ACKNOWLEDGEMENTS

The Asper Review of International Business and Trade Law gratefully acknowledges the support of the Asper Chair of International Business and Trade Law and the Canadian Credit Management Foundation. The Asper Review is equally appreciative of its patrons in the legal community. Appreciation is also extended to the plethora of anonymous internal and external referees who were generous with their valuable time and insights.

The Asper Chair of International Business and Trade Law was established in 1999 at the Faculty of Law of the University of Manitoba. Its mandate includes teaching, research and publication.

INFORMATION FOR CONTRIBUTORS

Manuscripts and communications should be directed to:

Editor-in-Chief
Asper Review of International Business and Trade Law
466 Robson Hall, Faculty of Law
Phone: 204.474.6159
University of Manitoba
Fax: 204.480.1084
Winnipeg, Manitoba R3T 2N2
E-mail: asperreview@gmail.com
All the articles in the Asper Review of International Business and Trade Law are externally refereed by independent academic experts after being rigorously peer reviewed by Manitoba faculty editors, as well as reviewed by student staff.
<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Frivolous Claims in the International Investment Regime: How CETA Expands the Range of Frivolous Claims that May be Curtailed in an Expedient Fashion</td>
<td>Ksenia Polonskaya</td>
</tr>
<tr>
<td>35</td>
<td>Citizen Enforcement of Environmental Regulation in Africa’s Extractive Industries</td>
<td>Evarestus Oshionebo</td>
</tr>
<tr>
<td>59</td>
<td>Buying Votes in the 21st Century: The Potential Use of Bitcoins and Blockchain Technology in Electronic Voting Reform</td>
<td>Brianna Bogucki</td>
</tr>
<tr>
<td>85</td>
<td>New Rules for a New World: How Technology and Globalization Shape Bankruptcy Venue Decisions</td>
<td>Laura N. Coordes</td>
</tr>
<tr>
<td>111</td>
<td>Rationalizing Defences to Enforcement under the New York Convention 1958</td>
<td>Dr. Chrisspas Nyombi &amp; Dr. Konstantinos Sillafis</td>
</tr>
<tr>
<td>131</td>
<td>Faltering Blocks in the Arguments against Unitary Taxation and the Formulary Apportionment Approach to Income Allocation</td>
<td>Alexander Ezenagu</td>
</tr>
<tr>
<td>163</td>
<td>The Standard of Corporate Liability for Crime: What Can Other Jurisdictions Learn from Canada?</td>
<td>Reem Radhi</td>
</tr>
<tr>
<td>181</td>
<td>Friends with Net Benefits: The Investment Canada Act and Stated Owned Enterprises</td>
<td>Robert Sroka</td>
</tr>
</tbody>
</table>
Frivolous Claims in the International Investment Regime: How CETA Expands the Range of Frivolous Claims that May be Curtailed in an Expedient Fashion

KSENIA POLONSKAYA

INTRODUCTION

A frivolous lawsuit is a claim “that is not legally tenable and as such is worthless”.¹ Such lawsuits have “no serious purpose or value or merit”.² Frivolous claims clog the judicial system, exhaust the resources of the

---


* PhD Candidate, Queen’s University (Faculty of Law). I am grateful to Professor Joshua Karton, Basil Alexander and Sarah Hamill for their feedback and encouragement that helped greatly in preparing this article.
other party to the dispute and, thus, bring the administration of justice into disrepute. Domestic legal systems empower courts to limit and deter frivolous lawsuits in a variety of ways. For example, Ontario’s *Rules of Civil Procedure* allow courts to dismiss an action “on the ground that the action is frivolous, vexatious, or is otherwise an abuse of process.” Among other options, a court could dismiss the proceedings, strike down the claim entirely or in part, or order costs to the party that brought a frivolous submission. The domestic courts can also impose sanctions on lawyers for bringing a frivolous submission.

However, not only domestic legal systems are vulnerable to such claims. In the system of investor-state arbitration (ISA), a mechanism whereby foreign investors may sue sovereign states before a tribunal of three arbitrators for property takings and other unfair treatment, the number of frivolous claims has also increased considerably. According to a report issued by the United Nations Conference on Trade and Development (UNCTAD), in several investment arbitrations foreign investors filed claims that were “without merit”, meaning that they had no real chance to succeed on the merits. Nonetheless, foreign investors are able to file and litigate such claims using

---

3 See *Thompson Bey v Iyer*, 2016 FC 990 at paras 8-9, which states that frivolous and vexatious litigation “clogs the courts and uses up the judicial resources [...] the court has a duty to control access to the judicial system by these types of litigants”.

4 Ontario, *Ontario Rules of Civil Procedure*, RRO 1990, Reg 194, RS at r 2.1.01(1), 2.1.01(1)(d), 2.1.02(1).

5 Paul M Perell & John W Morden, *The Law of Civil Procedure in Ontario* (Markham: Lexis Nexis, 2010) at 354. The authors explain that “the court may strike out a pleading [...] that is frivolous, vexatious or abuse of process”.

6 Sanford Levinson, “Frivolous Cases: Do Lawyers Really Know Anything at All” (1986) 24 Osgoode Hall LJ 353 at 359, 361-62. Levinson points out that the “Ontario Barrister’s oath which requires the vow that the barrister will not, among other things, promote suits upon frivolous pretences”. The article pinpoints that the courts can order costs against the solicitor in cases where one advanced a “frivolous position”. Levinson refers to M. Gold who apparently argued “that it may be part of a solicitor’s duty to the court that he only raise points which are fairly arguable. To take a thoroughly bad and unmeritorious point which results in extra costs to the parties may justify a costs order against a solicitor”. See M Gold, “The Court’s Authority to Award Costs Against Lawyers”, cited in E Gertner, ed, *Studies in Civil Procedure* (Toronto: Butterworths, 1979) at 79.


8 Ibid at para 7. The Report stated that “there have been fears about frivolous or vexatious claims that could inhibit legitimate regulatory action by Government”.

ISA with the result that states often find themselves dragged into the lengthy and costly process of arbitration.9

The problem of frivolous claims in the international investment regime first entered the spotlight during the negotiations of the Convention on the Settlement of Investment Disputes Between States and Nationals of Other States, also known as the ICSID Convention.10 During negotiations, a number of diplomatic delegates pointed out that the Convention needed a mechanism that would prevent foreign investors from drawing states into vexatious proceedings.11 At the time, Aron Broaches, a former President of the World Bank, reassured those delegates that the ICSID Secretary General could perform the function of dissuading claimants from bringing such claims.12 As history shows, the Secretary General of the ICSID Arbitration Centre has never performed such a function and has only succeeded in preventing the filing of claims manifestly outside of the jurisdiction of the ICSID Arbitration Centre.13

Currently, some international treaties for protecting foreign investment,14 in addition to the ICSID Arbitration Rules,15 do provide mechanisms for

---

9 See Tania Voon, Andrew Mitchell & James Munro, “Parting Ways: The Impact of Investor Rights on Mutual Termination of Investment Treaties” (2014) 29:2 ICSID Rev 451 at 452. In that work, the authors recognize the impact of particular disputes on states and on the regime [Voon].


13 Ibid.

14 See The Dominican Republic, Central America, United States Free Trade Agreement, art 10.20, online: <http://bit.ly/2vOufSi> [DR-CAFTA].

15 The ICSID Convention is in focus because the ICSID Arbitration Centre is the most frequently used venue for resolving disputes between investors and states. In addition, the ICSID Centre was the first centre to introduce a mechanism for the early dismissal of frivolous claims. Finally, the ICSID Tribunals developed a consistent body of case-law on meritless claims. See Parra, supra note 12 at 9, 16, 250. Note: it is an interesting issue whether the early
Asper Review

[Vol. XVII]

curtailing frivolous claims on an expedited basis. The recently concluded *European Union-Canada Free Trade Agreement* (CETA)\(^{16}\) also borrows from national rules of civil procedure in commitment to limiting frivolous claims. As this article will show, CETA uniquely builds on the existing experience of dismissing frivolous claims *in limine*.

This article submits that the existing mechanisms for early curtailment are insufficient and examines CETA’s approach to dismissing frivolous claims. In particular, this article argues that CETA provides for important institutional changes in the model of investment arbitration such as an appeals mechanism as well as procedures concerning remuneration and appointment of arbitrators that are likely to make a big difference in the rate of frivolous claim dismissals early in investment proceedings. So far, the topic of frivolous claims in investment arbitration has received only limited coverage.\(^{17}\) In particular, this work focused on frivolous claims in the context of indirect expropriation. Michele Potestà and Marija Sobat have also examined the application of the ICSID Rules on frivolous claims by ISA tribunals for

curtailing frivolous claims.\textsuperscript{18} Eric De Brabandere has focused on reviewing the early dismissal mechanism in the ICSID Rules as well.\textsuperscript{19} This article builds on the existing literature on frivolous claims in ISA by specifically addressing the current standard for early dismissal of frivolous claims and contributes to the existing “dialogue of authors”\textsuperscript{20} by making the case that CETA’s approach is superior to existing mechanisms.

The International Investment Regime: A Brief Overview

What is the International Investment Regime?

The international regime for the protection of foreign investment is embodied in more than 3000 International Investment Agreements (IIAs) which are primarily enforced through a mechanism known as Investor-State Arbitration (ISA). IIAs are of two types: (1) Free Trade Agreements (FTAs) and (2) Bilateral Investment Treaties (BITs).\textsuperscript{21} While BITs contain only investment rules,\textsuperscript{22} FTAs are “a full package deal” that include multiple chapters, among which only one deals solely with investment.\textsuperscript{23} IIAs are instruments subject to international law, and as such, the obligations created by them are to be interpreted and applied in light of public international law.

IIAs incorporate a variety of more-or-less standardized substantive investment rules, such as Fair and Equitable Treatment (FET),\textsuperscript{24} Full Security

\textsuperscript{18} Michele Potestà & Marija Sobat, “Frivolous Claims in International Adjudication: A Study of ICSID Rule 41(5) and of Procedures of Other Courts and Tribunals to Dismiss Claims Summarily” (2012) 3:1 J Intl Dispute Settlement 137.


\textsuperscript{20} The “dialogue of authors” is a term of art that I borrow from Abraham Drassinower. See Abraham Drassinower, What’s Wrong with Copying (Cambridge: Harvard University Press, 2015). In that book on copyright, Drassinower suggests that each and every author is a participant in such dialogue, providing a voice that contributes to the scholarly discussion.

\textsuperscript{21} Rudolf Dolzer & Margrete Stevens, Bilateral Investment Treaties (Leiden: Martinus Nijhoff Publishers, 1995) [Dolzer & Stevens].

\textsuperscript{22} Ibid.


\textsuperscript{24} FET is a standard of treatment with a highly contested meaning in international law. In broad terms, the standard requires a host state to treat a foreign investor “fairly” and “equitably”; however, most treaties often do not define the meaning of these terms. IIAs...
and Protection (FSP), Expropriation, Most-Favoured Nation Treatment (MFN), and umbrella clauses. The substantive investment standards (such as expropriation or FET) aim to protect foreign investors and their investments against various non-business risks while in the host state.


The FSP standard, like FET, is broad and often undefined. It is typically interpreted to include two elements: legal protection and physical protection. Some tribunals require FSP to be guaranteed only to foreign investments, while others require FSP to be applied both to foreign investors and foreign investments. For example, CETA requires a host state only to guarantee physical protection to foreign investors and foreign investment. For more, see Matthias Herdegen, Principles of International Economic Law (Oxford: Oxford University Press, 2016) at 469.

Expropriation is another international standard that IIAs accord to foreign investors. In short, under this standard, a state may not expropriate (i.e. “take” or “seize”) property of a foreign investor except under due process of law and with prompt and adequate compensation. Expropriation may be direct or indirect. Direct expropriation involves the actual taking or destruction of property by a state or its regulatory bodies. Indirect expropriation is a far broader and more controversial concept, it can encompass any governmental action or series of actions that impairs the value of an investor’s investment. See Dolzer & Stevens, supra note 21; Andrew Newcombe, Regulatory Expropriation, Investment Protection and International Law: When is Government Regulation Expropriatory and When Should the Compensation be Paid? (University of Toronto, 1999), online: <http://www.italaw.com/documents/RegulatoryExpropriation.pdf>.


Rudolf Dolzer & Christoph Schreuer, Principles of International Investment Law (Oxford: Oxford University Press, 2008) at 153-162 [Dolzer & Schreuer]. On umbrella clauses, see Christoph Schreuer, “Travelling the BIT Route of Waiting Periods, Umbrella Clauses and Forks in the Road” (2004) 5:2 J World Investment & Trade 249 at 250. Schreuer explains that “umbrella clauses have been added to some BITs to provide additional protection to investors beyond the traditional international standards. They are often referred to as ‘umbrella clauses’ because they put contractual commitments under the BIT’s protective umbrella. They add the compliance with investment contracts, or other undertakings of the host State, to the BIT’s substantive standards. In this way, a violation of such a contract becomes a violation of the BIT.”
However, such protections would be insufficient without a mechanism to enforce the standards under IIAs in case of non-compliance by the host state. ISA has become such an enforcement mechanism. ISA emerged to protect the property rights of foreign investors against the unilateral conduct of sovereign states (*ius imperium*).

Before further discussing the institutional design of ISA and exploring why frivolous claims are particularly problematic for it, it is essential to address the arguments in favour of why the international investment regime is valuable for foreign investors, individual states, and the international community as a whole.

Historical context helps to understand the significance of IIAs and ISA for its stakeholders. In the post-World War II environment, the international flow of foreign investment was crucial to the revival of the world economy. In capital-importing (i.e. developing) states, foreign investors often encountered various non-business risks, such as takings of property, discrimination, or lack of due process in local courts and administrative agencies. After being subjected to this mistreatment, foreign investors often asked their home states to interfere in order to protect their assets abroad. In some instances, the investors’ home states (primarily capital-exporting, developed states) engaged in diplomatic negotiations, lump sum agreements or inter-state dispute

---

29 Kate Miles, *The Origins of International Investment Law. Empire, Environment and the Safeguard of Capital* (Cambridge: Cambridge University Press, 2013) at 76. See also Asha Kaushal, “Revisiting History: How the Past Matters for the Present Backlash Against the Foreign Investment Regime” (2009) 50 Harvard Intl LJ 491 at 499 [Kaushal]. Particularly, Kaushal suggests that foreign investment protection “finds its roots in the post-World War II era [...] In the 1950s, newly independent states asserted control over their natural resources, using their national law to regulate foreign investments and investors. Nationalizations and expropriations of foreign concessions proliferated in Iran, Libya, Egypt, Cuba, Chile, and Venezuela, to name a few”.

30 Kaushal, *ibid* at 498 (suggesting that foreign investors often had to rely on the means of diplomatic protection); see also Stephan Wilske & Martin Raibe, “The Arbitrator as Guardian of International Public Policy? Should Arbitrators Go Beyond Solving Legal Issues?” in Catherine A Rogers & Roger P Alford, eds, *The Future of Investment Arbitration* (Oxford: Oxford University Press, 2009) 249 at 251 (arguing that diplomatic protection and military force akin to gunboat diplomacy were the most common ways to settle investment disputes).

31 Diplomatic negotiations, however, were not always successful. For instance, the British Government negotiated with the Soviet Government a sum to be compensated to Lena Goldfields Inc., a British company whose concession rights were expropriated. The diplomatic pressure continued for more than a decade, but without any success. See Arthur Nussbaum, “Arbitration Between the Lena Goldfields Ltd. and the Soviet Government” (1950) 36 Cornell L Rev 31 at 34.
settlement.\textsuperscript{33} In other instances, some home states did not shy away from even bringing disputes over foreign property into the battlefield.\textsuperscript{34} For example, during the Suez and Abadan crises, Britain and France sent their military forces to bring about favourable resolutions after the governments of Egypt and Iran respectively nationalized assets of British and French nationals.\textsuperscript{35} Both incidents resulted in the escalation of international tension and resulted in serious political crises.\textsuperscript{36}

Of course, such conflicts over foreign property threatened to interrupt the international flow of foreign capital needed to rebuild the post-war economy.\textsuperscript{37} In addition, such conflicts undermined the two foundational principles of the United Nations (UN), namely friendly cooperation among nations and peaceful resolution of disputes.\textsuperscript{38} To safeguard foreign nationals and depoliticize disputes over foreign property,\textsuperscript{39} some states and

\textsuperscript{32} For example, see Richard B Lillich & Burns H Weston, “Lump Sum Agreements: Their Continuing Contribution to the Law of International Claims” (1988) 82:1 AJIL 69 at 74. Lump-sum agreements guaranteed a certain sum to be compensated to the foreign investor independently of their actual losses. In some instances, the sums under such agreements were significantly less than the losses of foreign investors and of course did not meet the “prompt, adequate and effective compensation” standard. Quite often, states agreed to the conditions of the lump-sum agreements for political or economic reasons. As a result, foreign investors were strategically disadvantaged.

\textsuperscript{33} For example, in the Barcelona Traction case, Canada did not step in to represent its Canadian investor. The International Court of Justice concluded that a state has a right but not an obligation to exercise diplomatic protection. See Case concerning Barcelona Traction, Light and Power Company Limited (Belgium v Spain) (New Application, 1962), [1970] ICJ Rep 3 [Barcelona Traction].

\textsuperscript{34} James Cable, Intervention at Abadan: Plan Buccaneer (Hague: Springer, 1991) at 9.


\textsuperscript{39} José E Alvarez, The Public International Law Regime Governing International Investment (Martinus Nijhoff Publishers, 2011) at 59 [Alvarez]. Note: it is necessary to account for David Schneiderman’s work, which persuasively argued that the depoliticization of investment disputes has failed. See David Schneiderman, Resisting Economic Globalization: Critical Theory and
international organizations began enacting IIAs that included ISA mechanisms.

Institutional Design of Investment Arbitration

This section briefly discusses the conventional model of investment arbitration and its critiques. The next subsection addresses the structural changes of CETA with regards to the mechanism of investment arbitration. These subsections are important for the overall examination of why the current mechanisms for curtailing frivolous claims in limine are not sufficient.

i. The Conventional Model of Investor-State Arbitration

ISA tribunals are convened ad hoc for each dispute. ISA tribunals are typically composed of three arbitrators. One is appointed by the state, another by the investor, and the president of the Tribunal is appointed on the agreement of both parties. The arbitrators do not hold any form of tenure; they are separately appointed for each arbitration and receive remuneration directly from the parties to the dispute. This structure of arbitral tribunals and the procedure of arbitral appointments can be explained—and is typically justified—by the need to guarantee the neutrality of adjudicators and to secure the tribunals from any biases either in favour of the investor or the state.
Before submitting a claim, a foreign investor must satisfy the requirements set out in the IIA that pertains to the scope of its application: the definition of investor (i.e. jurisdiction *ratione personae* of the ISA tribunal) and the definition of investment (i.e. jurisdiction *ratione materiae*).\(^\text{43}\) Most often, IIAs include broad formulations of these definitions. Under the definition of investment, IIAs typically include any kind of asset (e.g. tangible and intangible property).\(^\text{44}\) The definition of investor requires the investor to possess a particular nationality in order to gain access to the protections guaranteed under the specific IIA—normally, the investor must be a national of a state that has ratified the IIA other than the host state where it made the investment.\(^\text{45}\)

### ii. Critiques of the Conventional ISA Model

Recently, the conventional ISA model has been subjected to a wide range of criticisms, wide enough that a survey of all the critiques of international investment regime\(^\text{46}\) is well beyond the scope of this article. Accordingly, this article only focuses on the three critiques that specifically relate to the discussion on why frivolous claims in investment arbitration are particularly problematic.

First, there is an ongoing debate regarding the extent to which ISA arbitration limits the ability of states to regulate domestically.\(^\text{47}\) It is clear that states fear that ISA could become pervasive in their regulatory space.\(^\text{48}\) Particularly, there is a persisting concern that foreign investors will threaten to bring a claim against a state in order to deter that state’s proposed welfare

---


\(^{45}\) Douglas, *supra* note 43 at 286.

\(^{46}\) Alvarez, *supra* note 39 at 75-94. Alvarez provides perhaps one of the most elaborate classifications of criticisms of the investment regime.


\(^{48}\) The newly negotiated IIAs illustrate this fear. For instance, CETA reaffirms the right of states to regulate. See CETA, *supra* note 16, Chapter 8 at Preamble, arts 8.9, 8.10. We can see similar pro-regulatory statements in the TPP and the TTIP.
It is necessary to mention that such a claim does not have to be viable in order to proceed to a full hearing on the merits, which will, of course, mean high costs for both parties. To avoid such costs, some states may therefore shy away from pursuing a particular policy—resulting in a so-called “chilling effect.”

Second, the investment rules of IIAs, as interpreted by ISA tribunals, have taken on meanings broader than many say they were originally intended to have. In particular, ISA tribunals have developed interpretations of the FET standard that are broader-than states intended. Some tribunals include the element of legitimate expectations of the investor when applying FET, while other tribunals contend that this element is outside of the scope of FET.

Kate Miles examines an attempt by the Indonesian Government to enact a forestry law that “affected more than 150 mining companies”. Miles submits that “the companies threatened the state with international investment arbitration to ensure inapplicability of the new legislation [...] the Government subsequently enacted regulation that effectively exempted from the new forestry law the mining companies”. Reportedly, the Indonesian Government refused to enact the forestry law soon after the Government received advice that it may face ISA claims “in the realm of US$31 billion if the forestry laws remained effective”. Accordingly, for Miles, the regulatory chill “has the potential to frustrate initiatives designed to implement national and global environmental objectives”. See Kate Miles, The Origins of International Investment Law: Empire, Environment and the Safeguard of Capital (Cambridge: Cambridge University Press, 2013) at 183.

