematic collection of data. Used with care, empirical research can contribute to the understanding of social phenomena and support an evolution of doctrines and design of institutions toward the goal of fair and accurate decision-making.

References

- C.G.G. Aitken and F. Taroni, 2008, Fundamentals of statistical evidence – a primary for legal professionals . 12 *International Journal of Evidence and Proof*, 181.
- J. Baldwin and G. Davis, 2003, Empirical Research in Law in P. Cane and M. Tushnet, eds., *The Oxford Handbook of Legal Studies*. Oxford: Oxford University Press.
- B.L. Berg, 1998, *Qualitative Research Methods for the Social Sciences*, 3d edn. Boston: Allyn and Bacon.
- J.A.F. Costa, 30 May 2008, American and European Investment Arbitrators: A Single Culture?, Paper presented to Law and Society Association and the Canadian Law and Society Association conference on Placing Law, Montréal.
- Y. Dezalay and B. Garth, 1996, *Dealing in Virtue*. University of Chicago Press.
- C.R. Drahozal, 2006, Arbitration by the Numbers: The State of Empirical Research on International Commercial Arbitration. 22 *Arbitration International* 291.
- S.D. Franck, 2007, Empirically Evaluating Claims About Investment Treaty Arbitration, 80 *North Carolina Law Review* 1.
- S.D. Franck, 2009, Development and Outcomes of Investment Treaty Arbitration, 50 *Harvard International Law Journal* 435.
- B.G. Garth, 1997, Observations on an Uncomfortable Relationship: Civil Procedure and Empirical Research, 49 *Alabama Law Review* 103.
- M.A. Hall and R.F. Wright, 2008, Systemic Content Analysis of Judicial Opinions, 95 *California Law Review* 63.
- M. Heise, 2002, The Past, Present, and Future of Empirical Legal Scholarship: Judicial Decision Making and the New Empiricism, *University of Illinois Law Review* 819. International Court of Arbitration (of the International Chamber of Commerce), "Facts and Figures on ICC Arbitration – Statistical Reports", online: International Chamber of Commerce <http://www.iccwbo.org/court/arbitration/id5531/index.html> (last accessed 16 December 2010).
- M. Kent, 2005, 'Forced' vs. Compulsory Arbitration of Civil Rights Claims, 23 *Law & Inequality* 95.
- A.E. Krasuski, 2004, Mandatory Arbitration Agreements Do Not Belong in Nursing Home Contracts With Residents, 8 *DePaul Journal of Health Care Law* 263.
- K. Krippendorff, 1980, *Content Analysis: An Introduction to Its Methodology*. Newbury Park: Sage Publications.
- Lemire v Ukraine*, Award of 18 September 2000 (ICSID Additional Facility Rules), ICSID Case No. ARB(AF)/98/1.
- Locabail v Bayfield Properties*, 2000, QB 451 (Court of Appeal of England and Wales).
- N. Majeed, 2004, Investor-State Disputes and International Law: From the Far Side, 98 *ASIL Proceedings* 30.
- K.S. McArthur and P.A. Ormachea, 2009, International Investor-State Arbitration: An Empirical Analysis of ICSID Decisions on Jurisdiction, 28 *Review of Litigation* 559.
- R.K. Merton, 1968, *Social Theory and Social Structure*, revised and enlarged edn, New York: Free Press.
- K.A. Neuendorf, 2002, *Content Analysis Guidebook*. London: Sage Publications.