Recently, the international investment regime has already been subject to such pressures, particularly from national stakeholders in developed states. For instance, in Germany, the UK and Australia, there were serious parliamentary debates regarding whether these states should at all participate in ISA. For example, see Luke Nottage, Investor-State Arbitration Policy and Practice in Australia (CIGI Paper No. 6 June 2016); Leon E Trakman, “Investor-State Arbitration: Evaluating Australia’s Evolving Position” (2014) 15:1 J World Investment & Trade 152; Marc Burgenberg, A History of Investment Arbitration and Investor-State Dispute Settlement in Germany, (CIGI Paper No.12 October 2016), online: <https://www.cigionline.org/publications/history-investment-arbitration-and-investor-state-dispute-settlement-germany>. In Canada, some prominent academics signed a petition that condemns the international investment regime for its pro-investor orientation. See Gus Van Harten et al, Public Statement on the International Investment Regimes - 31 August 2010, online: <http://www.osgoode.yorku.ca/public-statement-international-investment-regime-31-august-2010/> [Public Statement on ISA].

Third, ISA tribunals have produced inconsistent decisions when applying investment rules, even in analogous factual circumstances.\(^5\) This inconsistency threatens to undermine the credibility of investment arbitration as a mechanism for resolving disputes. In particular, the inconsistency is problematic in the context of frivolous claims because an adjudicator has no clear understanding of how to consistently interpret and apply legal norms which results in an inability to measure whether a claim can succeed on the merits.\(^5\)

This generation of variable outcomes contributes to a perception among states that ISA could impact their freedom of regulatory action. However, inconsistent decisions and variable outcomes do not mean that the ISA tribunals are random or unprincipled in their decision-making.\(^5\) At a minimum, ISA tribunals tend to follow previous decisions to crystalize the scope and meaning of investment rules, and it appears that the gradual accretion of precedents has made outcomes more consistent over time. However, ISA tribunals’ decisions are not binding on any other tribunals or

\(^{5}\) The CMS v Argentina and LG&E v Argentina cases are illustrative. The tribunal in LG&E v Argentina (with nearly “identical” facts to CMS v Argentina) came to an opposite finding [or outcome] than the CMS conclusion in interpreting “state of necessity”. The state of necessity is a doctrine in international law that a state can invoke as defense to justify a regulatory measure; if the defense is successful, a state is not required to pay a compensation. While the tribunals came to the opposite conclusion, they did so applying the same investment rules to nearly identical facts. ISA tribunals can thus be inconsistent in their application and interpretation of international law. Such inconsistency obviously makes it difficult to define trends, and it defeats any expectations on what claims are viable. In addition, it makes it far more difficult to evaluate the risks for states associated with litigating these claims. See LG&E Energy Corp, LG&E Capital Corp, and LG&E International Inc v Argentina (2006), 46 ILM 36 (International Centre for Settlement of Investment Disputes) at para 139; CMS Gas Transmission Company v The Republic of Argentina (2005), 42 ILM 788, 44 ILM 1205 (International Centre for Settlement of Investment Disputes). See also Todd Weiler, International Investment Law and Arbitration: Leading Cases from the ICSID, NAFTA, Bilateral Treaties and Customary International Law (Cameron May, 2005); Michael Waibel, “Two Worlds of Necessity in ICSID Arbitration: CMS and LG&E” (2007) 20 Leiden J Intl L 637.


\(^{5}\) Gabrielle Kaufmann-Kohler, “Arbitral Precedent: Dream, Necessity or Excuse?” (2007) 23:3 Arb Intl 357 at 357. Kaufmann-Kohler argues that “international arbitration lacks a doctrine of precedent, at least as it is formulated in the common-law system. Regardless, arbitrators increasingly appear to refer to, discuss and rely on earlier cases”.\(^5\)
on national courts, which means that the ISA tribunals could step away from a particular pattern in interpreting IIAs at any time.\textsuperscript{55}

iii. CETA’s Model of a Quasi-Investment Court

In response to these critiques, CETA departs in ways both subtle and dramatic from the conventional ISA mechanism. Under CETA, a claimant must still satisfy the traditional jurisdictional requirements to bring a claim. CETA, however, changes the procedure of the appointment of arbitrators, their method of remuneration, and the ethical requirements incumbent upon them.\textsuperscript{56}

More drastically, CETA replaces \textit{ad hoc} arbitral tribunals with a permanent “standing tribunal”.\textsuperscript{57} According to article 8.27 paragraph 2, the CETA Joint Committee, which is composed of the representatives of the EU and Canada, will appoint 15 arbitrators for renewable 5-year terms.\textsuperscript{58} Among these 15 adjudicators, there will be 5 nationals of the EU, 5 nationals of Canada and 5 nationals of third countries. Individual cases are to be resolved by a panel of three arbitrators.\textsuperscript{59} In contrast to the conventional approach (where the parties to the arbitration pick the arbitrators), the president of the CETA tribunal (who must be a national of a third country) selects the arbitrators for the particular dispute.\textsuperscript{60} The most peculiar feature of the CETA’s model is that the arbitrators will be paid a retainer to keep themselves available for adjudicating disputes.

\textsuperscript{55} Andrea K Bjorklund, “Investment Treaty Arbitral Decisions as Jurisprudence Constante” in Colin Picker, Isabella D Bunn & Douglas Arner, eds, \textit{International Economic Law: The State and Future of the Discipline} (Oxford: Bloomsbury Publishing, 2008) 265 at 268 [Bjorkund]. Bjorklund explains that the “tribunals can and do refer to decisions of other tribunals […] they do not do so, however, in the guise of looking to previous decisions as binding precedent, and indeed jealously guard their autonomy. Thus, tribunals frequently disavow any formal obligation to review prior case law”.

\textsuperscript{56} CETA, \textit{supra} note 16, art 8.27.


\textsuperscript{58} CETA, \textit{supra} note 16, art 8.27 at para 2.

\textsuperscript{59} \textit{Ibid}, art 8.27 at para 5.

\textsuperscript{60} \textit{Ibid}.
Perhaps the most significant change CETA makes is that it authorizes parties to bring an appeal to a so-called appellate tribunal.61 The appellate tribunal has the power to review awards, including to address errors “in the application or interpretation of applicable law”.62 This power means that the CETA appellate tribunal will supervise the application and interpretation of investment rules, which may result in greater consistency in arbitral practice under the treaty.63

Why Frivolous Claims Are Particularly Problematic in the International Investment Regime

In the domestic context, frivolous claims are problematic because such claims clog the judicial system, disrupt the proper administration of justice and impose unnecessary costs on innocent defendants. The current international investment regime, however, has no unified judicial system. Instead, conventional ISA tribunals function on an ad hoc basis where the parties of the arbitration bear the costs of the proceedings. So, why are frivolous claims a problem that requires a mechanism for early dismissal in the arbitral context? Although the institutional organization of any domestic judicial system is different from the international investment regime, frivolous claims are still problematic in investment arbitration because they (a) impose unreasonable costs upon states and (b) threaten to undermine the credibility of the international investment regime itself.

i. Costs

Frivolous claims can impose high costs. In the domestic context, Judge Easterbrook perfectly summarizes why frivolous claims are problematic: “suits are easy to file and hard to defend. Litigation gives lawyers opportunities to impose on their adversaries’ costs much greater than they impose on their.

---

61 Ibid, art 8.28.
63 In the international economic law context, the WTO Appeal Body gained scholarly praise for ensuring consistent case-practice in trade disputes. The constituency of arbitral practice of course serves for reinforcing the rule of law as the results become more predictable. In the context of frivolous claims, predictability in law’s application plays an important role because it gives the tribunal grounds to assert that this particular claim has no chance to succeed on the merits due to the pattern in how legal standards have been interpreted and applied. For more, see Peter Van den Bossche & Werner Zdouc, The Law and Policy of the World Trade Organization (Cambridge: Cambridge University Press, 2013) at 205.
own clients [...] litigation becomes a predatory instrument rather than a method of resolving honest disputes”.

Judge Easterbrook’s criticism of frivolous lawsuits applies to the ISA context. The arbitral process is claimed to be prompt and cost-efficient, but it does not always deliver on that promise. In fact, ISA arbitration can last four years or longer, and the costs of defending a claim can exceed 10 million dollars, which accounts only for arbitrators’ and lawyers’ fees and not the additional costs for experts, witnesses, investigations, and other requirements. For example, Poland and the Czech Republic respectively spent US$6 million and US$10 million to cover their costs in defending a single arbitration. Moreover, as the case of the Philippines shows, the costs of participating in investment arbitration can be burdensome for societies. The Transnational Institute reports, “the Philippines government spent US$58 million to defend two cases against German airport operator Fraport – the equivalent of the salaries of 12,500 teachers for 1 year, vaccination for 3.8 million children against diseases such as TB, diphtheria, tetanus, polio; or the building of 2 new airports”.

Arbitration also requires specialist legal expertise to strategically evaluate a claim. Most states (including some developed states) lack such expertise in-

---

65 Daniel Behn, “Legitimacy, Evolution, and Growth in Investment Treaty Arbitration: Empirically Evaluating the State-of-Art” (2015) 46 Geo J Intl L 364 at 376. Daniel Behn analysed a data-set of 77 investment arbitration cases. Behn explains that “cases in the data set take an average of four years to resolve [...] the longest case [...] took 129 months to resolve”.
66 UNCTAD Trade & Development, supra note 7 at para 14.
67 Ibid.
68 Pia Eberhardt & Cecilia Olivet, Profiting From Injustice (Brussels: Corporate Europe Observatory and the Transnational Institute, 2012) at 15, online: <https://www.tni.org/files/download/profitingfrominjustice.pdf> [Eberhardt].
69 The importance of legal expertise is crucial, as the contrasting cases of Philip Morris v Australia and Shell v Nicaragua demonstrate. In Philip Morris, Australia approached the arbitration strategically and filed for bifurcation of the proceedings, a strategy that proved successful in helping Australia win the case on jurisdictional grounds while minimizing costs. In Shell, Nicaraguan officials admitted that they lacked expertise on matters relating to ISA. The case was settled. See Shell Brands International AG and Shell Nicaragua SA v Republic of Nicaragua (2006), (International Centre for Settlement of Investment Disputes Case No. ARB/06/14), online: <http://investmentpolicyhubunctad.org/ISDS/Details/231>; Philip Morris v Australia, Award on Jurisdiction and Admissibility, (2015) PCA Case No. 2012-12, online: <http://www.italaw.com/sites/default/files/case-documents/italaw7303_0.pdf>; Jorun
house, but many states also lack the resources to hire expert outside counsel. The arbitration of Phillip Morris v Australia is illustrative for such a strategic approach to conducting arbitral proceedings. Australia’s success in having Philip Morris’ claim struck out as abuse of process rested on the fact it was able to bifurcate the investment proceedings. For the purposes of our inquiry, it is important to note that in addition to the Australian government solicitors, the respondents’ team included at least three law firms as counsel. Such legal advice is, of course, costly. According to the Sydney Morning Herald, the overall cost to Australia in that arbitration amounted to $50 million.

Of course, under international law all states are considered equal sovereigns. However, the broader social and economic reality shows that states are functionally not equal in their capacity to resist the pressure of costly arbitral claims and proceedings. Developing states may not possess “the technical expertise and even institutional capacity” to respond to the pressure of an ISA action (so-called “challenges of capacity”). A lack of capacity to uphold defences or actively engage in arbitration could result in an unfavourable settlement or cancellation of national regulatory actions even if the state would ultimately be vindicated had the arbitral process proceeded to its conclusion. The costs may become particularly burdensome when a single developing state has to face multiple arbitrations in the context of a national emergency.


70 The Essex Court Chambers, Eleven Wentworth Chambers, Selborne Wentworth Chambers all participated in the proceedings as legal counsel for Australia. See Philip Morris, ibid.


72 Some states have to use third-party funding and expertise to carry on a defence. For example, Uruguay was able to survive a legal battle against tobacco companies only with the financial and expert support of Michael Bloomberg and other anti-tobacco advocates. See Heather Wipflti, The Global War on Tobacco: Mapping the World’s First Public Health Treaty (Baltimore: JHU Press, 2015) at 132.


74 For example, Argentina is facing, or has faced, 37 claims arising out of actions taken to respond to its financial crises. See UNCTAD Trade & Development, supra note 7 at para 9; Schneiderman, supra note 39 at 4344.
For states with smaller economies, the threats of investment claims (even if such claims potentially lack merit) could be challenging to withstand. Gus Van Harten and Dayna Nadine Scott conducted an empirical study regarding the influence of ISA claims on governmental decision-making.\textsuperscript{75} They found that claims or even threats of claims led to changes in government policies, at least in Ontario. Their study provides support to the underlying concern of this article that the threat of meritless claims may have chilling effects.\textsuperscript{76} The threat to file an ISA claim may add pressure upon a state and provide an investor with leverage.

Developed states with larger economies may also suffer from the costs that frivolous claims impose. When a government has to uphold a defence against an apparently frivolous submission, national stakeholders could demand an explanation as to why tax dollars are spent for engaging in frivolous proceedings.\textsuperscript{77} The issue of costs thus results in another problem, which is that frivolous claims may ultimately undermine the credibility of ISA as a mechanism for resolving disputes between investors and states.

\textit{ii. Credibility of Investor-State Arbitration}

Under the international investment regime, only foreign investors have a right to bring a claim against states, not \textit{vice versa}. Hence, such a structure potentially empowers foreign investors to drag states into the process of arbitration. Notably, foreign investors, who are often multinational corporations, may submit their claims not to vindicate their property rights but instead to achieve other strategic purposes.\textsuperscript{78} In domestic settings, corporate actors sometimes launch so-called “sham litigation” by bringing lawsuits without merit to drive their competitors from the marketplace.\textsuperscript{79}

---


\textsuperscript{76} Ibid.


\textsuperscript{79} In the US Congress, one of the witnesses called sham litigation “an anticompetitive practice”, “a deadly weapon” that “results in extorting the tax payer’s money by using the threat
conduct may result in significant social costs because such litigation has anticompetitive effects. For example, when other companies cannot access the marketplace because they fear litigation, consumers suffer from the lack of competition, which allows monopolists to charge higher prices. Of course, sham litigation (like other kinds of frivolous litigation) is also problematic because it clogs the legal system and prevents the proper administration of justice.  

In addition, it is entirely plausible that a board of directors may want to demonstrate to their shareholders that they will pursue all possible avenues of litigation, including ISA. Under this scenario, a foreign investor may submit a non-viable claim and nonetheless continue to arbitrate on the jurisdiction and merit stages. If there is no procedure available to curtail such claims early in the process, states may be ‘dragged’ into continuous proceedings that involve serious costs.

The cost of frivolous claims is hardly the only reason why states may ultimately withdraw from the international investment regime. The cost of such claims, however, may be one of the contributive factors why states may consider a shift to state-to-state dispute settlement or to carving out particular industries from investment treaties. Tanya Voon, Andrew Mitchell and James Munro argue that even a single claim may be sufficient to generate backlash against the international investment regime. The precedents are well known; when Philip Morris launched its claim challenging Australia’s tobacco plain packaging regulations, Julia Gillard, the Prime Minister of Australia, made a public declaration that her government would reconsider its approach to investment arbitration in favour of a state-to-state dispute settlement system for resolving disputes that concern Australia. During the negotiations of the
Trans-Pacific Partnership (TPP), it was no coincidence that Australia insisted tobacco be carved out from the subject-matter scope of the treaty.\footnote{Cathleen Cimino-Isaacs, Trans-Pacific Partnership: An Assessment (London: Jeffrey J Schott Peterson Institute for International Economics, 2016) at 115.}

**Addressing Frivolous Claims in the International Investment Regime**

This section surveys arbitral practice of the ISA tribunals on dismissing frivolous claims and introduces the reader to the standards the investment tribunals apply to curtail frivolous submissions early in the proceedings. Accordingly, this section proceeds in two parts.

Firstly, this section discusses the standard “manifestly without legal merit” that the ISA Tribunals applied before CETA. Here, the analysis is focused on the practice of the ICSID Arbitration Centre, which remains one of the most frequently used venues for resolving disputes between foreign investors and states.

Secondly, this section discusses the novel standard that CETA introduces for limiting frivolous claims that is “claims in which the award in favour of claimant cannot be made”. There is no arbitral practice under CETA yet. Accordingly, CETA’s provisions on frivolous claims will be discussed in the context of the *Pac Rim v El Salvador* decision rendered under DR-CAFTA,\footnote{DR-CAFTA, supra note 14 at 10.20(4).} a treaty that incorporates similar wording on frivolous claims to the standard found in CETA.

**Claims Must Satisfy the High Threshold of “Manifestly Without Legal Merit”**

Antonio Parra (the former ICSID Deputy Secretary-General) explained that “recurring complaints from some respondent governments” generated considerable pressure to add a mechanism for the expedient dismissal of unmeritorious claims.\footnote{Antonio Parra, “The Development of the Regulations and Rules of the International Centre for Settlement of Investment Disputes” (2007) 41 Intl Lawyer 47 at 65 [Parra].} In response to states’ grievances, in 2005 the ICSID

---


\footnote{Antonio Parra, “The Development of the Regulations and Rules of the International Centre for Settlement of Investment Disputes” (2007) 41 Intl Lawyer 47 at 65 [Parra].}
Secretariat published a working paper, “Suggested Changes to the ICSID Rules and Regulations”, that proposed an expedited screening procedure for rejecting non-meritorious claims. In the paper, the Secretariat stated that:

It is suggested to make it clear, by the introduction of a new paragraph (5), that the tribunal may at an early stage of the proceeding be asked on an expedited basis to dismiss all or part of a claim on the merits. The change would be helpful in addressing any concerns about the limited screening power of the Secretary-General.\(^\text{88}\) [emphasis added]

Accordingly, the ICSID Secretariat introduced an amendment to the ICSID Arbitration Rules in 2006 by adding Article 41(5), which established an expedited procedure “for the early dismissal by arbitral tribunals of patently unmeritorious claims”.\(^\text{89}\) Rule 41(5) of the ICSID Arbitration Rules provides that:

Unless the parties have agreed to another expedited procedure for making preliminary objections, a party may, no later than 30 days after the constitution of the Tribunal, and in any event before the first session of the Tribunal, file an objection that a claim is manifestly without legal merit. The party shall specify as precisely as possible the basis for the objection. The Tribunal, after giving the parties the opportunity to present their observations on the objection, shall, at its first session or promptly thereafter, notify the parties of its decision on the objection. The decision of the Tribunal shall be without prejudice to the right of a party to file an objection.

\(^{88}\) ICSID Secretariat, “Suggested Changes to the ICSID Rules and Regulations” (2005) at 7, online:<https://icsid.worldbank.org/en/Documents/resources/Suggested%20Changes%20to%20the%20ICSID%20Rules%20and%20Regulations.pdf>. Frivolous claims were in the spotlight during ICSID negotiations back to the 1960s. A number of negotiating parties raised questions about potential vexatious and frivolous litigation that could result in significant losses of both time and money. Accordingly, one of the representatives submitted that a “provision was required to discourage groundless or vexatious claims”. The World Bank admitted that the comments of the delegates on frivolous claims required attention, and it offered the following solution: “However, as a practical matter it would appear to be proper in such an extreme case [of frivolous claim] for the Secretary-General of the Centre to seek to dissuade the claimant from proceeding with his claim”. Yet, under the final text of the ICSID Convention, neither the ICSID Centre nor the Secretary General of the Centre performs a role of dissuading claimants from submitting frivolous claims. Article 36(3) of the final version of the ICSID Convention specifies that the Secretary General can refuse to register the request for arbitration. See Memorandum of Meeting of Executive Directors on the Subject of Settlement of Investment Disputes, (13 March 1962) SecM 62-68 at para 2(d), (19 January 1962) SecM 62-17. However, Schreuer notes in his Commentaries to the ICSID Convention that the Secretary General simply rejects the request for arbitration “only if the lack of jurisdiction is so obvious that the request does not deserve consideration by a tribunal.” Christoph H Schreuer, The ICSID Convention: A Commentary (Cambridge: Cambridge University Press, 2001). The mechanism for early dismissal was eventually introduced not in the ICSID Convention, but in the ICSID Arbitration Rules, and only after 2005.

\(^{89}\) Parra, supra note 87 at 65.
pursuant to paragraph (1) or to object, in the course of the proceeding, that a claim lacks legal merit. \[90\] [emphasis added]

Accordingly, the ICSID Arbitration Rules require a claim to satisfy the standard “manifestly without legal merit” to curtail it expediently. Of course, an analysis of relevant arbitral awards is needed to understand what sort of claims are frivolous according to an ISA tribunal.

While there is no stare decisis in investment law, ISA tribunals “can and do refer to decisions of other tribunals” in order to clarify the meaning and the scope of investment rules. \[91\] Admittedly, investment tribunals may change their interpretation of Rule 41(5). Accordingly, this subsection examines the currently existing standard “manifestly without legal merit”. Firstly, this section considers the standard “manifestly”. Secondly, it reviews the standard “without legal merit”.

i. “Manifestly”

This subsection shows that in interpreting the standard “manifestly”, tribunals require a respondent to demonstrate that the claim lacks merit ‘on its face’. \[92\] The claim should not involve any novel issues of law or allegations of bad faith. \[93\] Tribunals refuse to apply the Rule 41(5) if a claim fails to meet these requirements.

In Trans-Global Petroleum, Inc v Jordan, a US investor concluded a concession agreement to explore a designated area of the Dead Sea. \[94\] The Jordanian government then made several attempts to remove Trans-Global Petroleum from the concession in favour of a local investor. \[95\] In Trans-Global, the government invoked Rule 41(5)—the first time the rule was cited in


\[91\] For example, see Bjorklund, supra note 55 at 268.

\[92\] Trans-Global v Jordan Trans-Global Petroleum, Inc v Hashemite Kingdom of Jordan (2008), ICSID Case No. ARB/07/25 (International Centre for Settlement of Investment Disputes) at para 84 [Trans-Global].

\[93\] MOL Hungarian Oil and Gas Company Plc v Republic of Croatia: Decision on Respondent’s Application under ICSID Arbitration Rule 41(5) (2014), ICSID Case No. ARB/13/32 (International Centre for Settlement of Investment Disputes) at para 53 [MOL Hungarian Oil].

\[94\] Trans-Global, supra note 92 at para 48.

\[95\] Ibid at para 52.
arbitration. The tribunal concluded that the word “manifestly” should mean “self-evident”, “clear”, “plain on its face” or “certain”.96

In subsequent arbitral awards, tribunals held that the “manifestly” standard could not be satisfied in cases where the parties “have locked horns over the possible differences” as to the interpretation of legal rules. For example, in MOL Hungarian Oil and Gas Company Plc v Republic of Croatia, the foreign investor sued Croatia for systemic obstacles, such as a refusal to issue licenses and permits that prevented the claimant from successfully exercising its concession rights.97 In applying Rule 41(5), the tribunal emphasized that both the claimant and the respondent:

have between them cited some half a dozen ICSID cases on the relationship between treaty claims and contract claims [...] [it] would seem to show that, whatever the merits of the Respondent’s objection, it cannot be deemed ‘manifest’ so as to meet the standard set by Rule 41(5).98

Hence, the summary review procedure under Rule 41(5) applies only to undisputed rules of law. PNG Sustainable Development Program Ltd v Papua New Guinea99 further illustrated this point. The proceedings concerned the claimant’s investment in a gold mine and a pit copper mine in Papua New Guinea.100 The government filed a preliminary objection, alleging that the claim did not satisfy the jurisdictional requirements of the applicable IIA and lacked legal merit.101 The tribunal emphasized that review in a summary fashion involves the application of “undisputed or genuinely indisputable rules of law to uncontested facts”.102 As a result, the tribunal dismissed the preliminary objection of Papua New Guinea under Rule 41(5). The tribunal justified its decision as follows:

[It] would in principle be inappropriate to consider and resolve novel issues of law in a summary fashion, which would inevitably limit the Parties’ opportunity to be heard and the Tribunal’s opportunity to reflect.103

---

96 Ibid at paras 83-84, 105.
97 MOL Hungarian Oil, supra note 93 at para 19.
98 Ibid at para 49.
100 Ibid at para 17.
101 Ibid at para 30.
102 Ibid at para 89.
103 Ibid at 94.
In cases of uncertainty, the existence of a legitimate dispute over the meaning of the law means that a claim cannot be “manifestly” without legal merit. In RSM Production Corporation and Others v Grenada, the tribunal added that “in cases of doubt or uncertainty as to the scope of a claimant’s allegation(s), any such doubt or uncertainty should be resolved in favour of the claimant”.

Allegations of bad faith also cannot be reviewed in summary fashion. According to the tribunal in Mol Hungarian Oil, serious “imputations of bad faith” and the allegations of the respondent that a claimant is “pursuing claims that it knows are baseless in order to obtain relief to which it knows it is not entitled” cannot not be adjudicated in limine and require full hearings.

ii. “Without Legal Merit”

In interpreting the phrase “without legal merit”, tribunals have concluded that “the adjective legal” in Rule 41(5) is clearly used in contradistinction to “factual”. Accordingly, tribunals have held that “the tribunal is in no position to decide disputed facts alleged by the parties” early in the proceedings. Accordingly, the tribunal has to address legal merits as if the facts alleged by the claimant were true. The tribunals in Trans-Global and in Brandes Investment Partners LLP v Venezuela reached the same interpretation of the adjective “legal” in the context of ICSID Rule 41(5).

The tribunals also concluded that a respondent could take advantage of Rule 41(5) both at the stages of jurisdiction and merits. Accordingly, respondent states could raise their preliminary objections under Rule 41(5) alleging both lack of jurisdiction and lack of merit on the substance of the dispute. The tribunal in Brandes Investment Partners v Venezuela noted:

---

104 RSM Production Corporation and others v Grenada: Award (2010), ICSID Case No. ARB/10/6 (International Centre for Settlement of Investment Disputes) [RSM Production].
105 Ibid at para 6.1.3.
106 MOL Hungarian Oil, supra note 93 at para 53.
107 Trans-Global, supra note 92 at para 97.
108 Ibid.
109 Ibid at para 93.
110 Trans Global, supra note 92 at para 93; Brandes Investment Partners LP v Venezuela: Decision on the Respondent’s Objection under Rule 41 (5) of the ICSID Arbitration Rules (2009), ICSID Case No. ARB/08/3 (International Centre for Settlement of Investment Disputes) at paras 58-59 [Brandes Investment].
Rule 41(5) does not mention “jurisdiction”. The terms employed are “legal merit”. This wording, by itself, does not provide a reason why the question whether or not a tribunal has jurisdiction and is competent to hear and decide a claim could not be included in the very general notion that the claim filed is “without legal merit”.111

After analyzing the case-law, it becomes apparent that arbitration tribunals set a high threshold to dismiss claims as meritless. To be dismissed as frivolous, claims must deal only with the application of “undisputed or genuinely indisputable rules of law to uncontested facts”.112 Accordingly, a claim must meet a high threshold of manifestly without legal merit to be dismissed under the ICSID Arbitration Rules.113

iii. Claims in Which the Award in Favour of Claimant Cannot be Made

The Dominican Republic-Central American FTA (DR-CAFTA)114 entails a mechanism for curtailment of frivolous claims that is distinctive from the ICSID Arbitration Rules and allows a broader range of claims to be curtailed in limine. According to art. 10.20 paragraph 4 of DR-CAFTA,

[A] tribunal shall address and decide as a preliminary question any objection by the respondent that, as a matter of law, a claim submitted is not a claim for which an award in favor of the claimant may be made.115 [emphasis added]

Art. 10.20 paragraph 4 appears to construe early dismissal of an investor’s claim in terms of viability or, more precisely “likelihood of success”. Arthur Rovine explains that “it is based on what American practitioners will understand as a motion to dismiss for failure to state a claim”.116 The tribunal must address and decide at the stage of preliminary objection whether a claim may succeed as a matter of law assuming that the facts alleged by the claimant are true. The burden to persuade the tribunal that a claimant has no case rests with the respondent.117

111 Brandes Investment, ibid at para 50.
112 PNG Sustainable Development, supra note 99 at para 89.
113 MOL Hungarian Oil, supra note 93 at para 44.
114 DR-CAFTA, supra note 14 at 10.20.4.
115 Ibid.
117 Pac Rim Cayman LLC v The Republic of El Salvador: Decision on Preliminary Objection 10.20.4 and 10.20.5 (2010), ICSID Case No. ARB/09/12 (International Centre for Settlement of Investment Disputes) at para 111 [Pac Rim].
The provisions of art. 10.20 paragraph 4 fell under the scrutiny of an ISA tribunal in Pac Rim v El Salvador, which provides a unique opportunity to examine how the ISA tribunal evaluates this “likelihood of success” standard and discusses exactly what threshold the claim must meet. Pac Rim Cayman, a US based company, brought a number of DR-CAFTA and non-DR-CAFTA based claims against El Salvador. According to Pac Rim’s submission, El Salvador violated its investment obligations under DR-CAFTA when it denied approval of an environmental permit and refused to prolong an exploration licence relating to operations in the country.\textsuperscript{118}

El Salvador brought a preliminary objection under DR-CAFTA and alleged that the Pac Rim’s claim was frivolous, in particular that an award in favour of the claimant could not be made. The preliminary objection by El Salvador in the Pac Rim case invited the ISA tribunal to discuss the standard of review under the art. 10.20 paragraph 4. The tribunal in Pac Rim stated that:

> [t]he Tribunal does not consider that the standard of review under Article 10.20.4 is limited to “frivolous” claims or “legally impossible” claims, contrary to the submissions of the Claimant. These words could have been used by the Contracting Parties in agreeing CAFTA; but all are significantly absent. Moreover, the implied addition of these or similar words would significantly restrict the arbitral remedy under Article 10.20.4, when the structure of this provision permits a more natural and effective interpretation consistent with its object and purpose.\textsuperscript{119} [Emphasis added]

Under such interpretation, art. 10.20 paragraph 4 appears not to require the claims to meet the standard “manifestly without legal merit”. The tribunal noted that for an early dismissal of a claim under art. 10.20 paragraph 4, “a tribunal must have reached a position, both as to all relevant questions of law and all relevant alleged or undisputed facts, that an award should be made finally dismissing the claimant’s claim at the very outset of the arbitration proceedings, without more”.\textsuperscript{120} The tribunal emphasized that it may refuse to dismiss a claim “depending on a particular circumstances of each case” even where “such a claim appeared likely (but not certain) to fail if assessed only at the time of the preliminary objection”.\textsuperscript{121}

Concerning any possible comparisons between the ICSID Arbitration Rule 41(5) and DR-CAFTA art. 10.20 paragraph 4, the Tribunal in Pac Rim

\textsuperscript{118} Ibid at paras 60-64.
\textsuperscript{119} Ibid at para 108.
\textsuperscript{120} Ibid at para 110.
\textsuperscript{121} Ibid at para 110.
noted as follows, “the Tribunal was also not materially assisted by comparisons […] with the New ICSID Arbitration Rule 41.5, which has different wording and does not share exactly the same object and purpose”.122 Clearly, the tribunal distinguishes between the scope of the ICSID Rule 41(5) and the DR-CAFTA provision and acknowledges that DR-CAFTA should be interpreted differently.

The DR-CAFTA rule on frivolous claims and its interpretation in the Pac Rim case are important for this article for two reasons. Firstly, DR-CAFTA provisions on frivolous claims resemble the wording in CETA, in which art. 8.10 also suggests that to curtail claims, “in which as a matter of law, an award in favour of claimant cannot be made”; thus the interpretation by the Pac Rim tribunal may chart the approach that future CETA tribunals can undertake with regards to preliminary objections of a similar sort. Secondly, DR-CAFTA integrates a conventional model of investment arbitration which embeds a number of institutional qualities that may influence the arbitral decision-making in assessing the likelihood for success of a particular claim.123

Of course, both of these identified features require further analysis, which this article intends to provide. Accordingly, in the following sections, this article will show what impact the flaws of the conventional ISA model may have on early dismissal of frivolous claims. This article will then discuss the CETA’s mechanisms for the dismissal of frivolous claims together with the institutional changes that treaties offer and will address the potential impact of these provisions upon early dismissal of frivolous claims.

Why the Institutional Design of the Conventional ISA Model Fails to Equip Arbitrators to Dismiss Frivolous Claims in limine

As has been discussed, under the international investment regime, arbitrators are in charge of early dismissal of frivolous claims. This section submits that the conventional model of ISA poorly equips arbitrators to perform this function. Accordingly, this section does not aim to persuade a reader that the arbitrators are “villains”124 but rather highlights the systemic features of the conventional ISA model that appear to decrease the ability for the arbitrators to exercise their function in the early curtailment of frivolous

---

122 Ibid at para 118.
123 See sections that follow.
claims. These specific features will be discussed in the shadow of the various criticisms of the international investment regime. This section will approach the discussion from two angles. Firstly, that the institutional design does not equip arbitrators well to dismiss frivolous claims in limine. Secondly, this article discusses perceived arbitral bias as one of the reasons why the conventional model of ISA may offer insufficient mechanisms to effectively curtail frivolous claims.

**ISA Tribunals May Be Inconsistent in Their Decision-Making and ISA Tribunals May Apply and Interpret the Investment Rules More Broadly Than Intended**

The mechanisms for early dismissal of claims require arbitrators to evaluate the merits of the claim, in particular by examining whether an award in favour of the claimant may be made. This evaluation of course is impossible without knowing the scope of the applicable investment standard. To put it differently, in order to evaluate the possibility of success of a claim as a matter of law, the arbitrators must first know what the content of the applicable legal rule is.

According to Michael Reisman, investment standards (e.g. FET, expropriation etc.) are “evaluation rules”, meaning that “those who apply it [have] to take account of a range of variables and to exercise judgment as to their contribution, in varying, idiosyncratic contexts, to realizing the goal or goals that have been specified”. In practice, it means that states formulate investment standards in a fashion that leaves room for arbitral discretion in their interpretation and application. Presumably, arbitral practice may result in “incremental development of substantive law” and crystallization of the scope of applicable investment standards. This way, the arbitrators may easily evaluate the merits of a claim early in the proceedings.

Yet ISA tribunals are famous for generating inconsistent decisions. Sometimes, the tribunals even ignore the interpretations of legal rules previously given by their colleagues even when the claim arose from the same

---


facts and under the same BIT. Under the conventional ISA model, there is no hierarchical structure among the tribunals that may (at least in theory) compel arbitrators to follow each other’s interpretations. Notably, the international investment regime rejects *stare decisis*, which in national legal systems promotes certainty of legal norms and helps to crystalize the content of those norms. As such, the argument is not that *stare decisis* shapes ‘good’ law (it might not) but that it renders considerable social benefits by ensuring a consistency of application that upholds the credibility of the adjudicatory system. Of course, the benefits of consistency cannot be confined to credibility alone but also extend to the proper functioning of the mechanisms of early curtailment of frivolous claims. In particular, a settled matter of law provides guidance to claimants on the pattern of permissible behaviour within the legal system and to the adjudicators concerning which claims can survive legal scrutiny when the facts alleged by a claimant are presumed to be true. Admittedly, *jurisprudence constante* can provide guidance for tribunals as to which claims are legally viable.

Yet the value of *jurisprudence constante* is limited as arbitrators are free to ignore the findings of their colleagues without any sanction. Accordingly, tribunals may evaluate the merits of the claims distinctively depending on how the arbitrators themselves understand, interpret and apply the investment rules. The ICSID standard “manifestly without legal merit” is set so high that it may even be immune to such an issue because the ICSID Arbitration Rules require a claim to be *prima facie* without legal merit.

The DR-CAFTA standard is different and requires the ISA tribunal to engage in evaluating the probability of success of a claim. The likelihood of success depends on the application and interpretation of substantial legal standards. In the conventional model of ISA, where tribunals interpret and apply investment rules differently (often broader or narrower meanings), it is


130 Waldron, *supra* note 128 at 3.


132 Ibid at 277.
difficult to identify a single consistent pattern of interpretation and application of investment rules. In absence of such a single mode of application, it is impossible to anticipate the outcome of a tribunal’s evaluation of success of a claim at the early stages.

In other words, probability of success is a category without clearly defined metrics, at least in investment arbitration where the principle of *stare decisis* does not govern the development of jurisprudence. Domestic legal systems embed mechanisms for the early curtailment of frivolous claims such as the summary of judgment rule and national courts may evaluate a claim on the merits because the governing principle of *stare decisis* determines the boundaries for permissible conduct by the claimant. In addition, the domestic system usually has an appeal mechanism that oversees the consistency of legal application and interpretation. These institutional features provide an infrastructure of support that permits the adjudicators (or in the domestic context, judges) to evaluate the merits of a claim early on. In such a setting, when the law is silent or the claimant introduces a novel issue, national courts most likely will proceed with a full trial on the merits. However, in the context of arbitration (where there is no *stare decisis* or hierarchy of courts that ensures consistent and coherent interpretation of legal rules), this is often not the case.

*Investment Arbitration Does Not Embed Any Incentives for Arbitrators to Curtail Frivolous Claims Early*

This subsection suggests that the conventional model of investment arbitration does not embed incentives for the arbitrators to curtail frivolous claims expediently. Particularly, the link between the arbitrators and the structure of their compensation and the process of their selection cannot be overlooked.\(^{133}\)

\(^{133}\) Paul Greenberg and James Hayley suggest that the structure of compensation is directly linked to independence of the judiciary. For example, if the compensation is low, it signals that the candidates in judges could potentially seek nonmonetary gains from the judicial position and the court’s system will become more prone to abuse. This is the case for judges because they are banned from undertaking other types of employment. For the arbitrators in the conventional arbitration model, this is not the case. The structure of compensation thus could be utilized to attract “properly motivated individuals” to access the bench. See Paul E Greenberg & James A Hayley, “The Role of the Compensation Structure in Enhancing Judicial Quality” (1986) 15 J Leg Stud 417 at 418, 425.
Before delving into this discussion, it is necessary to mention that this dissertation embraces the “homo economicus paradigm” in adjudication and as such views arbitrators as rational actors who act in self-interest in order to maximize utility and as such are not immune to incentives that promise economic and personal gain. Investment arbitration, as this section will show, lacks structures that would divorce arbitral action from such incentives.

Under the current investment arbitration model, arbitrators receive remuneration while participating in the dispute and as such their remuneration depends on the continuity of that dispute. Of course, the arbitrators may never admit that arbitral appointment attracts them by virtue of its monetary or non-pecuniary gains such as considerations of prestige or compensation. However, the structure of arbitral compensation shows that the arbitrators ultimately have a financial interest in the continuity of investment proceedings. Again, according to Schneiderman and Van Harten, the cost of arbitration is high. Lee M. Caplan underscores how arbitrators earn on average US$3,000 a day in addition to travel and accommodation allowances. For example, the arbitrator in Chevron and

---

134 Russell Smyth, “Do Judges Behave As Homo Economicus, and If So, Can We Measure Their Performance? An Antipodean Perspective on Tournament of Judges” (2004-2005) 32 Fla St UL Rev 1299 at 1301. Smyth explains that homo economicus is “sparked mainly by Posner's seminal work on judicial decision making [...] that argued that judges, like the rest of us, are self-interested and respond to incentives”. See also Richard A Posner, “What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)” (1993) 3 Sup Ct Econ Rev 1 at 2-3.

135 Sergio Puig, “Social Capital in the Arbitration Market” (2014) 25:4 Eur J Intl L 387 at 398. Puig indicates that “in addition to the visibility and reputational value of an appointment to an investor-state case, the financial incentives are considerable. At US$3,000.00 an eight-hour day (plus expenses), arbitrators make on average $200,000 per case”.


137 Eberhardt, supra note 68.


139 Lee M Caplan, “A Proposed Set of Arbitration Rules for Weaker Players” in Karl Sauvant, ed, Yearbook on International Investment Law & Policy 2009-2010 (New York: Oxford University Press, 2010) at 345. Particularly, Caplan submits that “As a rule, ICSID arbitrators are paid US $3000 per day [...] to the contrary, in ad hoc arbitration, arbitrators are free to charge their normal, private sector hourly fees, which may range from as much as US 700 to over US 1000 per hour”.

---
Texaco v Ecuador received a total remuneration of US$939,000. As such, the persisting concern of arbitral bias is not a matter of the conduct of individual arbitrators but arises from “the broader institutional context within which decisions are taken”.

The institutional structure of investment arbitration pertaining to the adjudicators does not provide for the administrative and economic safeguards typical of judicial independence in domestic legal systems, such as security of tenure, rotation of case assignments and prohibition of outside employment. The lack of these structures, particularly security of tenure, makes arbitrator remuneration dependent on proceedings continuing and on “whoever has the ability to make the claims and trigger the arbitrators’ appointments”. As Gus Van Harten has noticed, the current design of investment arbitration provides only the foreign investors with this ability. These features of the international investment regime make arbitrators vulnerable to financial incentives as motivation not to curtail a claim early. Accordingly, these features highlight why the existing mechanisms for early curtailment of frivolous claims in the international investment regime are not particularly useful given institutional features such as the structure of arbitral compensation and general uncertainty over the content of legal norms.

140 Eberhardt, supra note 68 at 35.
141 It is not to say that particular cultural background features of individual arbitrators (such as, age, gender, previous affiliation, current primary occupation) do not matter; they do. For a more general discussion on the importance of individual factors for judges, see Cohen, supra note 136 at 16; Jonathan Kastellec, “Racial Diversity and Judicial Influence on Appellate Courts” (2013) 57:1 American J Political Science 167. For more specific information on the investment arbitration context, see Joost Pauwelyn, “The Rule of Law without the Rule of Lawyers? Why Investment Arbitrators are from Mars, Trade Adjudicators from Venus” (2015) 109:4 AJIL 761; David Schneiderman, “Judicial Politics and International Investment Arbitration: Seeking an Explanation for Conflicting Outcomes” (2010) 30:2 Nw J Intl L & Bus 383 at 385.
143 Ibid.
144 Ibid.
Examining CETA’s Quasi-Investment Court in Light of Early Curtailment of Frivolous Claims

CETA contains two articles that aim to combat frivolous claims early in arbitral proceedings. CETA’s art. 8.32 focuses on claims that are manifestly without legal merit, and art. 8.33 focuses on claims that are unfounded as a matter of law.\footnote{Since CETA introduces two standards for the early dismissal of frivolous claims, it appears worthwhile to consider (at least briefly) the issue of the interaction of CETA’s articles 8.32 and 8.33. Particularly, whether or not it is useful to have both standards in CETA. CETA imposes restrictions on using both mechanisms simultaneously by the state-respondents. Accordingly, CETA specifies that a respondent cannot submit an objection under article 8.32(1) “if the respondent has filed an objection pursuant to Article 8.33”. In turn, Article 8.33(3) stipulates that “[i]f an objection has been submitted pursuant to Article 8.32, the Tribunal may, taking into account the circumstances of that objection, decline to address” an objection under article 8.33(1). Presumably, CETA introduces such restrictions to ensure that the state respondents will not unnecessarily prolong proceedings by submitting multiple preliminary objections.} Art. 8.32 copies the wording of ICSID Arbitration Rule 41(5). Hence, it is plausible to assume that the tribunals will follow an interpretation similar to the ICSID standard. On the other hand, art. 8.33 is distinct in its formulation from both art. 8.32 and ICSID rule 41(5). Art. 8.33(1) appears to follow DR-CAFTA in expanding the range of claims that can be dismissed in a summary fashion. These claims are not limited only to submissions that “manifestly” lack “legal merit” but also include claims that cannot potentially succeed. As such, CETA provides state-respondents with two fast-track procedures for early dismissal of a frivolous claim.

However, since there is obviously no arbitral practice under CETA yet, it is difficult to foresee exactly how CETA tribunals will examine the viability of claims and measure the potential of a claim’s success early in such proceedings. Nonetheless, it is still possible to explore the features of CETA’s ISA dispute resolution model in light of its provisions on the early curtailment of frivolous claims.