- K. Palm, 2006, Arbitration Clauses in Nursing Home Admission Agreements: Framing the Debate, 14 *Elder Law Journal* 453.
- R v Valente*, 1985, 2 SCR 673. Supreme Court of Canada..
- E. Rolph, E. Moller, and J.E. Rolph, 1997, Arbitration Agreements in Health Care: Myths and Reality, 60 *Law and Contemporary Problems* 153.
- G.C. Sisk and M. Heise, 2005, Judges and Ideology: Public and Academic Debates About Statistical Measures, 99 *Northwestern University Law Review* 743.
- S. Stemler, 2001, An Overview of Content Analysis, 7(17) *Practical Assessment, Research & Evaluation* (electronic journal).
- B. Trujillo, 2005, Patterns in a Complex System: An Empirical Study of Valuation in Business Bankruptcy Cases, 53 *UCLA Law Review* 357.
- A.L. Tyree, "Fact Content Analysis of Case Law: Methods and Limitations" (1981) 22 *Jurimetrics Journal* 1.
- P.H. Schuck, 1989, Why Don't Law Professors Do More Empirical Research, 39 *Journal of Legal Education* 323.
- W. Michael Tupman, 1986, Case Studies in the Jurisdiction of the International Centre for Settlement of Investment Disputes, 35 *International and Comparative Law Quarterly* 813.
- United States, *Report of the Advisory Committee on International Economic Policy Regarding the Model Bilateral Investment Treaty* (30 September 2009), Annex B (Particular Viewpoints of Subcommittee Members), Statement of Mark Kantor.
- G. Van Harten, 2007, *Investment Treaty Arbitration and Public Law* (OUP).
- G. Van Harten, June 2008, "A Case for an International Investment Court", Society of International Economic Law Inaugural Conference Working Paper.
- R.P. Wagner and L. Petherbridge, 2004, "Is the Federal Court Succeeding? An Empirical Assessment of Judicial Performance" .152 *University of Pennsylvania Law Review* 1105.
- J. Yackee, 2008, Conceptual Difficulties in the Empirical Study of Bilateral Investment Treaties, 33 *Brooklyn Journal of International Law* 405.
- J. Yackee, 2010, Do Bilateral Investment Treaties Promote Foreign Direct Investment? Some Hints from Alternative Evidence, University of Wisconsin Legal Studies Research Paper No. 1114.
- F. Yala, 2005, The Notion of 'Investment' in ICSID Case Law: A Drifting Jurisdictional Requirement?, 22 *Journal of International Arbitration* 105.

Appendix A

Investment treaty arbitration – issues template – June 2009

Issue 1: 'corporate person investor' – expansive approach	Issue 1: 'corporate person investor' – restrictive approach
<p>Flexible approach to claims by corporate persons, indicated by:</p> <p>(a) flexible approach to claims by foreign holding companies, including:</p> <ol style="list-style-type: none"> (1) prioritization of corporate form over control of the investment vehicle or rejection of an implied origin-of-capital test. (2) allowance of a claim by a foreign company that is owned and likely controlled by nationals of the host state. (3) allowance of a claim by a shareholder whose investment in the host state is owned by an intermediary company of a third state. <p>Note: allowance of a claim by a foreign company or natural person whose ownership of the investment extends through a chain of companies running from the host state to the home state, but does not extend into any third state or end in the host state, does <u>not</u> qualify as an expansive approach.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning. On the present issue, a treaty does not satisfy this threshold unless it expressly and specifically allows claims by foreign holding companies regardless of their ultimate ownership and control. A treaty that refers generally to 'direct and indirect' ownership of the investment (thus dealing with <i>Barcelona Traction</i>) does not satisfy this threshold.</p> <p>Note: a tribunal's reasoning on minority shareholder claims and interests may relate to the issue of 'corporate person investor' and to that of 'minority shareholder interest' (see below). A degree of coding overlap may therefore be unavoidable, leading in some cases to entry of an award under both issues. That said, a distinction should be maintained between these issues, where possible, on the basis that the</p>	<p>Restrictive approach to claims by corporate persons, indicated by:</p> <p>(a) restrictive approach to claims by foreign holding companies, including:</p> <ol style="list-style-type: none"> (1) use of veil-piercing or of an indirect control test or of a substantial connection text in order to preclude jurisdiction/ admissibility. (2) refusal of a claim by a foreign company that is owned and likely controlled by nationals of the host state. (3) refusal of a claim by a shareholder whose investment in the host state is owned by an intermediary company of a third state. <p>Note: refusal of a claim by a foreign company or natural person whose ownership of the investment extends through a chain of companies running from the host state to the home state, even where it does not extend into any third state or end in the host state, nevertheless <u>does</u> qualify as a restrictive approach.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning. On the present issue, a treaty does not satisfy this threshold unless it expressly and specifically precludes claims by foreign holding companies on the basis of their particular ownership and control or it expressly and specifically provides for application of the relevant test outlined in (1) above.</p>