As discussed above, the CETA parties have changed the conventional model of ISA in two important ways where frivolous claims are concerned. The first feature is the appeal body, may play a significant role in ensuring the consistency of arbitral practice when applying and interpreting investment rules. As a result, CETA tribunals will have a clearer understanding of which claims may succeed on the merits given how the investment rules have been interpreted and applied. Of course, this will impact the early curtailment of frivolous claims, particularly under art. 8.33, which requires that CETA...
tribunals examine whether a claim has any potential of success. In particular, when clear and consistent standards are in place, CETA tribunals will have an opportunity to see at an early stage whether a claim is viable. This effect has long-term benefits for both states and investors. For example, foreign investors will know and understand the parameters of viable claims, which will help guide decisions and actions in investment proceedings. The availability of such fast-track procedures may take some pressure off of states when foreign investors bring claims or threaten to do so.

Second, CETA changes the procedure for the appointment of arbitrators. CETA creates a standing arbitration tribunal with 15 arbitrators of particular qualification and vests the state-parties with the power to appoint. Such architecture creates a ‘scarcity’ of arbitral resources and may result in greater incentive for the arbitrators to curtail frivolous submissions. In addition, arbitrators will be kept on retainer under the CETA model. It is plausible to suggest that such retainers may disincentivize arbitrators from prolonging proceedings for financial gain. However, David Schneiderman argues that this outcome is improbable because arbitrators will still remain dependent upon investors to launch arbitrations. According to Schneiderman:

> [a]rbitrators will continue to remain dependent upon investors – and only those with sufficiently deep pockets – to finance litigation against states. These reforms do little, then, to remove the financial incentives tribunal members already have to read treaty standards expansively so as to ensure future employment.\(^\text{146}\)

Accordingly, while the arbitrators are on retainer, they will still be paid by the hour and may still have a financial incentive not to dismiss the claims early.

**CONCLUDING REMARKS**

Frivolous claims in the international investment regime are particularly problematic because they undermine its credibility and threaten to impose serious costs on state-respondents. Prior to CETA, the ICSID Arbitration Rules offered the primary mechanism for dismissing claims that manifestly lack legal merit, which was a high threshold to satisfy. Later on, DR-CAFTA introduced another standard where an award in favour of the claimant cannot be made. Both mechanisms appear to be useful only to a limited extent given that the mechanisms are linked to the conventional ISA model.

CETA not only incorporates both mechanisms but also links them to profound institutional changes. CETA alters the conventional model by changing the procedure for the appointment of arbitrators, the rules of remuneration of arbitrators and introduces an appeal body. There is good reason to be optimistic in light of these changes where early dismissal of frivolous claims is concerned, particularly the introduction of a supervising appellate body, which could ensure consistency in applying and interpreting investment rules. This will provide guidance as to which investors’ claims may have a chance to succeed on the merits, something the current system sorely lacks. It is plausible to assume that when arbitrators possess greater certainty regarding the pattern in applying legal rules, there is a greater chance that unfounded claims will be curtailed at the early stages without causing negative effects either on the credibility of the international investment regime or through imposing high costs upon states. However, CETA does have potential for improvement in many areas, especially concerning the structure of arbitral compensation. By continuing to fail to address the underlying financial incentives arbitrators have to avoid dismissing claims early, the structural problems of the past may persist and will likely require further consideration if the international investment regime as a whole is to maintain its credibility and efficacy going forward.
1. Introduction

African countries have conventionally responded to the socially irresponsible behaviour of resource extraction companies by enacting ‘command-and-control’ rules and regulations. The ‘command-and-control’ approach to regulation involves statutory prescription of behavioural norms for industry participants. This approach is premised on the belief that, because violations of the prescribed norms attract legal sanctions, companies will be incentivized to comply with the prescribed norms so as to avoid those sanctions. The problem, however, is that ‘command-and-control’ rules are for the most part ineffective in regulating the activities of extractive companies in Africa. The ineffectiveness of ‘command-and-control’ regulation in Africa is traceable to several factors including the incapacity and inefficiency of regulatory agencies; inadequate funding of regulatory agencies; corruption; lack of political will to enforce laws and regulations against extractive companies; lack of independence on the part of regulatory agencies; and the
complicity of African governments in the social misbehaviour of extractive companies through state participation in mineral extraction projects.¹

In the recent past, however, some African governments appear to have recognized this apparent shortfall in regulatory enforcement. Hence, in the last few years, African governments have attempted to diversify the avenues and mechanisms for regulating extractive companies. This article examines the diversification of regulatory mechanisms through private citizen enforcement of regulatory laws in three African countries: Nigeria, South Africa and Zambia. Although occasional references are made to other African countries, I have chosen to focus on these three countries because of their common experience with regard to corporate social irresponsibility in their extractive sectors. The article argues that, while there are specific laws enabling citizen enforcement of regulatory provisions in these countries, certain structural and institutional obstacles including rampant poverty, high cost of litigation and state participation in resource exploitation prevent the full utilization of these statutory provisions by African citizens.

2. Statutory Provisions Enabling Citizen Enforcement of Environmental Regulation

As discussed below, a few African countries have enacted statutory provisions permitting citizens to enforce environmental regulations. Such statutory provisions are not unique to African countries. Some developed countries including the United States of America accord private citizens and non-governmental organizations (NGOs) the right to enforce environmental laws in circumstances where regulatory authorities are unwilling or unable to enforce such laws.² The empowerment of private citizens to enforce environmental laws is intended “to supplement government action, to make up the balance of necessary enforcement at times when under-funded or over-worked agencies could not ensure that all laws are complied with.”³

The incapacity of the public agencies charged with regulating the conduct of extractive companies in Africa makes a compelling case for the

¹ Evaristus Oshionebo, Regulating Transnational Corporations in Domestic and International Regimes: An African Case Study (Toronto: University of Toronto Press, 2009) at 71-78 [Oshionebo, “Regulating”].
³ Ibid at 2.
diversification of regulatory enforcement. In this regard, a few countries in sub-Saharan Africa have enacted statutory and constitutional provisions enabling private enforcement of regulatory standards. In South Africa, a citizen’s right to enforce regulatory rules is constitutionally protected through the Bill of Rights enshrined in the Constitution. Section 38 of the Constitution of the Republic of South Africa grants citizens “the right to approach a competent court, alleging that a right in the Bill of Rights has been infringed or threatened”. The Constitution casts a wide net with regard to the persons who may institute public interest litigation in South Africa. These persons include:

(a) anyone acting in their own interest; (b) anyone acting on behalf of another person who cannot act in their own name; (c) anyone acting as a member of, or in the interest of, a group or class of persons; (d) anyone acting in the public interest; and (e) an association acting in the interest of its members.

One of the rights guaranteed under the South African Bill of Rights is the right to a clean and healthy environment. In this regard, section 24 of the Constitution provides that:

Everyone has the right (1) to an environment that is not harmful to their health or well-being; and (2) to have the environment protected, for the benefit of present and future generations, through reasonable legislative and other measures that (a) prevent pollution and ecological degradation; (b) promote conservation; and (c) secure ecologically sustainable development and use of natural resources while promoting justifiable economic and social development.

In effect, these constitutional provisions allow South African citizens to enforce environmental laws provided they can establish that the alleged infringement directly impacts their right to a clean and healthy environment.

The legal standing of South African citizens to enforce environmental laws is equally apparent in the National Environmental Management Act 1998 which provides that:

Any person or group of persons may seek appropriate relief in respect of any breach or threatened breach of any provision of this Act, including a principle contained in

---


6 Ibid, s 24.
Chapter 1, or of any provision of a specific environmental management Act, or of any other statutory provision concerned with the protection of the environment or the use of natural resources -

(a) in that person’s or group of person’s own interest;
(b) in the interest of, or on behalf of, a person who is, for practical reasons, unable to institute such proceedings;
(c) in the interest of or on behalf of a group or class of persons whose interests are affected;
(d) in the public interest; and
(e) in the interest of protecting the environment.\(^7\)

In addition, South Africa’s *National Environmental Management Act 1998* empowers private prosecution of environmental offences.\(^8\) Thus, South African citizens can institute proceedings against extractive companies for environmental offences where public regulatory agencies fail to institute such proceedings against companies.\(^9\)

Similarly, Zambia has specifically empowered its citizens to enforce environmental provisions in the *Mines and Minerals Development Act 2015*\(^10\) and the *Environmental Management Act 2011*.\(^11\) For example, section 87(7) of the *Mines and Minerals Development Act 2015* provides that:

A person, group of persons or a private or State organisation may bring a claim and seek redress in respect of the breach or threatened breach of any provision relating to damage to the environment, biological diversity, human and animal health or to socio-economic conditions -

(a) in that person’s or group of persons’ interest;
(b) in the interest of or on behalf of, a person who is, for practical reasons, unable to institute such proceedings;
(c) in the interest of, or on behalf of, a group or class of persons whose interests are affected;
(d) in the public interest; and
(e) in the interest of protecting the environment or biological diversity.\(^12\)

Nigeria has yet to fully embrace a rights-based approach to natural resource governance. Generally speaking, statutory enactments governing Nigeria’s resource sector do not vest a right on citizens to enforce

---

\(^7\) *National Environmental Management Act 107 of 1998*, s 32(1) (South Africa) \[Environmental Management Act (South Africa)\].

\(^8\) Ibid, s 33.

\(^9\) Ibid.


\(^11\) *Environmental Management Act of 2011*, ss 108-110 (Zambia) \[Environmental Management Act (Zambia)\].

\(^12\) *Mines and Minerals Act (Zambia)*, *supra* note 10, s 87(7).
environmental laws and regulations. This perhaps explains why Nigerians have long resorted to common law tort litigation in their private attempts to enforce environmental laws against oil and gas companies.\(^\text{13}\) However, tort litigation often offers little hope because of the difficulties involved in obtaining evidence to prove negligence on the part of oil and gas companies.\(^\text{14}\) Proof of negligence in oil-related litigation is particularly difficult because of the technical and scientific nature of oil activities.\(^\text{15}\) Hence, in the recent past, Nigerian citizens have increasingly relied on the Constitution to hold oil and gas companies accountable for their social irresponsibility.

The Constitution of the Federal Republic of Nigeria 1999 does not expressly provide for environmental rights but it expresses the environmental objectives of Nigeria as the protection and improvement of the environment, including the safeguarding of water, air, land, forest and wild life.\(^\text{16}\) However, the Constitution vests a right on Nigerians to enforce the fundamental human rights provided for in the Constitution, including the right to life; right to dignity of human person; right to personal liberty; right to fair hearing; right to private and family life; right to freedom of thought, conscience and religion; right to freedom of expression; right to peaceful assembly and association; right to freedom of movement; right to freedom from discrimination; and right to own property.\(^\text{17}\) Interestingly, some of the human rights provisions in the Nigerian Constitution can be enforced against private entities including extractive companies because, as the Supreme Court of Nigeria held in Attorney General of Ondo State v Attorney General of the Federation and 35 Others,\(^\text{18}\) the provisions of the Constitution are not confined to governments but also apply to private persons, companies and private organizations. Thus, where the activities of an extractive company violate any of the fundamental human rights guaranteed under the Nigerian Constitution, the victim can sue the company to remedy the wrong.\(^\text{19}\)

In addition to these constitutional provisions, the human rights


\(^\text{15}\) Frynas, supra note 13 at 124.


\(^\text{17}\) Ibid, ss 33-43.


provisions enshrined in the *African Charter on Human and Peoples’ Rights*\(^2\) (African Charter) apply in African countries that have ratified the Charter. Nigeria has ratified and domesticated the African Charter through the *African Charter on Human and Peoples’ Rights (Ratification and Enforcement) Act*.\(^2\) This Act provides that the African Charter has “force of law in Nigeria and shall be given full recognition and effect and be applied by all authorities and persons exercising legislative, executive or judicial powers in Nigeria.”\(^2\) The African Charter guarantees a number of rights including the right to life (Article 4); right to “enjoy the best attainable state of physical and mental health” (Article 16); and the right to a clean and satisfactory environment (Article 24). The ratification of the African Charter enables Nigerian citizens to enforce the provisions of the Charter in both domestic courts and international tribunals such as the African Court on Human and Peoples’ Rights and the Court of Justice of the Economic Community of West African States (“Court of Justice of ECOWAS”).\(^3\)


It is practically impossible for regulatory agencies to detect and punish every violation of regulatory standards. It is equally the case that, sometimes, regulatory agencies may choose not to enforce regulation against erring companies so as to conserve scarce resources. At other times, regulators may prioritize their enforcement targets by focusing their enforcement efforts on major infractions while ignoring minor infractions. Further, as discussed below, because many African states participate directly in resource exploitation, regulatory agencies often elect not to enforce regulations against


\(^2\) Ibid.

\(^3\) This court was established by the *Protocol to the African Charter on Human and Peoples’ Rights on the Establishment of the African Court on Human and Peoples’ Rights*, online: <http://www.achpr.org/instruments/court-establishment/>.

\(^4\) The Court of Justice of the Economic Community of West African States was created by member states of the Economic Community of West African States on May 28, 1975. See *Economic Community of West African States: Revised Treaty*, online: <http://www.ecowas.int/wp-content/uploads/2015/01/Revised-treaty.pdf>.
TNCs engaged in joint venture projects with African governments. In all of these instances, a void is unwittingly created in the regulatory process. Citizen suits help to fill this void because they complement and, in some cases, supplement regulatory enforcement by state agencies.

This is particularly the case where state agencies lack the expertise “and financial resources to pursue all the enforcement cases that they might desire.” The complementary value of citizen suits is apparent in Africa where regulatory agencies lack the capacity and resources to enforce regulation. In fact, lack of institutional capacity is at the heart of many of the ills plaguing Africa’s natural resource sector, including the social irresponsibility of extractive companies. In this context capacity means the ability, expertise and resources to regulate companies operating in the extractive sector. The incapacity of regulatory agencies in Africa is compounded by the lack of financial and technical resources. In Nigeria for example, the National Oil Spills Detection and Response Agency (NOSDRA), an agency responsible for investigating oil spills, reportedly relies on oil and gas companies to provide scientific expertise and diagnostic equipment for its investigations. Regulatory agencies in Africa also often lack the political will to enforce regulations against extractive companies due primarily to the fact that these companies are considered by African governments as vital to their national economies.

Moreover, regulatory agencies are sometimes captured by the very entities they are supposed to be regulating, thus creating a void in regulatory enforcement. Public interest litigation could ameliorate the regulatory void arising from the capture of regulatory agencies by industry participants. Regulatory capture occurs where a regulatory agency becomes a tool of the companies and entities they are supposed to regulate and where the agency is “governed by the commercial interests of the parties it regulates, rather than

26 Africa Progress Panel, supra note 4 at 96.
27 Ibid at 61.
the public interest. As this author has argued elsewhere, transnational corporation (TNCs) in Africa’s extractive sector appear to have captured regulatory agencies, and hence, for the most part, these agencies either refuse to enforce regulation against TNCs or make regulatory choices that suit the interests of the TNCs.

Regulatory capture is particularly apparent in Nigeria where, as mentioned previously, the NOSDRA relies on oil and gas companies to provide scientific expertise and diagnostic equipment for its own investigations. The NOSDRA’s reliance on oil and gas companies for scientific expertise and equipment not only allows these companies to dictate the terms of oil spill investigations in Nigeria, including where and when the investigation will occur, but also makes it difficult for the NOSDRA to act as an unbiased and independent regulator.

Aside from the complementary attributes citizen suits bring, the involvement of citizens in the enforcement of regulation in Africa could lessen the burden on regulatory agencies and free up scarce resources which could then be devoted to other areas of need. For example, the cost and expense involved in public interest litigation are borne by citizen plaintiffs and, where such litigation is successfully concluded, the regulatory agency which ought ordinarily to institute the suit is spared the financial and material resources that it would have spent prosecuting the case.

Citizen suits could set regulatory precedents for both regulators and companies where there is successful litigation regarding statutory provisions on corporate conduct. Unlike regulatory agencies which are vested with statutory power to compel compliance with regulatory provisions, private citizens must invoke the judicial system and obtain an order of a competent court in order to compel companies to comply with regulations. Such suits enable citizens to seek judicial clarification and interpretation of regulatory provisions. In the process of interpretation, courts often set behavioural standards for companies and governments. In this sense, private citizen suits are mechanisms for standard-setting in the extractive sector. Litigation by

---

33 Ibid at 9.
34 Oshionebo, “Regulating”, supra note 1 at 217.
private citizens against governments and extractive companies could even ratchet up standards in the extractive sector.

This appears to be the case in countries such as Nigeria where section 33(1) of the Constitution, which guarantees the right to life, has been creatively utilized by citizens to enforce environmental standards against oil and gas companies. In Gbemre v. Shell, the court held that the right to life guaranteed under the Nigerian Constitution includes the right to a clean and pollution-free environment and that Shell acted in breach of the Constitution by polluting the environment through its gas-flaring activities. Although the Nigerian Constitution does not expressly recognize environmental rights, the court drew an immutable link between the right to life and the quality of the environment by interpreting section 33(1) of the Constitution broadly. The court expanded the scope of the right to life to include a right to a clean environment, thus setting a new environmental standard in Nigeria.

In South Africa, citizen suits have equally helped to shape and set the parameters for public participation in environmental regulation, as illustrated by the case of EarthLife Africa (Cape Town) v Director General, Department of Environmental Affairs and Tourism and Another. In this case, EarthLife Africa, an environmental NGO, sought to invalidate the authorization by the Director-General of the Department of Environmental Affairs and Tourism (DEAT) for the construction of a nuclear reactor. Relying on the Environmental Impact Assessment Regulations, which provide that all interested parties shall be “given the opportunity to participate in all the relevant procedures contemplated in these regulations”, the Plaintiffs alleged that the process leading to the authorization was flawed because the DEAT did not afford the Plaintiffs an opportunity to comment on the final Environmental Impact Report prior to granting the authorization. The Western Cape High Court held that the DEAT’s approach was “fundamentally unsound” and continued as follows:

The regulations provide for full participation in ‘all the relevant procedures contemplated in these regulations’. The respondents seek to limit such participation to the ‘investigation phase’ of the process (as contemplated by regs 5, 6 and 7). After

35 Gbemre v Shell, supra note 19.
36 2005 (3) SA 156 (S. Afr.) [EarthLife Africa].
submission of the EIR, however, the ‘adjudicative phase’ of the process commences, involving the DG’s consideration and evaluation, not only of the EIR, but also — more broadly — of all other facts and circumstances that may be relevant to his decision. There is nothing in the Act (ECA) or the regulations that expressly excludes public participation or application of the audi rule during this ‘second stage’ of the process. In line with settled authority, therefore, it follows that procedural fairness demands application of the audi rule also at this stage.37

This decision is significant not only because it nullified the authorization for the construction of the nuclear reactor but also because it delineated the parameters for public participation in environmental decision-making in South Africa. More specifically, the decision established the principle that, in South Africa, public participation is required right up to the stage of the final decision on the matter in question.

The standards-setting potential of citizen suits is equally apparent at the continental level where African citizens have utilized the African Charter to compel African governments to enhance the implementation and enforcement of regulation in the extractive sector. This is exemplified by cases such as SERAC & CESR v Nigeria38 and SERAP v Federal Republic of Nigeria,39 both of which were instituted by NGOs. In SERAC & CESR v Nigeria,40 the African Commission on Human and Peoples’ Rights (“African Commission”) held that Nigeria acted in violation of Articles 16 and 24 of the African Charter because it failed to take reasonable regulatory steps to prevent the state oil company, the Nigerian National Petroleum Corporation (NNPC), and Shell Petroleum Development Corporation (Shell) from polluting and degrading the environment as a result of their oil and gas operations.41 In particular, the African Commission held that Nigeria’s failure to regulate the NNPC and Shell deprived local communities of the right to health and the right to a clean and healthy environment.42 Equally significant is the case of

37 Ibid at para 89.
40 SERAC & CESR v Nigeria, supra note 38.
41 Ibid at para 50-54.
42 Ibid at para 52-54.
SERAP v Federal Republic of Nigeria\(^\text{43}\) where the Court of Justice of ECOWAS held that the failure on the part of the Nigerian government to effectively regulate the environmental activities of oil and gas companies amounted to a violation of Article 24 of the African Charter, which guarantees a right to a satisfactory environment favourable to human development.\(^\text{44}\)

Perhaps more significantly, the decisions in SERAC & CESR v Nigeria and SERAP v Federal Republic of Nigeria set important standards for regulating the operations of extractive companies in Africa. In SERAC & CESR v Nigeria, the African Commission held that Article 24 of the African Charter imposes an obligation on African governments “to take reasonable and other measures to prevent pollution and ecological degradation, to promote conservation, and to secure an ecologically sustainable development and use of natural resources.”\(^\text{45}\) According to the African Commission, governments desirous of complying with Article 24 of the African Charter must take several steps including:

... ordering or at least permitting independent scientific monitoring of threatened environments, requiring and publicising environmental and social impact studies prior to any major industrial development, undertaking appropriate monitoring and providing information to those communities exposed to hazardous materials and activities and providing meaningful opportunities for individuals to be heard and to participate in the development decisions affecting their communities.\(^\text{46}\)

The African Commission emphasized that African “[g]overnments have a duty to protect their citizens, not only through appropriate legislation and effective enforcement but also by protecting them from damaging acts that may be perpetrated by private parties.”\(^\text{47}\)

Similarly, in SERAP v Federal Republic of Nigeria\(^\text{48}\) the Court of Justice of ECOWAS established important regulatory standards for the governments of West African countries with regard to their regulatory obligations under Article 24 of the African Charter. The court held that Article 24 of the African Charter imposes “an obligation of attitude and an obligation of result” on African governments.\(^\text{49}\) The court held further that:

---

\(^{43}\) SERAP v Nigeria, supra note 39.

\(^{44}\) Ibid at paras 100-112.

\(^{45}\) SERAC & CESR v Nigeria, supra note 38 at para 52.

\(^{46}\) Ibid at para 53.

\(^{47}\) Ibid at para 57.

\(^{48}\) SERAP v Nigeria, supra note 39.