<p>reasoning will relate to (1) the issue of 'corporate person investor' where the claim involves the use of a holding company, whether or not by a minority shareholder, so as to facilitate a claim where it might not otherwise be possible, or (2) the issue of 'minority shareholder interest' where the claim has been brought by a minority shareholder of the domestic investment, thus necessarily raising the issue of the nature and scope of the claimant's interest as a minority shareholder.</p>	
<p>Issue 2: 'natural person investor' – expansive approach</p>	<p>Issue 2: 'natural person investor' – restrictive approach</p>
<p>Flexible approach to claims by natural persons, indicated by:</p> <p>(a) allowance of claim against the only state of which the claimant is a citizen;</p> <p>OR</p> <p>(b) allowance of claim against a state of which the claimant is a citizen without confirmation that the citizenship upon which the claim is based is dominant and effective.</p> <p>OR</p> <p>(c) allowance of claim based on flexible application of the requirement for foreign nationality as customarily applied to natural persons.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning.</p>	<p>Restrictive approach to claims by natural persons, indicated by:</p> <p>(a) refusal of claim against the only state of which the claimant is a citizen;</p> <p>OR</p> <p>(b) refusal of claim against a state of which the claimant is a citizen following confirmation that the citizenship upon which the claim is based is not dominant and effective.</p> <p>OR</p> <p>(c) refusal of claim based on strict application of the requirement for foreign nationality as customarily applied to natural persons.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning.</p>

Issue 3: 'concept of investment' – expansive approach	Issue 3: 'concept of investment' – restrictive approach
<p>Flexible approach to concept of 'investment', indicated by:</p> <p>(a) where a claim is under the ICSID Convention, non-application of the <i>Fedax</i> criteria, including by focusing primarily on the definition of investment in the BIT or other investment treaty (a.k.a. subjective theory of investment under ICSID);</p> <p>OR</p> <p>(b) where the claim is under the ICSID Convention, liberal application of the <i>Fedax</i> criteria to include as 'investment' any activities that are stand-alone and that go beyond conventional FDI project activities, in line with 'the liberal movement, favourable to an extension of the jurisdiction of ICSID tribunals to every kind of economic rights' (Yala, p. 108);</p> <p>OR</p> <p>(c) whether or not the claim is under the ICSID Convention, rejection of any requirement for actual transfer of capital into the respondent state as a feature of investment (unless there are extenuating circumstances such as corrupt practices that block an investor from doing so);</p> <p>OR</p> <p>(d) whether or not the claim is under the ICSID Convention, inclusion of non-traditional categories of ownership within the concept of 'investment', e.g. sales office, market share, or corporate governance rights in contract where the asset is not part of conventional FDI project activities.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning. On the present issue, a treaty that expressly and clearly incorporates the relevant asset that is included by the tribunal as part of a non-traditional conception of 'investment' will satisfy this threshold unless the claim is an ICSID claim and either (a) or (b) above apply.</p>	<p>Restrictive approach to concept of 'investment', indicated by:</p> <p>(a) where a claim is under the ICSID Convention, strict application of the <i>Fedax</i> criteria (i.e. rejection of subjective theory focusing primarily on the definition of investment in the BIT or other investment treaty) to limit 'investment' to conventional FDI project activities or otherwise to deny claim;</p> <p>OR</p> <p>(b) whether or not the claim is under the ICSID Convention, adoption of a requirement for actual transfer of capital into the respondent state as a feature of investment;</p> <p>OR</p> <p>(c) whether or not the claim is under the ICSID convention, exclusion of non-traditional categories of ownership where not linked directly to conventional FDI project activities.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning.</p>

Issue 4: 'minority shareholder interest' – expansive approach	Issue 4: 'minority shareholder interest' – restrictive approach
<p>Flexible approach to claims by minority shareholders, indicated by:</p> <p>(a) allowance of claim by a minority shareholder without limiting the claim to the shareholder's interest in the value and disposition of the shares (as opposed to interests of the domestic firm as a whole).</p> <p>OR</p> <p>(b) allowance of claim by a minority shareholder where the treaty does not clearly and specifically allow it.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning. On (b) above, most treaties allow expressly and specifically claims by minority shareholders by defining investment to include ownership of stock or shares in a domestic company.</p>	<p>Restrictive approach to claims by minority shareholders, indicated by:</p> <p>(a) limitation of such a claim to the extent of the claimant's minority shareholder interest in the value and disposition of the shares;</p> <p>OR</p> <p>(a) preclusion of claim by a minority shareholder due to lack of control over the investment (in circumstances where, for example, the treaty does not define investment to include ownership of stock or shares in a domestic company).</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning.</p>
Issue 5: 'permissibility of investment' – expansive approach	Issue 5: 'permissibility of investment' – restrictive approach
<p>Flexible approach to approval/ permissibility of 'investment', indicated by:</p> <p>(a) Evident onus on state to show that investment was not affirmatively approved or was based on corrupt practices.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning.</p>	<p>Restrictive approach to approval/ permissibility of 'investment', indicated by:</p> <p>(a) Evident onus on investor to show that investment was affirmatively approved or was not based on corrupt practices.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning.</p>
Issue 6: 'parallel claims' – expansive approach	Issue 6: 'parallel claims' – restrictive approach
<p>Flexible approach to parallel claims, indicated by:</p> <p>(a) allowance of treaty claim in the face of:</p> <p>(1) a treaty-based duty to exhaust (or other claim-related condition such as a time-limited duty to pursue) local remedies which has clearly not been satisfied by the claimant, whether or not the claim relates to</p>	<p>Restrictive approach to parallel claims, indicated by:</p> <p>(a) refusal or delay (in order to permit resolution of aspects of the dispute in another forum) of a treaty claim in face of:</p> <p>(1) a [treaty-based] duty to exhaust (or other claim-related condition such as a time-limited duty to pursue) local remedies which has clearly not been</p>