\(^{49}\) Ibid at para 100.
Article 24 of the Charter thus requires every State to take every measure to maintain the quality of the environment understood as an integrated whole, such that the state of the environment may satisfy the human beings who live there, and enhance their sustainable development. It is by examining the state of the environment and entirely objective factors, that one judges, by the result, whether the State has fulfilled this obligation. If the State is taking all the appropriate legislative, administrative and other measures, it must ensure that vigilance and diligence are being applied and observed towards attaining concrete results.  

According to the court, the mere enactment of legislation and the mere creation and funding of regulatory agencies by governments do not necessarily mean that the governments have complied with their regulatory obligations under Article 24 of the African Charter. Even where governments take such positive steps they “may still fall short of compliance with international obligations in matters of environmental protection if these measures just remain on paper and are not accompanied by additional and concrete measures aimed at preventing the occurrence of damage or ensuring accountability, with the effective reparation of the environmental damage suffered.”

Furthermore, because statutory and constitutional provisions regarding citizen enforcement of regulation allow citizens to sue regulatory agencies for their actions or inaction, citizen suits could enjoin these agencies from taking actions which are inimical to the interests of citizens. In *EarthLife Africa*, for example, an environmental NGO successfully sued to enjoin a regulatory agency in South Africa from authorizing the construction of a nuclear reactor unless the agency followed established legal standards. In this regard, citizen suits could potentially awaken the consciousness of regulatory agencies and perhaps prompt them to act to protect public interest.

In addition, citizen suits could act as stimulus for the enhancement of regulation either through amendments to extant laws or through enactment of new statutes. For example, it has been reported that the African Commission’s decision in *SERAC & CESR v Nigeria* prompted the Nigerian government to take several regulatory steps including the establishment of the NOSDRA and the revision and amendment of environmental laws and regulation to better monitor and control the operations of oil and gas.

---

51 *Ibid* at para 105.
52 *Earthlife Africa*, *supra* note 36.
companies.\textsuperscript{53} Relatedly, citizen suits could shame governments and extractive companies into enforcement of, and compliance with, regulatory standards. As discussed above, failure by governments to regulate the activities of extractive companies amounts to a breach of the governments’ obligation to protect human rights under the African Charter. Thus, a judicial pronouncement to the effect that an African government has breached its Charter obligations potentially stains the international reputation of the government and could attract international condemnation and criticism. The desire to avoid such unpleasant consequences will likely incentivize any responsible government to act proactively in the future to ensure that it regulates the conduct of extractive companies operating in its jurisdiction. Likewise, private citizen suits against extractive companies could prompt these companies to settle their disputes with local host communities and through that process the concerns of local communities could be addressed to the satisfaction of both parties.

Additionally, statutory and constitutional provisions permitting citizen enforcement of regulation empower activist groups in Africa to participate in the regulatory process. In South Africa, it has been observed that citizen litigation promotes participatory democracy because it enables “economically powerless or otherwise marginalized members of society to insist that those in power pay attention to their needs.”\textsuperscript{54} The participation of activist groups in the regulatory process is particularly significant in Africa because, for the most part, regulatory agencies in Africa lack both the capacity and the political will to regulate extractive companies. As mentioned previously, the incapacity of regulatory agencies creates a void in the regulatory process; citizen enforcement of regulatory rules could help to fill this void.

In this sense, citizen enforcement of regulation could potentially help to restrain the financial power of extractive companies, which they often use to extract investment terms that are inimical to the interests of host countries.\textsuperscript{55} Many TNCs operating in the extractive sector are financially more powerful than host African governments. In reality, the economic power of these TNCs

often dwarfs the power of host African countries.\textsuperscript{56} Thus, it is often the case that extractive TNCs use their power and influence to dissuade African governments from regulating their operations so as to enhance their profits.\textsuperscript{57} The fear that citizens of host countries could sue to enforce regulatory laws could potentially discourage TNCs from leveraging their power and influence to diminish enforcement of regulation by host African governments.

4. Impediments to Citizen Enforcement of Regulation

Although citizen enforcement of regulatory rules has produced positive results with regard to the regulation of extractive companies in Africa, it is doubtful whether such positive outcomes can be sustained due to the multiple impediments and hurdles that stand in the way of citizens desirous of enforcing regulatory laws against extractive companies. First, many African states participate directly in the exploitation of natural resources through contractual arrangements such as joint venture agreements, production sharing agreements and equity participation. For example, the Nigerian National Petroleum Corporation, a state-owned company, is engaged in joint ventures with major oil and gas companies operating in Nigeria.\textsuperscript{58} Similarly, the Ghana National Petroleum Corporation (GNPC) is statutorily authorized to “hold an initial participating carried interest of at least fifteen per cent” in the exploration and development of petroleum and it has the option to acquire an additional participating interest in petroleum operations.\textsuperscript{59} Consequently, the GNPC has entered into joint venture agreements with several oil and gas companies including Tullow; Sabre Oil and Gas; and Kosmos Energy.\textsuperscript{60}

In the hard rock mining sector, countries such as Botswana, Ghana, Guinea, Namibia, South Africa, and Zambia actively participate in the production of minerals. The government of Botswana owns a significant

\textsuperscript{57} Oshionebo, “Regulating”, \textit{supra} note 1 at 110.
\textsuperscript{59} Petroleum (Exploration and Production) Act of 2016, s 10(14) (Ghana).
equity interest in several diamond mining companies, including a 50 per cent stake in Debswana Diamond Company Limited,\(^{61}\) while the government of Namibia owns a mining company called Epangelo as well as a 50% stake in Namdeb, a diamond mining company.\(^{62}\) The government of Ghana is statutorily empowered to “acquire a ten percent free carried interest” in mining ventures in Ghana.\(^{63}\) The government of Ghana may in fact increase its participation in any mining venture with the consent of the holder of the mining lease.\(^{64}\) Similarly, the government of Guinea engages in mining activities “on its own behalf, either directly or through the Public Limited Company responsible for management of the mining patrimony acting alone or in association with third parties in the mining sector.”\(^{65}\) South Africa holds interests in mining projects through state-owned companies including Alexko; the African Exploration, Mining and Financing Corporation; and Industrial Development Corporation.\(^{66}\) In Zambia, the government has power to acquire mining rights over an identified area which “shall be reserved for government investment and shall not be subject to an application for the acquisition of a mining right by any person.”\(^{67}\) Mining rights so acquired by the Zambian government are vested in an investment company owned by the government.\(^{68}\)

State participation in resource projects could have unintended deleterious or chilling effects on citizen enforcement of regulatory rules in Africa. Given the direct involvement of African states in the exploitation of natural resources, private citizens could be fearful that litigation against companies involved in joint-venture projects with their government could trigger reprisals by the government. Reprisals could include arrest and detention of citizens and the filing of frivolous charges (sometimes referred to as Strategic Lawsuits against Public Participation) by government prosecutors.


\(^{63}\) Minerals and Mining Act of 2006, s 43(1) (Ghana).

\(^{64}\) Ibid, s 43(2).

\(^{65}\) Amended 2011 Mining Code, art 16 (Guinea), online: <http://bit.ly/2wVArt93>.

\(^{66}\) World Bank, supra note 62 at 17.

\(^{67}\) Mines and Minerals Act (Zambia), supra note 10, s 17(1), (2).

\(^{68}\) Ibid, s 17(3).
against citizen advocates.\textsuperscript{69} In addition, judges could be fearful of such reprisals by the government because, in many African states, judges need to be in the good books of the government to ensure career progression or advancement. Until very recently, Nigerian judges were reluctant to adjudicate matters involving the government and oil and gas companies\textsuperscript{70} and injunctions were rarely granted against oil and gas companies engaged in joint venture projects with the Nigerian government.\textsuperscript{71} While the justification offered by Nigerian judges for their refusal to grant injunctions against oil and gas companies was that such injunctions would cause a stoppage of trade and possibly lead to loss of employment,\textsuperscript{72} it is equally clear that Nigerian judges were desirous of protecting the economic interests of the government in oil and gas projects.\textsuperscript{73} Judicial decisions that protect the economic interests of the government would obviously endear the judge(s) to the government, thus ensuring career progression for the judge(s).

Secondly, the ability of private citizens to enforce regulatory rules in Africa is impeded by widespread illiteracy and poverty. Natural resource extraction projects are often located in remote and rural communities which are mainly inhabited by poor and illiterate people. Rural communities in Africa also lack the scientific knowledge to determine whether their environmental rights have been violated by extractive companies. Thus, these communities may not be aware that they could seek redress in the law courts.

Thirdly, unlike host communities which are predominantly poor and ignorant, companies operating in Africa’s extractive sector are financially, economically and technologically powerful. These companies often have more power than host African governments and are thus able to hire the services of the best lawyers and scientists to defend suits instituted against the companies.\textsuperscript{74} Given the enormous financial resources possessed by extractive companies, it is not surprising that lawyers representing these companies often intentionally engage in ‘long-winded litigation that essentially wears out


\textsuperscript{70} Frynas, supra note 13 at 144.

\textsuperscript{71} Ibid at 122-23.

\textsuperscript{72} Ibid at 122 (citing the case of Irou v Shell-BP, Unreported Suit No. W/89/71).

\textsuperscript{73} See Emeseh, “Limitations”, supra note 29 at 603 (asserting that the Nigerian judiciary prioritized the economic interests of the government over the protection of the environment). See also Frynas, supra note 13 at 122-23 (asserting that “the economic interests of the oil industry appeared to be more important to the judge than the course of justice.”)

\textsuperscript{74} Emeseh, “Limitations”, supra note 29 at 604.
the weaker opponent.” These lawyers often file frivolous motions and interlocutory appeals aimed at prolonging the litigation process.

Fourthly, the ill-effects of extractive activities could take several years to manifest and even where host communities subsequently become aware of the infringement of their rights, it may be too late for them to sue for legal redress as their cause(s) of action could become statute-barred under a Statute of Limitations. Jedrzej Frynas puts it best when he argues that

... a legal claim may become statute barred because of the latency period. The full effects of oil operations are not always immediately apparent. Damage such as long-term soil degradation requires a latency period for its development. The injury may not be immediately visible or may go undiscovered for a period of time. Conducting proper scientific studies may take many years to assess changes in vegetation or soil fertility, some of which can only be observed in the long-term. Because of the long latency period, potential litigants may not always have enough time to file a suit within the statutory period of limitation.

Zambia has attempted to ameliorate this problem by providing that the right of citizens to bring an action with regard to harm caused by mining or mineral processing operations lapses

... after a reasonable period from the date on which the affected person or the community could reasonably be expected to have learned of the harm or damage, taking due account of -

(a) the time the harm or damage may take to manifest itself; and
(b) the time that it may take to correlate the harm with the mining or mineral processing operations, having regard to the situation or circumstance of the person or community affected.

Generally speaking, ‘reasonable period’ is determined on an objective basis. However, the determination of ‘reasonable period’ under Zambia’s Mines and Minerals Development Act, 2015 also connotes a certain degree of subjectivity to the extent that the Act requires consideration of the specific “situation or circumstance of the person or community affected.” For example, the illiteracy or ignorance of a community adversely affected by mining operations would have to be taken into account in determining the period within which the community could reasonably be expected to have learned of the harm or damage.

---

75 Ibid at 603-4.
76 Maitland & Chapman, supra note 28 at 7.
77 Frynas, supra note 13 at 129-30.
78 Mines and Minerals Act (Zambia), supra note 10, s 87(6).
79 Ibid, s 87(6)(b).
The Zambian provision is laudable because it ensures that citizens are not prevented from seeking legal redress for mining-related environmental harms on the basis of civil procedure rules enshrined in the Statute of Limitations. African countries desirous of empowering their citizens to participate in the regulatory process would do well to emulate the Zambian provision.

Aside from general statutes of limitations, the statutes establishing state-owned companies in Africa may bar litigation against these companies in certain instances. For example, the statute establishing the Nigerian National Petroleum Corporation provides that

... no suit against the Corporation, a member of the Board or any employee of the Corporation for any act done in pursuance or execution of any enactment or law, or of any public duties or authority, or in respect of any alleged neglect or default in the execution of such enactment or law, duties or authority, shall lie or be instituted in any court unless it is commenced within twelve months next after the act, neglect or default complained of or, in the case of a continuance of damage or injury, within twelve months next after the ceasing thereof.  

Following this provision, any suit instituted against the NNPC outside of the statutory period of twelve months is improper and statute-barred. Given the scientific and complex nature of oil and gas activities, it could be difficult for host communities to detect violations of their environmental rights within twelve months. Besides, violations may not manifest within the 12 month limitation period because, as noted above, oil-related damage may require a latency period. Thus, host communities may not be able to sue to enforce their rights within the limitation period. The NNPC has successfully invoked this limitation period in a number of cases including the case of Eboigbe v NNPC, wherein the plaintiffs claimed that, because they were illiterates, they were not aware that pipelines constructed by the NNPC on their farmlands violated their legal rights. The suit was filed several years after the alleged violation, prompting the Supreme Court of Nigeria to hold that the action was statute-barred. The twelve months limitation period essentially insulates the NNPC against civil actions and explains “why little litigation has arisen against the NNPC for damage arising from oil operations” in Nigeria. This observation is significant given that, as noted earlier, state-owned companies

---

81 Frynas, supra note 13 at 129-30.
83 Frynas, supra note 13 at 129.
such as the NNPC are often involved in joint venture activities with extractive companies.

Fifthly, public interest litigation can only truly be said to complement the regulatory efforts of public agencies if the judiciary is independent of the executive arm of government. In many African countries the judiciary is hardly independent and, as noted earlier, judges rely on the government for career advancement and progression. Furthermore, there could be reprisals against judges whose decisions are perceived as ‘anti-government’ and such reprisals could manifest in very subtle forms such as transfer of judges to inhospitable jurisdictions; denial of promotion or elevation to higher courts; and withholding of funds to the judiciary altogether.

Sixthly, the cost of litigation in Africa is prohibitively high, thus very few Africans can afford the expense involved in public interest litigation. Many African citizens are either ignorant of their rights or, where they are aware of their rights, are too poor to afford the expenses involved in enforcing their rights. In addition, the litigation process in many African countries involves inordinately long delays which not only add to the cost of litigation but could also cause citizens to be disinterested in bringing their legal claims to court for adjudication.\textsuperscript{84} In these circumstances the right conferred on private citizens to enforce regulatory rules could become meaningless.

Relatedly, the fear that costs could be awarded against unsuccessful litigants could dissuade citizens from enforcing regulatory rules against extractive companies. A simple way to ensure that citizens are not dissuaded from seeking judicial enforcement of regulatory rules is to prohibit courts from awarding costs against private citizens and NGOs involved in public interest litigation. Zambia has adopted this position by statutorily prohibiting the award of costs against private persons involved in public interest litigation, even in instances where the action instituted by such private litigants proves unsuccessful.\textsuperscript{85} To that end, the \textit{Mines and Minerals Development Act 2015} provides that costs shall not be awarded against private citizens suing to enforce the environmental, health and safety standards under the Act “if the action was instituted reasonably out of concern for the public interest or the interest of protecting human health, biological diversity


\textsuperscript{85} Mines and Minerals Act (Zambia), \textit{supra} note 10, s 87(8).
and in general, the environment.”

Similarly, South African courts are enjoined not to award costs against private citizens who fail to secure the relief sought in respect of any breach or threatened breach of the provisions of the National Environmental Management Act, 1998 “if the court is of the opinion that the person or group of persons acted reasonably out of a concern for the public interest or in the interest of protecting the environment and had made due efforts to use other means reasonably available for obtaining the relief sought.”

Finally, although NGOs are actively engaged in public interest litigation in Africa and often help private citizens enforce their rights, the legal regimes in some African countries hinder the ability of NGOs and private individuals to initiate public interest litigation. In Nigeria, for example, there is the doctrine of *locus standi* which is often invoked by Nigerian courts as the basis for striking out public interest suits instituted by NGOs. *Locus standi* simply means the legal capacity or standing to institute proceedings in a court of law. In the Nigerian context, the doctrine of *locus standi* holds that, to be entitled to the standing to institute an action, a plaintiff must establish that they are directly affected by the act complained of and that they have a peculiar or personal right which has been infringed or which faces a real likelihood or threat of being infringed. Thus, in Nigeria, a general interest common to the public at large does not accord any particular citizen or group of persons the standing to sue to protect the interest.

The debilitating impact of the doctrine of *locus standi* on public interest litigation in Nigeria is apparent in *Oronto Douglas v Shell Development Company Limited and Others*, a case in which a private citizen sought to compel Shell to conduct an environmental impact assessment in relation to an oil project as required under the Environmental Impact Assessment Act. The court struck out the action on the grounds that the plaintiff had no standing to institute proceedings.

---

86 Ibid. See also *Environmental Management Act 2011 (Zambia)*, supra note 11, s 110(4).
87 *Environmental Management Act* (South Africa), supra note 7, s 32(2).
90 See *Adesanya v Nigeria*, *Olawoyin v AG Northern Region*, ibid; Oluyede, supra note 88 at 370.
the action given that the plaintiff’s personal rights were not infringed.\textsuperscript{92}

Equally disturbing is the decision in Amos v Shell Petroleum Development Company of Nigeria Limited\textsuperscript{93} where the court invoked the doctrine of \textit{locus standi} to deny the plaintiff’s access to justice. In this case, the defendants constructed a dam on a creek resulting in the flooding of the plaintiffs’ farmlands. The court held that the plaintiffs lacked standing to institute the action because, under Nigerian law, the creek was a public waterway. Thus, the Attorney-General was the appropriate party to institute an action in regard to the creek. The strictness of the doctrine of \textit{locus standi} in Nigeria has compelled a Nigerian academic to conclude that “the \textit{locus standi} rule effectively precludes access to court by public-spirited citizens and environmental NGOs seeking the enforcement of environmental regulations.”\textsuperscript{94}

The procedural hurdle of \textit{locus standi} is designed in part to ensure that ‘strike suits’ are eliminated from the judicial system. ‘Strike suits’ are suits instituted against companies for the primary purpose of coercing the companies into monetary settlement.\textsuperscript{95} In effect, ‘strike suits’ are frivolous suits that are often motivated by a desire to extract monetary compensation from companies. While it is arguable whether the empowerment of private citizens to enforce regulatory provisions could promote ‘strike suits’ against extractive companies, the reality is that, even in countries such as South Africa where citizens are empowered to enforce regulations, there is no evidence of a floodgate with regard to citizen litigation.\textsuperscript{96} ‘Strike suits’ are unlikely to emerge in Africa not only because of the prohibitive cost of litigation which impedes access to the courts,\textsuperscript{97} but also because, in most

\textsuperscript{92} It should be noted the Court of Appeal set aside the trial decision on the grounds that the trial judge committed an error by deciding the issue of \textit{locus standi} without examining the statement of claim and without allowing the plaintiff to adduce evidence in support of the statement of claim. See Orono Douglas v Shell Development Company Limited and Others, [1999] 2 N.W.L.R. (Part 591) 466 (Nigeria).


\textsuperscript{94} Amechi, \textit{supra} note 84 at 391. See also Emeseh, “Limitations”, \textit{supra} note 29 at 603. Emeseh asserts that the strictness of the doctrine of \textit{locus standi} in Nigeria “has stifled the development of public interest litigation in the country.”

\textsuperscript{95} See Michael B Dunn, “Pleading Scienter after the Private Securities Litigation Reform Act: Or, a Textualist Revenge” (1998) 84 Cornell L Rev 193 at 195.


\textsuperscript{97} \textit{Ibid} at 171.
instances, the statutes empowering private citizens to enforce regulations do not provide for monetary compensation even where a court finds a company in breach of statutory provisions. Thus, successful private litigants are not entitled to monetary compensation under these statutes. Rather, the remedies granted by the courts are non-pecuniary and restorative in nature.

That being said, there are specific instances where a statute may enable award of damages in public interest litigation. For example, damages could be awarded where a plaintiff proves that, although the action was instituted to protect public interest, they suffered adverse impacts as a result of the actions of the defendant. This appears to be the situation envisaged under Zambia’s Environmental Management Act 2011, which provides that the court may award damages to a plaintiff in order to compensate the plaintiff for any loss that they may have suffered or to remedy any adverse effect caused by the defendant’s act or omission.\(^98\) Private citizens may also be entitled to damages where the action of a company “constitutes an infringement of a constitutionally entrenched fundamental right”.\(^99\) The Constitutional Court of South Africa appears to support this view when it held in *Fose v Minister of Safety and Security*\(^100\) that a “loss occasioned by the breach of a right vested in the claimant by the Supreme law” ought to be compensable by damages. This is because section 38 of the Constitution of South Africa, which enables citizens to enforce environmental rights, vests discretion in courts to grant “appropriate relief” to litigants. ‘Appropriate relief’ would of course include monetary damages. While it is conceptually possible for citizens to recover damages for breach of constitutional rights in South Africa, such damages may not be awarded in public-interest litigation unless the claimant is able to establish that the injury or loss they suffered as a result of the breach is distinct from the injury or loss sustained by the public at large. It is equally doubtful whether South African courts would grant monetary damages where there are alternative remedies to rectify the breach in question.\(^101\)

---

98 Environmental Management Act (Zambia), supra note 11, s 110(1), (3).
99 Elmarie van der Schyff, “South Africa: Public Trust Theory as the basis for Resource Corruption Litigation” (Open Society Foundations, August 2016), online: <http://www.opensocietyfoundations.org/sites/default/files/legal-remedies-6-schyff-20160802_0.pdf> [van der Schyff].
100 1997 (3) SA 786 (CC) (S. Afr.).
101 van der Schyff, supra note 99.
5. Conclusion

This article examines statutory and constitutional provisions enabling private citizens to enforce regulatory provisions against companies in Africa’s extractive sector. It situates the utility of these provisions within the broader context of the empowerment of African citizens to participate in the regulatory process. It notes, however, that the utilization of these statutory provisions is hindered by a number of factors including state participation in resource extraction; widespread poverty and ignorance on the part of African citizens; and the prohibitive cost of litigation across the continent. These obstacles notwithstanding, it is imperative for African governments to diversify the mechanisms and avenues for regulating the activities of extractive companies given the incapacity of regulatory agencies in Africa. Such diversification can be achieved by empowering private citizens, local communities and NGOs to enforce statutory provisions, particularly those relating to the environment.
INTRODUCTION

With technological advances happening almost constantly, the world and the systems used to run it are changing. Processes are becoming automated and digitized, allowing organizations and individuals to operate complex systems run by computer and eliminating human error and bias. Given the complexity, cost, and time required to conduct a fair and representative election, it was only a matter of time until governments attempted to apply new technological advances to assisting in the running of a democratic system of voting. Thus far, however, few governments have even attempted to institute electronic voting systems, with even fewer implementing them successfully. While some countries, such as Estonia, have largely succeeded, creating an electronic platform to improve the convenience of voting has been difficult, with security being a primary concern.