<p>a contract.</p> <p>(2) a contractually-agreed dispute settlement clause that was consented to by the claimant or a closely related company, where the claim appears to relate to a contractual dispute but regardless of whether any claim has been brought in the contractually-agreed forum and regardless of whether the treaty claim is based on an umbrella clause.</p> <p>(3) a fork-in-road clause, where the claimant or a closely related company ('closely related' meaning a company owned and likely controlled by the investor) has brought a parallel claim via the relevant 'path' of the fork-in-road (i.e. in a domestic court or a domestic or international tribunal, according to the relevant path of the fork-in-road) and the claim arises from the same underlying factual dispute</p> <p>(4) an actual claim pursuant to another treaty [that can lead to a damages award in favour of the investor], arising from the same factual dispute.</p> <p>Note: allowance of treaty claims in any of these circumstances will typically be based on the strict application of the <i>lis pendens</i> or <i>res judicata</i> rule, using a civil law-based 'triple identity' test to require that (i) the parties, (ii) the cause of action, and (iii) the dispute all be identical before a parallel treaty claim can be stayed.</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning.</p>	<p>satisfied by the claimant, whether or not the claim relates to a contract.</p> <p>(2) a contractually-agreed dispute settlement clause, where the claim appears to relate to the contract but regardless of whether or not any claim has been brought in the contractually-agreed forum and regardless of whether the treaty claim is based on an umbrella clause.</p> <p>(3) a fork-in-road clause, where the claimant or a closely related company has brought a parallel claim via the relevant path of the fork-in-road and the claim arises from the same underlying factual dispute.</p> <p>(4) an actual claim pursuant to another treaty, arising from the same factual dispute.</p> <p>Note: refusal or delay of a treaty claim in any of these circumstances may be based on a flexible approach to <i>lis pendens</i> or <i>res judicata</i>, common law doctrines of issue estoppel or <i>forum non conveniens</i>, abuse of process, incorporation of an effective duty to resort to local remedies as a component of a substantive standard, or the rationale that a fork-in-road clause entails a choice by the claimant not only of forum but also of available remedies and causes of action (i.e. the investor (and closely related companies) can choose to bring either a treaty claim or a claim before a domestic court or domestic or international tribunal, but not both).</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning.</p>
---	---

<i>Issue 7: 'scope of MFN treatment' – expansive approach</i>	<i>Issue 7: 'scope of MFN treatment' – restrictive approach</i>
<p>Flexible approach to most-favoured-nation treatment, indicated by:</p> <p>(a) extension of MFN to non-substantive/ treatment-oriented provisions of other treaties (e.g. so as to include dispute settlement provisions of other treaties).</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning. On the present issue, this threshold is crossed where the treaty states explicitly that the MFN clause either does or does not apply to the dispute settlement (or other non-substantive/ treatment-oriented) provisions in other treaties. If the treaty makes a more general statement, for example that the MFN clause applies to 'all matters' or to matters of 'treatment' in other treaties, then this does <u>not</u> satisfy this threshold for an express and specific resolution of the issue.</p>	<p>Restrictive approach to most-favoured-nation treatment, indicated by:</p> <p>(a) refusal to extend MFN to non-substantive/ treatment-oriented provisions of other treaties (e.g. so as to include dispute settlement provisions of other treaties).</p> <p>Note: an award is non-classifiable where the issue is dealt with expressly and specifically in the text of the relevant investment treaty and this forms the basis of the tribunal's reasoning.</p>