One potential option that has been suggested for use in electronic voting (“e-voting”) is that of bitcoin and blockchain technology, a decentralized system of currency exchange that can be coded to exchange different digital indicators of value, from money, to ownership of fungible goods, to votes.
While there has been some controversy about the use of bitcoins, particularly given that they are a largely unregulated currency prone to wild market fluctuations, the unregulated nature of the technology is irrelevant to its potential to be used for voting and will therefore not be discussed in this paper. Instead, this document will describe the development and use of bitcoins using blockchain technology, the characteristics of a potential electronic voting system that would be required in order to make it functionally equivalent to current voting methods in Canada, and the applicability of blockchain and bitcoin technology to such a scheme.

The Development of Bitcoin and Blockchain Technology

Though there are now several different types of currency and transactional systems that utilize the blockchain (or similar) technology, blockchain technology was originally a decentralized payment system built to facilitate the transfer and use of bitcoins (BTC), a digital currency not issued by a central bank. Bitcoins, known as a ‘cryptocurrency’ because of the cryptographic software used to regulate the system through which they are transacted,¹ are “neither a commodity currency (backed by gold or some other commodity), nor a fiat currency (used by convention as a result of legal edict).”² As Abramowicz explains, bitcoins are not simply a currency based on the digital equivalent of dollars, but can also be used to “create and enforce property rights.”³ Indeed, bitcoins can be assigned non-monetary values in order to transact ownership of assets, such as stock in companies or real property, through the blockchain system.⁴

Bitcoins and blockchain technology were developed in order to allow verified peer-to-peer transactions without the need for a central server or verification and facilitation by a third party.⁵ Bitcoin technology (and digital currency in general) was developed as an alternative to server-based and third-

---

³ Ibid.
⁴ DutchChain, “The real value of bitcoin and cryptocurrency technology – the Blockchain explained” (14 October 2014) at 00h.2min.45s, online: Youtube <https://www.youtube.com/watch?v=YIVAluSL9SU>.
party-dependant commerce systems.\textsuperscript{6} Third-party-dependant commerce systems, such as banks, require complex processes to make and reverse transactions, resulting in inefficiencies such as higher transaction costs, minimum amounts for transactions, and an unnecessary cost associated with making “non-reversible payments for non-reversible services” since the cost associated with maintaining a complex system is distributed across a variety of transactions, including simple, one-way exchanges.\textsuperscript{7} While server-based systems cut down on the personnel requirements often employed by third party systems, if the server fails, all transaction records are lost and further transactions are unable to be processed.\textsuperscript{8} These downtimes result in high costs, with one estimate suggesting that $26.5 billion per year is lost due to server downtimes ($55,000 in annual revenue for small businesses, $91,000 for midsized business, and over $1,000,000 for large businesses) in addition to the negative effects technology failures have on employee morale and company reputation.\textsuperscript{9}

The reason why server systems and facilitation of transactions by third parties has thus far been the norm in e-commerce systems is due to the question of trust that arises in any transaction. Using a digital currency without someone to verify the transaction can result in the problem of double spending, where less-savoury individuals attempt to spend the same digital dollars twice.\textsuperscript{10} Typically, the third party or server (such as eBay or PayPal) will maintain digital ledgers for its customers; when one customer orders a transfer of funds to another, the third party will deduct the money from the sender’s account and move it to the receiver’s. Without a third party to verify and facilitate the transaction, the sender’s account does not get debited the amount owed to the receiver, giving the sender the chance to attempt to use the ‘already spent’ money a second time.\textsuperscript{11}

\begin{footnotes}
\footnotetext[6]{The Memo, “The Blockchain Explained” (11 November 2015) at 00h.00m.52s, online: YouTube <https://www.youtube.com/watch?v=wyfm92qqSh8> [The Memo].}
\footnotetext[7]{Nakamoto, supra note 5 at 1.}
\footnotetext[8]{The Memo, supra note 6 at 00h.00m.52s.}
\footnotetext[9]{Chandler Harris, “IT Downtime Costs $26.5 Billion in Lost Revenue”, InformationWeek (24 May 2011), online: <http://www.informationweek.com/itdowntimecosts$265billioninlosst revenue/d/d-id/1097919>.}
\footnotetext[10]{Jerry Brito & Andrea Castillo (2013) “Bitcoin: A Primer for Policy Makers” at 3-4, online: <https://www.mercatus.org/system/files/Brito_BitcoinPrimer.pdf> [Brito].}
\footnotetext[11]{Ibid.}
\end{footnotes}
How Do Bitcoins and Blockchains Work?

The solution to the issues posed by third-party-dependant and server-based systems is, ostensibly, a peer-to-peer transaction system based on “cryptographic proof of trust” that allows individuals to digitally transfer money and ownership of property without the middleman (or middle-server).\textsuperscript{12} Instead of being controlled by a central server or third party, the transaction ledger is distributed to all “nodes” in a decentralized global network of computers. The transactions are collected into a block of data, and nodes called ‘bitcoin miners’ compete to solve a complex mathematical equation in order to verify the transactions in the block who are then, if successful, paid for their efforts with newly-minted bitcoins.\textsuperscript{13} The first bitcoin miner to solve the equation verifies the transactions made since the last ledger update (that the parties have the correct information to proceed and enough currency to make the deal go through) and adds these transactions to the digital ledger as a “block” of information, creating a “linear sequence of linked data blocks” that serve as the ledger history for all transactions made on the system.\textsuperscript{14} Each block also records a ‘hash,’ a one-way software function that acts as a fingerprint of the data from the previous block of transactions, and a ‘nonce’, a random number that must be matched with the hash to satisfy the mathematical equation. This hash function is unable to be reproduced, making it easy to see whether a block has been tampered with and making hacks or other attempted modifications easy to detect (thus preserving the integrity of the chain). This is because in order to “to generate a new block, miners must find a nonce value that, when hashed with additional fields...results in a value below a given threshold”,\textsuperscript{15} and creating a false hash that satisfies the math problem when combined with the nonce is nearly impossible.\textsuperscript{16}

Once a miner has solved the problem required to verify the block, the winning miner transmits the data block to other nodes who did not win the mathematical competition, and they accept the addition of the block to the

\textsuperscript{12} Nakamoto, \textit{supra} note 5 at 1.
\textsuperscript{13} \textit{Ibid} at 4.
\textsuperscript{14} TEDx Talks, “Block Chain Revolution | Giovanna Fessenden | TEDxBerkshires” (20 July 2016) at 00h.4m.40s, online: <https://www.youtube.com/watch?v=oMhZTEQZPJi> [TEDx Block Chain].
\textsuperscript{16} \textit{Ibid}.
chain by verifying that the equation was solved correctly and that the transactions are valid (with correct public and private keys, and sufficient funds available to complete the transaction).

If a majority of the miners verify that the transactions are correct, the block is added and each node then updates its copy of the ledger accordingly. The mathematical calculations required to solve the problem also progressively increase in difficulty, so that “[a]s more processing power is dedicated to mining, the protocol will increase the difficulty of the math problem, ensuring that Bitcoins are always mined at a predictable and limited rate.” Since the maximum number of minable bitcoins has been capped at 21 million, this should ideally also prevent the devaluing of the currency through inflation (though it may be insufficient in protecting against deflation).

Satoshi Nakamoto, a pseudonym used by the originator of bitcoin and blockchain technologies, defined digital coins “as a chain of electronic signatures.” Being digital, each coin is a bit of code that “[e]ach owner transfers...to the next by digitally signing a hash of the previous transaction and the public key of the next owner and adding these to the end of the coin. A payee can verify the signatures to verify the chain of ownership.” An individual’s public key is akin to a username, while their private key is akin to a password. To transfer bitcoins, the sender will send the coins to the receiver’s public key, and “sign” it with their own private key.

---

17 Ibid
18 TEDx Talks, “New Kids on the Blockchain | Lorne Lantz | TEDxHamburgSalon” (24 March 2016), online: YouTube <https://www.youtube.com/watch?v=A1Vbrxkqjwc> [TEDx New Kids].
19 Brito, supra note 10 at 7.
20 Ibid. The last “satoshi” (0.0000001) of a bitcoin is expected to be mined in 2140, after which miners will be paid with transaction fees rather than new bitcoins.
22 Nakamoto, supra note 5 at 2.
23 Ibid.
24 Brito, supra note 10 at 5; Nakamoto, supra note 5 at 2 (diagram).
These transactions are also stamped with a hash that acts as a timestamp and confirmation of previous data in the chain. Since the hash function describes the data from the previous block, the information from the previous block must already be there in order to be included in the hash, thereby validating the previous data’s existence at the point in time at which the last block was added.\footnote{Nakamoto, supra note 5 at 2 (diagram).}

Nodes are programmed to always attempt to extend the longest chain, and not all chains need to receive every block for it to be accepted; since only a simple majority of nodes are required in order to accept a block,\footnote{TEDx Block Chain, supra note 14.} any of the rest of the nodes who miss a block can request their chains be updated once they receive subsequent blocks, and update their chains accordingly.\footnote{Nakamoto, supra note 5 at 4.}
The Non-Digital Value of Bitcoins: Are Bitcoins Worth Anything Away from a Computer?

The simple answer to whether bitcoins are worth anything anywhere other than the internet appears to be ‘mostly, no’. While banks, stock exchanges, and other institutions are exploring the potential uses for blockchain and bitcoin technology, since bitcoins are neither a commodity-based or fiat currency there are few places where bitcoins can be exchanged for actual cash dollars that can be spent at locations that do not take bitcoins as payment. This is distinct from other network-based electronic money transfer systems such as credit and debit cards. While those cards also work on an electronic system, they are backed by fiat or commodity-based currency, allowing them to be used interchangeably with regular money. Bitcoins, by contrast, can only be transacted in certain places, likely due to the reluctance of most mainstream retailers to accept an unbacked digital currency.

The main ways to buy and sell bitcoins with the result of receiving actual cash are through the use of certain websites, a company that acts a bitcoin broker, or by buying or selling coins to an individual by meeting in an online forum or in person. Additionally, some internet-based companies will allow you to pay for goods and services using bitcoins through the use of a digital wallet, or to put bitcoins towards digital reward cards that can be used on the websites for brick-and-mortar stores that do not currently accept bitcoins as payment. ‘coloured coins’ may be used to assign something other than monetary value to bitcoins. The use of coloured coins “allows [the] attaching metadata to Bitcoin transactions and leveraging the Bitcoin infrastructure for issuing and trading immutable digital assets that can represent real world value.” Essentially, small packets of data can be

28 See Daniel Gasteiger’s speech for examples of potential future uses of this technology. TEDx Talks, “Blockchain Demystified | Daniel Gasteiger | TEDxLausanne” (25 April 2016) at 00h.5m.35s, online: YouTube <https://www.youtube.com/watch?v=40ikEV6xGg4>.
30 Alexandra Posadzki, “ATMs that swap cash for Bitcoins coming to Canada this fall”, The Canadian Press (8 September 2013), online: <http://www.ctvnews.ca/canada/atms-that-swap-cash-for-bitcoins-coming-to-canada-this-fall-1.1445199>.
attached to bitcoins and the transfer of that data can be equated later to the transfer of physical assets.

Advantages to the Blockchain System

There are several advantages of the blockchain system of currency transfer. The first is that the system provides “immutable preservation of data integrity.”

Due to the fact that each transaction is imprinted with the hash information of the previous block, when combined with the consensus mechanism required to approve blocks it becomes nearly impossible to change or falsify information in subsequent blocks. To do so, the attacker would have to modify a block; this would reflect in the hash of the next block, and affect all subsequent blocks. Since a majority of nodes must accept the block of transactions as valid in order for the block to be added to the blockchain, the attacker would essentially have to change the desired block and then create a false chain longer than the ‘real’ chain on every computer that has access to the chain around the world, beginning from the block that they want to modify. While making such a change is theoretically “possible, [it is] computationally impractical” to do so.

Furthermore, if hackers attempt to subvert the system by generating a second, competing block chain – essentially creating a new block from which nodes would continue to branch off, thereby creating a new chain branch from a certain point rather than attempting to change aspects of the data in an already-existing blockchain – he or she would have to do so faster than other, legitimate chains were being built, since nodes always attempt to extend the longest chain. For example, if there were ten blocks in the blockchain, and the hacker wanted to change block six, he or she would have to create blocks to replace blocks six through ten (with proper transactions and authentication), and then attempt to create an eleventh block, all before the original chain added its eleventh block. Nakamoto further suggests that even if a hacker were successful in creating an alternate chain, the only thing the hacker would be able to accomplish is undoing previously-made transactions, as legitimate nodes will recognize invalid transactions (for

33 TEDx Block Chain, supra note 14 at 00h.04m.05s.
35 TEDx Block Chain, supra note 14 at 00h.07m.25s.
36 Ibid.
37 Nakamoto, supra note 5 at 6-7.
example, transactions that transfer money that does not exist) and refuse to accept blocks containing them.\textsuperscript{38} A hacker would therefore be unable to spend bitcoins that are not his or her own.

The relevance of this level of security and data integrity as it applies to voting is that it makes it difficult, if not impossible, to corrupt the transactional chain and artificially falsify voting transactions after they have occurred. Going back and changing someone’s vote would be extremely difficult once voters have established their vote on the blockchain, as it would likely require changing the transaction (either by changing the destination of the vote or by removing the vote completely) as discussed above. In fact, Lorne Lantz has suggested the use of e-voting over the blockchain as a way for corrupt governments with a history of rigging elections to regain the support and trust of their voters.\textsuperscript{39}

Other advantages provided by the blockchain system are those provided by having a distributed, decentralized system.\textsuperscript{40} Having a decentralized system rather than a system with a centralized server and access point nodes inherently protects against server outages and IT issues that occur as a result of many people attempting to access a specific server at the same time. This is a situation that seems likely to occur during the electronic voting process when people all over the country attempt to access the same server (or set of servers) at the same time to cast their votes. A decentralized system would also protect against Denial of Service (DoS) attacks like the one experienced by the NDP in 2012, wherein the NDP’s server was “bombard[ed]… with repeated attempts at communication to try to slow it down or crash it altogether”, leaving delegates unable to cast votes for the next NDP leader.\textsuperscript{41}

In the case of a blockchain system, if one node in the blockchain system fails, the remaining nodes continue processing data independently of the downed node, unlike when a central server goes down, and takes the rest of the system with it until it can be repaired or restarted. Furthermore, once the downed node comes back online, it can ‘catch up’ to the other nodes on the blockchain and update its copy of the ledger once the missing blocks in the chain are detected as missing. This should allow for relatively uninterrupted system access for voters, as even if one computer were to go down voters

\textsuperscript{38} Ibid.
\textsuperscript{39} TEDx New Kids, supra note 18 at 00h.13m.01s.
\textsuperscript{40} TEDx Block Chain, supra note 14 at 00h.07m.52s.
could continue voting (making transactions) while other computers continued forming blocks and verifying transactions (votes placed). The blockchain system could possibly face efficiency issues here, however, since voters could need to wait from several minutes up to an hour – the time it takes the miners to solve the mathematical problem allowing them to add blocks to the chain, verify transactions, and then to build on additional blocks to ensure the transaction will not be lost due to a fork in the chain – to ensure that their votes had gone through and had been successfully recorded.

A third advantage of using blockchain technology is that it solves the problem of validating user identity without requiring the user to provide personal information. Typically, financial institutions that facilitate transactions between parties require the parties to provide a certain amount of disclosure as to their identities in order to verify that they are who they claim to be and they have the funds they wish to transact. For example, PayPal may require you to connect a bank account in addition to proving other information such as name, address, and credit card information. While this information may or may not be shared with the other party, parties cannot remain completely anonymous due to the facilitating party’s knowledge of their identifying information. Furthermore, even with security measures in place, identity theft remains a major issue in Canada, with almost $10.5 million in losses recorded from over 20,600 people in 2014. A study funded by the U.S. Department of Justice spoke to individuals incarcerated for identity theft, and found that “[r]egardless of [the offenders’] chosen lifestyle, they were primarily motivated by the quick need for cash and see identity theft as an easy, relatively risk-free way to get it.”

Of particular interest is one anti-theft quality of blockchain systems that can protect identities both in regular commercial transactions and when voting: the fact

42 “Frequently Asked Questions – The Verification Process” (2016) PayPal (website), online: <www.paypal.com/ca>. Because PayPal allows anyone to buy and sell, even individuals wanting to remain solely as buyers may be required to verify their accounts after transferring a certain amount of money.


that the blockchain IDs used in transactions can be assigned anonymously, without the need for identity validation.\textsuperscript{45}

In contrast to regular transactions wherein the nature of the transaction and identity of the parties are known to a third party but kept from the public, the blockchain system makes the transactions public, while making it possible to keep the users almost completely anonymous,\textsuperscript{46} or by incorporating pseudonyms where it is necessary for an individual to identify themselves (i.e. to ensure that the same person who is spending bitcoins is the person who initially received them).

\textbf{Figure 3. Comparison of privacy models}

![Figure 3. Comparison of privacy models](image)

Bitcoin users create bitcoin addresses through the use of ‘wallets.’ These wallets typically generate a new bitcoin address for every potential transaction, since “once addresses are used, they become tainted by the history of all transactions they are involved with.”\textsuperscript{47} Since anyone can see the address to which bitcoins were sent, it is important to generate a new address for every transaction in order to avoid building transaction history available to the public.

Bitcoin addresses are generated from a combination of a public and private keys. Since the addresses are randomly generated and only used once, it is nearly impossible to track certain keys to certain bitcoin users, though this does not account for other methods of identification, such as IP address logging\textsuperscript{48} or the utilization of bitcoin transfer websites that require you to provide personal information.\textsuperscript{49} Furthermore, repeated transactions using

\begin{footnotes}
\item[45] Nakamoto, supra note 5 at 6.
\item[46] Ibid.
\item[47] Copes & Vieraitis, supra note 44.
\item[48] “Protect your privacy” Bitcoin.org (website), online: <https://bitcoin.org/en/protect-your-privacy>.
\item[49] Brito, supra note 10 at 8.
\end{footnotes}
bitcoins can be utilized in analyses that may allow multiple similar transactions to be linked to the same user through the amount and type of transaction, or location of the transaction. In a study by Androulaki et al., the authors estimated that “behaviour-based clustering techniques can unveil, to a large extent, the profiles of 40% of Bitcoin users, even if these users try to enhance their privacy by manually creating new addresses.”

While these results do present concerns about the level of anonymity and identity protection that is actually achieved relative to the amount that is promised, the parameters of the study included an assumption that the individual attempting to connect profiles had some familiarity with the area and circumstances in which transactions took place, and was basing their analysis on 25 transactions per user. Since individuals would likely be casting far fewer than 25 votes at a time if blockchain technology were employed in electronic voting, the potential for identification of a user would be ameliorated by the lack of information; however, given the importance of total anonymity in the voting process, this concern should not be taken lightly.

Incorporating Blockchain Technology into the Canadian Voting Process

In their report “Establishing a Legal Framework for E-Voting in Canada”, Schwartz and Grice describe several “attributes and values that Canadians currently have under the paper-based system” that an electronic voting system should emulate or to which it should be functionally equivalent in order to justify the use of such systems for Canadian elections. Whether blockchain technology may be an appropriate solution for the concerns described will be addressed for the majority of the points. While the report focusses on both the requirements needed to be fulfilled by electronic voting in general and characteristics of a legislative framework to implement it, this paper will focus on the former. The modified list of ideal characteristics is:

facilitated accessibility and reasonable accommodation, voter anonymity..., accurate and prompt results, comprehensible and transparent processes, system security and

50 Androulaki et al, supra note 15, s 7.
51 “[W]e assume that A can have access to the (public) addresses of some vendors along with (statistical) information such as the pricing of items or the number of their clients within a specified amount of time.” Ibid, s 3.1.
53 The list has been edited for topicality and brevity.
risk assessment, detection of problems and remedial contingencies..., effective and independent oversight, and cost justification and efficiency.\textsuperscript{54}

Essentially, a blockchain-based voting system would require some type of confirmation of voter eligibility, after which voters would be assigned a random key pair. Voters would then be able to vote by sending small amounts of bitcoins to their desired candidates. Finally, votes will be secured on the blockchain and a tally will be computed. Overall, a blockchain-based voting system should include all of the advantages of blockchain and bitcoin technology while complying with the values outlined by Schwartz and Grice as being necessary for the construction of an effective electronic voting system.

**Facilitated Accessibility and Reasonable Accommodation**

The first requirement of internet voting is that it be as or more accessible to voters than current paper-based voting methods. While on its face internet voting is commonly touted as having the potential to be more accessible to Canadians than paper-based voting due to the pervasiveness of the internet, important and novel obstacles arise when attempting to implement and carry out internet-based voting. One of the main arguments against the adoption of internet voting is due to the presence of a “digital divide”, a divide in accessibility to voting services due to the inequity of internet availability to certain segments of the population.\textsuperscript{55} While factors such as one’s location, gender identity, and level of education do not seem to have strong effects, familiarity and amount of computer savvy and age are important factors, with younger individuals being far more likely to vote online than their older counterparts.\textsuperscript{56} Furthermore, a report by the Canada-Europe Transatlantic Dialogue cited a variety of reasons for further divide, particularly based on economic inequality for users with cheaper, more easily accessible computers and internet connections.\textsuperscript{57}

While ideally a blockchain-operated system would be relatively easy-to-use (in order to minimize the learning curve for those unfamiliar with using certain types of technology), if the voter is unable to transmit their vote through the system due to a faulty or non-existent internet connection, the

\textsuperscript{54} Schwartz & Grice, supra note 52 at 6-7, 29-30.


\textsuperscript{56} Canada-Europe Transatlantic Dialogue, A Comparative Assessment of Electronic Voting, (Ottawa: Elections Canada, 2010) at 18 [Canada-Europe].

\textsuperscript{57} Ibid at 16.
efforts put into making a simple system will be moot. While internet connections may be available at polling stations, it is conceivable that the connection would slow down with everyone attempting to transmit their vote. Further, while there exists mobile technology (such as using iPads over a 3G or 4G network) that could bring the internet to those who do not have access to it despite being in an area where internet is available, access to these systems may also be compromised in rural or otherwise out-of-service areas. Schwartz and Grice suggest that it is necessary to consider measures that will help provide increased access to required technologies, whether it be by allowing people to use their personal electronics or by providing resources such as help lines to assist those unfamiliar with using the internet. In any case, non-electronic means of voting should also be available for those uncomfortable voting through electronic means, to make sure all Canadians feel comfortable exercising their right to vote.

**Voter Anonymity**

Schwartz and Grice emphasize the requirement to protect voter anonymity and ensure that voters are not being subjected to undue influence when they go to vote. In a process familiar to anyone who has participated in an election, initially voters are required to identify themselves to preceptors at polling stations, who confirm their eligibility to vote and hand them a ballot. The ballot contains no personal information, and once the individual records his or her vote, the ballot is put into a secured ballot box to be counted once polls close. In contrast, absentee voters receive and submit their ballots by mail. They receive a blank ballot in an envelope, which itself is in an outer envelope. The voter marks their ballot, encloses it in the inner envelope, and then signs the outer envelope – the voter’s signature and information on the outer envelope will allow Elections Canada to verify the voter as someone eligible to vote, and the inner envelope keeps the individual verifying the information on the outer envelope from seeing how the person voted.

A similar double envelope process using public key cryptography has been used for electronic voting in the past; essentially, the voter would “pack” their vote into a ciphertext that would be encrypted in a bitcoin “envelope” using Elections Canada’s public key. The voter would then “sign” an outer envelope

---

58 Schwartz & Grice, *supra* note 52 at 18.
59 *Ibid* at 29.
using their private key, which would be used by Elections Canada to identify
the voter as someone eligible to vote. Before processing, the private key would
be “stripped off” and subjected to other encryption methods to ensure
Düsseldorf, 2015) [unpublished] [Meter]. This paper also contains an extensive review of
various types of electronic voting systems, as well as their pros and cons.} While this specific electronic process is mentioned due to being
particularly analogous to current paper-based voting methods, other methods
have been investigated and summarized in literature on the topic.\footnote{Ibid. See also the discussion in Canada-Europe, supra note 56.}

Another, more troubling aspect of online voting is its potential to
courage the exercise of undue influence.\footnote{Schwartz & Grice, supra note 52 at 156.} This issue pervades many types
of voting, and could be particularly prevalent when an individual is in an
unmonitored environment – for example, their home or workplace.\footnote{Ibid. See also Timothy Frye, Ora John Reuter & David Szakonyi, “Political Machines at
Politics 195.} The issue is also more likely to arise where individuals are allowed to vote
unmonitored, but may only cast one ballot.\footnote{Meter, supra note 60 at 61.} This is an inherent issue with
the use of blockchain technology, and may be one of the biggest barriers to its
adoption along with the difficulty of bridging the digital divide. While it is
theoretically possible to override transactions before they are accepted by the
blockchain (by double-spending your bitcoins, or in this case, ballotcoins and
appending a higher transaction fee to your latter purchase and hoping that in
doing so miners are encouraged to process your latter transaction first, leaving
you with no ballotcoins with which to complete the first transaction),\footnote{“How to Cle
ar a Stuck Bitcoin Transaction” (29 July 2015), Bitzuma, online:
<bitzuma.com/posts/how-to-clear-a-stuck-bitcoin-transaction/>.} it is
difficult to do so. This problem has been ameliorated in Estonia, where
individuals are allowed to cast multiple ballots with only the last one
counting,\footnote{Meter, supra note 60 at 25.} but this could be difficult for blockchain voting methods where
private keys are anonymized for every transaction (using the same private key
is unwise because it risks voter anonymity). Anonymizing private keys brings
its own challenges, however, if the new votes come from different keys,
election officials will be unable to tell which votes are true and should be
counted. In theory, allowing multiple votes should assist an individual in
being able to cast a vote at a time when they are not in the presence of the

\footnotetext[60]{Christian Meter, Design of Distributed Voting Systems (MA Thesis, Heinrich-Heine-Universität
Düsseldorf, 2015) [unpublished] [Meter]. This paper also contains an extensive review of
various types of electronic voting systems, as well as their pros and cons.}
\footnotetext[61]{Ibid. See also the discussion in Canada-Europe, supra note 56.}
\footnotetext[62]{Schwartz & Grice, supra note 52 at 156.}
\footnotetext[63]{Ibid. See also Timothy Frye, Ora John Reuter & David Szakonyi, “Political Machines at
Politics 195.}
\footnotetext[64]{Meter, supra note 60 at 61.}
\footnotetext[65]{“How to Clear a Stuck Bitcoin Transaction” (29 July 2015), Bitzuma, online:
<bitzuma.com/posts/how-to-clear-a-stuck-bitcoin-transaction/>.}
\footnotetext[66]{Meter, supra note 60 at 25.}
person wishing to exert influence, even if they had been forced to vote previously, but systems would need to be designed to accommodate this.

**Accurate and Prompt Results**

Promptness, and particularly accuracy, of results are crucial to the voting process. Schwartz and Grice stress the importance of making sure sufficient time is allowed for electronic votes to be cast (and if necessary, recast) and for election regulators to respond to technical difficulties.\(^{67}\) Furthermore, it is important that individuals have faith that the voting system will accurately record their vote. Schwartz and Grice suggest that individuals trust paper ballots due to their track record over the extensive period of time during which they have been in use.\(^{68}\) This theory is corroborated by Smith, who found that voters feel most confident in the accuracy of paper-based voting systems, and least confident in mobile-phone-based voting systems.\(^{69}\) According to Blockchain Technologies Corp. CEO Nick Spanos, currently paper ballots are the only truly accurate methods for counting elections, since in that case the actual, physically marked ballots are available, whereas if the validity of election results calculated by a machine are called into question the only real option for determining whether there was an error is by looking at the software’s source code.\(^{70}\)

Unlike other methods of electronic vote calculation that utilize software to record and count the votes, the blockchain system does not rely primarily on specialized software to calculate votes, but instead counts each vote as a transaction that must be validated by a majority of network users. With a large number of people required to approve each block of transactions, erroneous transactions should be quickly weeded out, helping to ensure the validity of each vote. Furthermore, each blockchain transaction provides the sender with a receipt, allowing them to make sure that their vote has been sent to the particular candidate. Therefore, while blockchain technology does suffer from the slight stigma affecting electronic voting methods, it is likely

---

67 Schwartz & Grice, *supra* note 52 at 8.
better equipped than many systems to provide voters with perceptions of accuracy and peace of mind.

Promptness may be another issue with blockchain technology. Since each block takes about ten minutes to update, and there are a finite number of transactions in a block, it may take an extended period of time to confirm whether someone’s vote has been received. While this is less likely to be a problem for individuals voting at home or at work (assuming the transaction succeeds), if an individual travels to a polling station because they do not have internet capabilities at their home they may be required to wait for confirmation that their vote has been processed. While not ideal, the use of electronic calculation may mean that votes will be recorded and counted much more quickly than they would be if they had been processed by hand.

Comprehensible and Transparent Processes

Schwartz and Grice emphasize the point that “transparency is directly related to the amount of information that is available to the public, as well as to intermediaries.”71 Ironically, while one of the key aspects of blockchain technology is that it is nearly completely transparent, its origin is shrouded in secrecy. The seminal paper creating bitcoin and blockchain technology, “Bitcoin: A Peer-to-Peer Electronic Cash System”72 was written under the pseudonym Satoshi Nakamoto; at this point, the paper’s original author has yet to be confirmed. Previously, this paper discussed Schwartz and Grice’s observation that not only do people have to know how to use an electronic voting system, but they also need to believe that the system will accurately and efficiently record their votes.73 Whether or not the level of uncertainty surrounding the origin of blockchain technology will diminish voter faith in the blockchain system as a potential electronic voting method remains to be seen. While it is possible that the secrecy may present warning signs to some (both individuals with low-level computer skills who cannot bring themselves to believe in a system whose origin cannot be traced to a certain individual, and those tech-savvy individuals who see blockchain technology as a

71 Schwartz & Grice, supra note 52 at 41.
72 Nakamoto, supra note 5.
73 “Even though one can look at technological attempts to minimize the risks associated with Internet voting, it is crucial to recognize that it is not sufficient to render the voting system intrinsically trustworthy. The system must in practice be trusted by the Canadian public, and not only by a particular group of government bureaucrats or information technology specialists who work on the project.” Schwartz & Grice, supra note 52 at 18.
potentially risky system to base the election on), others may be satisfied simply knowing how the system works and how to use it, albeit with computer-literate people wanting additional details on the process.

Despite the technology’s secretive origins, blockchain technology is itself extremely transparent in that it relies on a public ledger of all transactions made on the system where anyone – including election officials, technical analysts, or the average voter – can view which transactions have gone through. However, just because the data is available does not mean that it is of any utility. The blockchain method of transferring bitcoins is a relatively new concept. Its multi-layered complexity may stymie the average observer from making use of the data, and the idea that people’s votes would be publically viewable by anyone in the world and would be protected by the ‘magic’ of the blockchain system may be difficult to justify to individuals who are not technologically savvy. That being said, this can be remedied by hiring individuals familiar with blockchain technology to check the data and report on trends that Canadians would like to know about. Information about the system should be made public, but in its simplest terms, voters would simply need to know how their vote was being kept secure, and how to actually use the system to vote.74 Of the various barriers to implementation that legislators would need to overcome, public faith in (and comprehension of) the system is not insignificant, but it is also not insurmountable.

System Security and Risk Assessment

System security is one of the primary concerns with any form of electronic voting, but blockchain technology is unique in that its security comes through its transparency. While other countries have restricted access to their program source codes or made only certain parts of their source code available to public scrutiny,75 nothing would be withheld on the blockchain system except the identities of voters, which would be coded by private keys. In a typical transaction a new private key would be generated by an individual’s wallet for each new transaction; since Elections Canada would need to verify that the person voting is eligible to do so, an alternative method of generating a private key may need to be used. For example, in Estonia, keys are stored in a “tamper-resistant hardware security module” and can only be accessed via a

74 Ibid at 44.
75 Schwartz & Grice, supra note 52 at 47.
password from a quorum of the National Electoral Committee. Alternatively, allowing a random private key to be generated and then using the receipt of the vote as confirmation could suffice, though this may cause issues with knowing which votes to disregard if multiple votes are allowed.

A major security concern is the ability to keep the identities of voters private while still making sure they are eligible to vote. While one study claims to have been able to show that user’s profiles can be discerned under certain conditions, the study relied on 2011 statistics that estimated the number of bitcoin users to be around 60,000 and that the user seeking to identify transaction makers had access to certain location-specific information. Since the number of bitcoin users has grown exponentially to over 10.1 million users as of November 25th, 2016, it is likely to be considerably more difficult to discern the locations of users, particularly if users across the country are able to vote at their convenience over a period of time.

The blockchain system’s transparency relative to other systems also has other advantages, though admittedly the level of transparency demanded by the blockchain system raises other questions about how to keep voting tallies private until votes can be verified and counted. The fact that the system’s data is publically available would allow anyone with familiarity with the system to monitor transactions and report errors to Elections Canada. Even if an individual with less-than-honest intentions was to attempt to exploit vulnerabilities in the system, the system would be significantly resistant to hacking due to the computational difficulty of being able to rewrite blocks with enough speed to update the blockchain across all of the computers in the network. Should an attacker attempt to override the system, he or she would be limited to “undoing” their vote, rather than being able to cancel, make, or “steal” other votes. Given that security and transparency are two important characteristics of an electronic voting system, the fact that the

---

78 Ibid, s 3.1.
79 The number of blockchain users can be estimated by the number of blockchain wallets being used, since while users may have more than one wallet, they must have at least one wallet to utilize the system. See “Blockchain Wallet Users” (Graph), online: <https://blockchain.info/charts/my-wallet-n-users>.
blockchain system is secure because of its transparency could make it a strong choice as a non-proprietary electronic voting system.

Detection of Problems and Remedial Contingencies

One issue that might arise from the use of bitcoins for voting is if a voting transaction becomes “stuck” in transfer. Because bitcoin miners choose which transactions they process, transactions with higher transaction fees are more likely to get processed quickly, while transactions with lower fees assigned will be processed more slowly, or not at all. This is because transactions without fees are often considered “spam” transactions that will not be approved and added to the blockchain. So long as a transaction fee of at least one satoshi per byte (0.00000001 BTC) is applied, a transaction should go through, though if the transaction fee is lower it may take longer for votes to be registered. Additionally, there is the issue of transaction fees. If a truly immutable blockchain is desired, it should be built on a public blockchain that requires a fee to use – a fee that would need to be borne by the Government of Canada. If the government was to foot the bill for casting bitcoin votes on a public blockchain, questions still arise as to how to distribute the funds required.

Since the registration of votes is of paramount importance, an e-voting system should allow voters themselves to confirm whether their vote has been registered. Should it be decided that voter receipts are a good idea, the provision of such receipts could be done through a bitcoin wallet app (which people would need in order to vote anyway) or through other means – Norway has created a notification system that allows people to verify their votes through verification codes transmitted via text message. This potential delay is also a compelling reason to follow Schwartz and Grice’s suggestion of having a week-long voting period that ends prior to the last day of ballot casting. Allowing a buffer period will ensure that votes that get stuck will have extra time to go through, increasing the probability that all votes cast electronically will be counted. Furthermore, if multiple votes were allowed as


81 Schwartz & Grice, supra note 52 at 57-58.
they are in Estonia, having a buffer period of several days before election day would allow officials to reconcile multiple votes.\textsuperscript{82}

In terms of remedial contingencies, Schwartz and Grice state that electronic voting should work at most in tandem with physical ballot-casting methods, and non-electronic voting options should be available at all times.\textsuperscript{83} Given the permanent nature of blockchain-based transactions, strategies for reconciling multiple votes could be necessary in order to mitigate the risk of failed voting transactions, with the added benefit of reducing the risk of vote coercion. If people can attempt to vote both online and in person, as in Estonia,\textsuperscript{84} permitting multiple votes would allow those who attempt to vote online unsuccessfully to still register a vote through traditional voting means. This may require adding a level of electronic oversight to traditional voting methods to determine which vote of many made by an individual is the most recent (i.e. if the latest vote is the only one counted, we must be able to determine which was the latest vote cast). However, given the permanent nature of blockchain transactions the addition of the electronic component may be necessary to mitigate the risk of failed votes by allowing multiple voting opportunities.

**Effective and Independent Oversight**

One issue that arises is the lack of oversight that necessarily comes with people using internet voting. While the government can create and put out apps to facilitate electronic voting and ensure that everyone is operating on the same base software (e.g. the government could create a bitcoin wallet app specifically for voting), the reality is that unless everyone’s personal machine is inspected and secured, there will necessarily be less oversight of the voting process and equipment in an uncontrolled voting environment than in the traditional system where everyone votes in a controlled setting.\textsuperscript{85} In order to mitigate risks associated with unobserved voting (such as coercion or the use of malware on voters’ devices to interfere with the voting process), Schwartz and Grice recommend stringent testing of the system before voting and auditing afterwards, and suggest that instead of running e-voting as similarly

---

\textsuperscript{82} *Ibid* at 40.
\textsuperscript{83} *Ibid*.
\textsuperscript{84} *Ibid*.
\textsuperscript{85} Schwartz & Grice, *supra* note 52 at 18.
as possible, that special oversight rules be created to reduce risks specifically associated with voting by electronic means.\textsuperscript{86}

In addition to testing and retesting the technology itself, Schwartz and Grice suggest that a board comprised of members with “technical expertise, independence, reliability and multiparty support” be created in order to oversee the implementation of electronic voting.\textsuperscript{87} Given the complexity of the blockchain system, it would be wise to assign a board to direct implementation and oversight of the system.\textsuperscript{88} Doing so would ideally “help address concerns that may arise from the novelty and technological complexity of conducting an election partly by Internet”.\textsuperscript{89} Appointing a board supported by various political parties, members of the judiciary, and experts on the use of blockchain technology would also greatly assist in building Canadians’ trust of the internet voting system, which Pieters and Becker say is necessary if online voting methods are to be adopted.\textsuperscript{90} It will also ensure that the system as a whole is scrutinized for inconsistencies, errors, and compliance with current election practices and legislation, and make sure that parts of the system do not get overlooked. Since peoples’ levels of confidence in electronic voting systems still fall behind their confidence levels in paper-ballot-based voting,\textsuperscript{91} having a diverse board is particularly necessary in order to reduce the impression of overreliance on technicians and programmers caused by the complexity of the system, thereby lending the system greater credibility as an alternative voting method.

Cost Justification and Efficiency

The final major factor that would require consideration before implementing a blockchain-based electronic voting system is the overall cost of implementing the system. The type of equipment used in electronic voting greatly impacts the cost to run such a program, and will vary widely depending on whether the program and equipment is run in a controlled or

\begin{flushright}
\textsuperscript{86} Ibid at 62, 65. \\
\textsuperscript{87} Ibid at 66. \\
\textsuperscript{88} Ibid. \\
\textsuperscript{89} Ibid. \\
\textsuperscript{91} Smith, supra note 69 at 68, 80-81.
\end{flushright}
uncontrolled voting environment. In a controlled voting environment, all of the machines people use are provided to voters. Before, during, and after the election the machines need to be stored and maintained, adding these costs to the cost incurred when machines were initially acquired. Different machines have different costs related to their upkeep and how they interact with the voting system; for example, depending on what kind and how many machines the e-voting system uses, costs could vary wildly. Further, because the blockchain system requires transactions to be uploaded and disseminated via the internet, taxpayer dollars would also need to be put towards setting up Wi-Fi or mobile hotspots in areas where no internet exists. This would potentially necessitate the purchase or rental of routers and modems, as well as a deal with internet services providers to provide fast, reliable internet to polling sites. Furthermore, once the election ends this equipment would either need to be disposed of or stored and maintained, incurring continuous costs.

In contrast, e-voting in an unsecured environment could cut down on equipment costs, since individuals would be able to use their personal internet-connected devices. While Elections Canada would have the option of developing a special voting app that would act as a bitcoin wallet for transacting votes, it would not be necessary for them to do so as long as they utilized a previously-established cryptocurrency that operated on an existing system (like bitcoins). The purchase, storage and maintenance may still be required; for example, accommodation may be required for individuals without access to internet-linked computers. Furthermore, if people were allowed to cast multiple votes (including at least one paper ballot) in order to mitigate the risk of coercion and votes not being registered in a timely manner, some electronic equipment would have to be purchased in order to keep track of which votes were cast when. Lastly, the implementation of an electronic voting system would require the hiring of various experts in blockchain technology, e-voting systems, computer maintenance, and electronic security. While these costs may decrease once the system is in place and less troubleshooting is needed, one cannot deny that personnel costs associated with developing and implementing an electronic system would likely be high. It is unclear whether the costs of implementing an e-voting system would be outweighed by its utility.

93 Schwartz & Grice, supra note 52 at 19.
Implementation

In brief, blockchain technology is a type of technology for facilitating electronic transactions that has the potential to be used as a platform for electronic voting in Canada. While the technology already exists and is theoretically useable by anyone with internet access, time and patience will be required in order to bring this technology to the point where it could be used elections on a wide-scale.

A major obstacle that would need to be overcome is the opacity of the system. Regardless of one’s opinion on the propriety of using an unregulated currency transfer system created by an anonymous originator, the fact remains that the blockchain system of transfer is a fairly abstract concept and may be difficult to comprehend for the average voter. The question that needs to be asked is: do voters care how the system works? The answer for many may be “no”, and so long as they are informed enough to use the system and understand that it is secure, that may be sufficient rationale for the majority of the electorate to get behind the technology. For those who want more information on the system and how it factors into the voting process, more information could be provided, particularly on a website or through pamphlets and other advertising. Additionally, telephone or webchat support resources could be created for people who still have questions. The main information to impress upon voters is how the system keeps votes confidential and what the electronic voting process looks like.

In addition to support available before voting, technical support must be available at the polls as well. Given the digital divide precluding some people from accessing the internet, Elections Canada may feel the need to provide Wi-Fi zones or even internet access points (computers connected to the internet) at polling stations for those individuals who wish to vote online but would not otherwise have internet access. Providing these kinds of resources (as well as resources to accommodate individuals who may be housebound or otherwise unable to go to the polls) will require purchase, maintenance, and storage of physical equipment (such as modems, wireless routers, and computer equipment), as well as the purchase of internet access from a local service provider sufficient to meet the demand of voters. Ideally the temporary strain of an election on the wireless network could be reduced by allowing internet voting over a period of time, somewhat like early polling stations provide now. Other electronic voters using personal devices should also be able to vote for a period of time from anywhere with internet access, both to make voting more convenient and to reduce the stress on provided internet connections. Of course, if Elections Canada decides that it cannot
justify the expense of outfitting polling stations with internet access, they could simply specify electronic voting may only be done at home or with one’s own personal device, and all voting at the polls will be done with traditional paper ballots; however, this may run afoul of accessibility requirements.

Possibly the most important aspect of the voting system would be independent oversight. The entire process would have to be supervised by neutral people from varying political parties, academics, members of the judiciary who can interpret whether sufficient accommodations are being made and whether the voting process complies with current election guidelines, and, in this case, individuals familiar and comfortable enough with blockchain technology to monitor the system and troubleshoot technical issues. This oversight board would likely be assembled by a committee dedicated to its creation, in order to ensure a balanced representation of all interests. Further, the board would likely be involved in all aspects of the process to bring blockchain voting to Canadian elections, from alpha and beta testing of the system, to use of the chosen technology in an election, troubleshooting, and reviewing both voting data and results (in an ongoing process throughout the election) as well as the operation (and efficiencies and inefficiencies) of the system itself. The board would also have to analyze issues and present potential solutions for the next election; as such, it may be an idea to have board terms overlap, so that each member sees at least two elections and can compare the results of and improve upon the processes used during each. If the public feels that it can trust the oversight board, ideally they will also develop confidence in the voting system the board represents.

Lastly, and perhaps most obviously, there are considerable technical considerations that will need to be addressed in order to introduce blockchain-based voting as a viable voting option. These considerations may include the development and testing of new or existing wallet apps, testing of various data encryptions methods, determining whether to allow individuals to vote in uncontrolled environments and, if so, how to maximise their voter security and minimise the potential for coercion. Testing will need to be done to determine whether individuals who already have bitcoin accounts can use their current wallet apps, whether it is possible to allow individuals multiple votes despite blockchain transactions being irreversible, and whether current theoretical models of blockchain voting systems could actually be put into
practice. In short, the main implementation obstacles are likely to be public confidence (in the system and the oversight board), technical issues pertaining to the use and development of the voting technology itself, and the purchase, maintenance and storage of equipment.

Conclusion

This paper has described an electronic voting system that operates on the blockchain. Essentially, in order to adopt such a system, voters would be registered to vote and would be assigned a random key pair. It is possible that specialty software would need to be used to ensure that voters are eligible to vote, since otherwise anyone with bitcoins could vote by sending small transactional amounts to the address affiliated with their desired candidates. In the end, votes will be visible and a tally will be computed. While the blockchain system has several advantages in terms of security and transparency, criticisms are warranted given the vital importance of the voting system in allowing Canadians to exercise their Charter-enshrined right to vote. Implementing an electronic voting system operating on blockchain technology will incur upfront and continuing costs, require independent oversight by a multi-disciplinary and multi-party board to ensure the system’s integrity, and would likely require accommodation for voters who do not have access or knowledge of how to use electronic voting technology. Additionally, the system would have to run alongside the paper balloting system, and it is yet undetermined whether the utility and potential convenience of an electronic voting system would outweigh concerns about system security and the accuracy and promptness of results.

Nonetheless, if appropriately applied, blockchain technology would be an excellent method of transmitting vote information, as it is specifically designed to maintain transaction anonymity and fairness. Its two main advantages are its ability to entirely anonymize and protect voter identities, and the fact that it is relatively tamper-proof with strong inherent safeguards in place to prevent hacking and interference by individuals with less-than-honourable intentions. While blockchain technology remains a potential solution to many of the issues that arise in a centralized e-voting system, much consideration is still needed if Canada wants to take advantage of this decentralized transaction system as a possible avenue to electronic voting in the future.

See e.g. Lee, supra note 34.
New Rules for a New World: How Technology and Globalization Shape Bankruptcy Venue Decisions

Laura N. Coordes*

INTRODUCTION

In 2014, Mt. Gox, a bitcoin exchange, filed for bankruptcy proceedings in Japan. The bankruptcy trustee sought to open concurrent proceedings in the United States and Canada in order to protect Mt. Gox's assets from seizure by creditors in these countries. Both the Canadian and the US courts had to determine the location of Mt. Gox’s center of main interests.

---

* Associate Professor, Arizona State University Sandra Day O'Connor College of Law. Many thanks to Karen Bradshaw, Rhett Larson, Erin Scharff, Bijal Shah, and participants at the TAU Workshop for Junior Scholars and the Asper INTLaw Conference for their comments on previous drafts.

1 Bitcoin is a digital payment system that has features similar to a currency. Bitcoins are stored in a user’s “digital wallet,” which functions like a bank account. Bitcoin exchanges, such as Mt Gox, function as a medium between bitcoin traders, allowing parties “to buy or sell bitcoins using different currencies.” Tal Yellin, Dominic Aratari & Jose Pagliery, “What is Bitcoin?” (2017) CNN Money, online: <money.cnn.com/infographic/technology/what-is-bitcoin>.


3 Ibid.
(“COMI”) in order to decide what protections Mt. Gox could be granted in each country.

Although the US played a relatively small role in Mt. Gox’s liquidation, the Mt. Gox proceedings generated much interest in US bankruptcy circles, with many wondering whether the US Bankruptcy Code was equipped to handle this new type of debtor. Could bankruptcy rules and procedures, developed largely before “bitcoin” became a word, accommodate the challenges posed by this new type of entity? For example, do bitcoins qualify as property of the bankruptcy estate (meaning they are assets that could be distributed to creditors in bankruptcy), or are they merely held in trust by the debtor for customers? Assuming bitcoins are property of the estate, where are they located—on a server, with the customer, or with the debtor? Furthermore, where is the debtor itself located? Is it in its jurisdiction of incorporation? Wherever it operates its principal place of business? What if, as was the case with Mt. Gox, all business was conducted online?

The court’s ability to answer these questions is critical for ensuring the smooth functioning of domestic and cross-border bankruptcies in an increasingly globalized world. Without knowing where a debtor’s assets are or what they consist of, we cannot determine what assets are available in bankruptcy for distribution to creditors. Without being able to determine the debtor’s location, we do not know where a bankruptcy case involving that debtor should proceed. The chief difficulty in dealing with a debtor such as Mt. Gox in the United States is that the Bankruptcy Code does not provide clear guidance for answering these questions.

Globalization and technological change have created new possibilities and new challenges in commercial bankruptcy practice, in both the US and abroad. Debtors may now exist entirely online, with no “brick-and-mortar” presence to anchor them to any particular location. Companies today increasingly amass substantial, valuable assets, including vast amounts of consumer information, which they may seek to sell in bankruptcy to other entities. A bankrupt company’s assets may include intangibles that are

---


5 For example, a bitcoin exchange may not even be eligible for bankruptcy in the US depending on its business model. There may be an argument that a bitcoin exchange could constitute a stockbroker, a commodity broker, a bank, or a clearing organization, making it ineligible for Chapter 11 bankruptcy relief. See ibid.

6 See Brian Schaller, “RadioShack Bankruptcy Case Highlights Value of Consumer Data”, (8 June 2015), Information Law Group, online:
difficult to categorize or value, such as social media accounts, domain names, and electronic data. Using the Internet, even purportedly “local” companies can have a global reach, shipping products all over the world.

Technological change also threatens to alter traditional US bankruptcy legal policy. Location has historically played a significant role in bankruptcy, and in practice, a bankruptcy case’s location can affect how the case proceeds, the amount of information relayed to the judge and other parties, and the stakeholders who can participate in the case. Concerns over the determination of a bankruptcy case’s location contribute to perceptions of the legitimacy and fairness of the bankruptcy system.

Scholars are generally of two minds when it comes to case placement in bankruptcy. Although many scholars believe that location and case placement should not be manipulated, others are not convinced and suggest that forum shopping might, in certain cases, produce beneficial outcomes. For both sides, however, technology and globalization might, at first glance, appear to mitigate any lingering concerns about case placement. For example, increased technological capabilities may make a bankruptcy case’s physical location less important if parties who cannot physically appear in court can appear remotely.

This paper counteracts claims that location is no longer a key consideration in bankruptcy law and emphasizes that technological change and globalization instead have the ability to exacerbate harmful forum shopping trends. Bankruptcy law faces substantial challenges in adapting to globalization and technological change, and this paper demonstrates that these changes threaten to make bankruptcy venue even more manipulable than it already is. Thus, the article argues that bankruptcy venue reform must

<http://www.infolawgroup.com/2015/06/articles/privacy-law/radioshack-bankruptcy-case-highlights-value-of-consumer-data/>. The article describes the sale of RadioShack’s assets, which included consumers’ personal information, to General Wireless.


8 See generally Laura N. Coordes, “The Geography of Bankruptcy” (2015) 68 Vand L Rev 381 at 399 [Coordes].

9 Ibid. See specifically the discussion surrounding concerns over forum shopping in bankruptcy.

10 See Samir D Parikh, “Modern Forum Shopping in Bankruptcy” (2013) 46 Conn L Rev 159 [Parikh]. The author proposes ways to alter the resources and incentives of parties who forum shop.

occur in order to ensure the integrity of the bankruptcy system in the face of these changes.

Scholars have debated whether to reform the US bankruptcy venue rules for years; however, the debate received renewed attention in 2012, when the American Bankruptcy Institute ("ABI") created the Commission to Study the Reform of Chapter 11 ("Commission"). The Commission examined the Bankruptcy Code and proposed reforms to Chapter 11 to bring this area of bankruptcy practice up to date with the changes occurring in the business world.\textsuperscript{12} In December 2014, the Commission released an extensive final report proposing numerous recommendations.\textsuperscript{13} Unfortunately, the Commission did not reach a consensus on what changes, if any, were needed for bankruptcy venue reform and did not include any recommendations for venue reform in its report.\textsuperscript{14}

The Commission’s inability to reach a consensus on venue reform is in many ways unsurprising given the strong positions scholars and advocates have staked out on both sides of the debate. Yet, as this paper will demonstrate, it will become increasingly necessary to precisely identify rules that facilitate appropriate access to bankruptcy courts.

By exploring the impact of technology and globalization on bankruptcy venue, this paper raises several broader questions, including how these changes have altered bankruptcy policy, why a bankruptcy case’s location remains relevant in light of technological advances, and how to approach changes to the bankruptcy venue rules. Technological change and globalization have the capacity to fundamentally alter the nature of the businesses that may enter bankruptcy, as well as the number and type of stakeholders that are affected when a business experiences financial distress. While exploring these changes, this article posits that explicit procedures for determining the physical location of a bankruptcy case, currently lacking in the Bankruptcy Code, are more necessary than ever if bankruptcy is to continue to meet the needs of all stakeholders in a case.

Ensuring that US bankruptcy law can accommodate changes due to globalization and technology is important for two additional reasons of particular interest to those outside the United States. First, many countries

\textsuperscript{12} American Bankruptcy Institute, “Purpose of the Commission”, (June 2011) American Bankruptcy Institute (website), online: <http://commission.abi.org/purpose-commission>.

\textsuperscript{13} US, Commission to Study the Reform of Chapter 11, Final Report and Recommendations (Virginia, American Bankruptcy Institute, 2014), online: <https://abiworld.app.box.com/s/vvircv5xxv83aavl4dp4h> [Commission Report].

\textsuperscript{14} Ibid at 310-14.
use US bankruptcy law, and particularly Chapter 11 of the Bankruptcy Code, the chapter used primarily to reorganize business entities, as the model for their own insolvency and restructuring laws. And second, as the Mt. Gox case illustrates, determination of COMI in the cross-border context plays a significant role in the protection debtors can receive around the world. Similarly, understanding where a global debtor is “located” for the purposes of situating a bankruptcy case within the United States has a substantial impact on the progression of the case.

This paper proceeds in four parts. Part I traces the origins of the US bankruptcy venue rules, sketching out how and why they were developed and why Congress designed a national system of bankruptcy courts. Part II discusses the opportunities and challenges that have arisen with respect to bankruptcy venue due to technological change and globalization. Part III explains how process fairness and the interests of justice mandate continued attention to bankruptcy venue reform and analyzes how possible reforms might address the challenges of globalization and technological change. Part IV concludes by emphasizing the need to revisit bankruptcy venue reform in spite of the ABI Commission’s failure to reach an agreement on whether or how reform is needed.

I. Location in Bankruptcy Practice and Policy

This Part describes the rationale behind the traditional importance of location to bankruptcy law and introduces the current bankruptcy venue rules. Although the US bankruptcy system is a federal system of laws, bankruptcy has its roots in the individual states. When bankruptcy law first emerged in the United States in the late 19th and early 20th centuries, the states were the “locus of reorganization law”. Any entity that could go bankrupt, whether it was a municipality, a business, an individual, or a railroad, was necessarily subject to both federal bankruptcy law and individual state laws. Eventually, Congress devised a series of uniform, federal laws to

---

15 To take one example, Canada’s Companies’ Creditors Arrangement Act has many similarities to Chapter 11 of the US Bankruptcy Code, including, for instance, allowing for debtor-in-possession financing. See Companies’ Creditors Arrangement Act, RSC 1985, c C-36.

govern bankruptcies, but the concept of bankruptcy as a predominantly local matter remained prevalent in many respects.\textsuperscript{17}

Indeed, the importance of state and local law within the federal bankruptcy system remains obvious today. For example, courts look to state law to determine the extent of the property rights a debtor has in bankruptcy.\textsuperscript{18} State law also provides certain remedies to parties in bankruptcy, such as a longer lookback period for fraudulent transfers that federal law does not.\textsuperscript{19} Perhaps most importantly, when a company that is a major employer in a city or town goes bankrupt, the entire region can experience negative effects.\textsuperscript{20}

Recently, members of Congress have expressed an ongoing interest in ensuring that the bankruptcy process remains “within the regions and communities that have the most significant vested interest in the outcome” of a case.\textsuperscript{21} Thus, despite bankruptcy’s status as a federal system, Congress has continually recognized the importance of state and local concerns in a bankruptcy case through both public statements and statutory provisions.

Nowhere is the importance of location more prevalent in the bankruptcy laws than in the bankruptcy venue rules. When Congress created the federal bankruptcy laws, it could easily have designated one federal court to specialize in and address all bankruptcy issues, similar to the way other specialty courts are designed. Instead, Congress created bankruptcy courts in each of the 50 states and designed venue rules to give debtors guidance and limitations on where their bankruptcy cases could be filed. Under the bankruptcy venue statute, a debtor may file in one of three locations: (1) in its state of


\textsuperscript{18} \textit{Butner v United States} (1979), No. 77-1410, [1979] 99 S Ct 914 at 77-1410, 440 U.S. 48 \textit{[Butner]}. The judgment holds that a property issue arising in bankruptcy proceedings was resolved by referring to state law rather than a federal rule of equity.

\textsuperscript{19} See \textit{Bankruptcy Code}, 11 USC 2012, s 544, which discusses a bankruptcy trustee’s avoidance powers under state law.

\textsuperscript{20} See Nancy Sarnoff, “Enron’s Collapse May Have Ripple Effect on Downtown Office Market”, \textit{Houston Business Journal} (Dec 9, 2001), online: <http://www.bizjournals.com/houston/stories/2001/12/10/newscolumn3.html?page=all>, which speculates about negative effects on Houston due to the Enron bankruptcy [Sarnoff].

incorporation, (2) in the location of its principal place of business or principal assets, or (3) where a case concerning the debtor’s affiliate is pending.\textsuperscript{22}

In recognizing the importance of location in bankruptcy, Congress reaffirmed a link between venue and geographic location that dates back to English law, to a time when jurors were required to have personal knowledge of a case.\textsuperscript{23} As the court system in England developed, jurors’ personal knowledge took on less significance, but the requirement that a plaintiff “lay” an action in the county where his or her claim arose remained.\textsuperscript{24} Even in “transitory” actions, those actions (relating to debt or contract) that had no particular tie to a given location, the requirement of linking the location of the lawsuit to another relevant location remained.\textsuperscript{25} In these transitory actions, plaintiffs were permitted to situate a case either in the location where the claim arose, or in the location where the defendant or the defendant’s property could be found.\textsuperscript{26} These rules demonstrate that even in early times, courts recognized the importance of bringing a lawsuit in a location that had some meaningful tie to the source of the dispute or the parties affected by it.

When Congress enacted the Federal Judiciary Act of 1789, it sought to incorporate these English principles.\textsuperscript{27} After determining that US public opinion favored even more protections for defendants, however, Congress went one step further, requiring venue to be laid in the location where the defendant in the case resided.\textsuperscript{28} To ensure that no party would have to travel too far to litigate a claim, Congress established a series of local, federal courts across the country. These courts, the predecessors to today’s bankruptcy and federal district courts, were intended to provide “fairness and convenience to defendants.”\textsuperscript{29}

These concepts of fairness and convenience underlie today’s bankruptcy venue rules even though traditional “defendants” do not exist in a bankruptcy case. Thus, in bankruptcy, the venue rules serve to protect the interests of all

\textsuperscript{22} \textit{Judiciary and Judicial Procedure}, 28 USC 2012, s 1408.

\textsuperscript{23} \textit{In re Patriot Coal Corporation, et al., Debtors}, 482 BR 718 at 737 (United States Bankruptcy Court 2012) (No 12-12900 (SCC) \[Patriot Coal\]; see also Willicom Wirt Blume, \textit{Place of Trial of Civil Cases}, 48 Mich L Rev 1 (1949) \[Blume\].

\textsuperscript{24} \textit{Patriot Coal}, ibid at 737; see also 92 C.J.S. \textit{Venue} §§ 3, 4, 7-9, 26, 27.

\textsuperscript{25} \textit{Patriot Coal}, ibid at 737.

\textsuperscript{26} Shirley M Sorter, “Venue Problems in Wisconsin” (1972) 56 Marq L Rev 87 at 89.

\textsuperscript{27} \textit{Patriot Coal}, supra note 23 at 737.


\textsuperscript{29} \textit{Patriot Coal}, supra note 23 at 738. In fact, some citizens threatened to leave the Union if they were not provided with a local court. See Blume, supra note 23 at 36.
parties, debtor and non-debtor alike, by ensuring that a case takes place in a geographically relevant location.

Technological change and globalization have brought forth a slew of large, multinational corporations that often seem to have few concrete ties to any particular location. Nevertheless, corporate bankruptcy cases continue to have substantial impacts on the regions out of which the bankrupt debtor primarily operates. These impacts may in turn affect local economies. A case in point is that of the bankruptcy of Evergreen Solar, a solar panel manufacturer based in Massachusetts. Evergreen Solar filed bankruptcy after the State of Massachusetts had invested large amounts of money, resources, and human capital in the company. When Evergreen Solar filed for bankruptcy, about 800 Massachusetts employees were laid off, the State became a creditor in the proceedings, and the State’s finances were negatively affected. Other cities have also faced negative consequences from company bankruptcy filings. The City of Detroit filed for bankruptcy a few years after the federal government bailed out the struggling automobile companies headquartered there, and the City of Houston’s downtown business district visibly struggled after the very public demise of Enron.

In spite of Congress’s evident attempt to prevent debtors from filing a case anywhere they desire, in recent years, debtors have begun using the Bankruptcy Code’s venue provisions to forum shop, or file a case in a location that is primarily only convenient for the debtor or senior lenders, rather than for all parties involved in the case. Indeed, from 2005-2011, 70% of the largest two hundred public-company filings took place in just two bankruptcy courts, in the Southern District of New York, and in Delaware.

32 Ibid.
33 Ibid.
34 Suzy Khimm, “Why Didn’t the Auto Bailout Save Detroit?”, MSNBC (19 July 2013), online: <http://www.msnbc.com/all-in/why-didnt-the-auto-bailout-save-detroit> (describing both how the auto bailout helped forestall Detroit’s demise and how Detroit was affected by the auto bankruptcies in the region).
35 Sarnoff, supra note 20.
The next Part analyzes how technology and globalization may exacerbate the trend of bankruptcy forum shopping, while also examining some ways in which these advances may alleviate some of forum shopping’s harmful effects.

II. Globalization, Technological Change and Bankruptcy

Technology and globalization have changed many aspects of bankruptcy practice. The following subsections explore new challenges and opportunities that have arisen with respect to various components of bankruptcy law, with particular emphasis on the effects on choice of venue.

i. Debtors

Technological change has broadened the scope of debtor entities eligible to file under the Bankruptcy Code. Entirely new types of companies have emerged in recent years, including businesses based solely online and businesses that specialize in new types of technology. These companies, unlike their more traditional, brick-and-mortar predecessors, may have few tangible assets and may exist in the physical world only as a server in an otherwise empty room. Intangible assets, such as domain names and intellectual property rights, may comprise the bulk of these debtors’ property. These companies lack an easily identifiable physical location for venue purposes as the physical location of the server, divorced from any ties to the people who run the company, does not necessarily serve as a logical location for situating a bankruptcy case.

One example of this new type of debtor is Mt. Gox, a now-defunct exchange for bitcoin. When Mt. Gox filed for bankruptcy protection in Japan, it announced that approximately 850,000 bitcoins belonging to the customers and the company had gone missing, likely stolen by hackers. Although the company later found some of the “stolen” bitcoins in a digital

---


wallet, the bulk of the bitcoins remained missing. The struggles that Mt. Gox faced in tracking and valuing its intangible assets may become typical as companies digitize more assets.

The rise of the start-up culture in the past few decades has coincided with many high-technology companies bursting onto the scene and subsequently struggling or failing completely. Thus, it is becoming increasingly critical for the Bankruptcy Code to accommodate new debtor entity types in emerging fields. Companies need to know whether and how they can file for bankruptcy and what will happen to their nontraditional assets if they do.

One difficulty with extending current bankruptcy rules to new types of companies is that companies that exist entirely online operate differently than do companies with a brick-and-mortar presence. This presents a challenge to bankruptcy judges and creditors, who must examine a company’s operations and decide whether that company’s plan of reorganization treats creditors fairly and is likely to succeed. For example, when Mt. Gox, the bitcoin exchange previously discussed, entered bankruptcy, the court had to determine how to treat this entirely new type of debtor entity: an exchange trading in virtual, digital currency. The Mt. Gox proceedings raised critical questions as to what assets the company actually possessed, where those assets were said to be located, and whether those assets could be considered “property” that Mt. Gox could use to satisfy its creditors.

---

39 Ibid.
40 Sommer Nicole Louie, “Comment: The Inadequacy of Bankruptcy Protection for the Biotechnology Industry” (2005) 22 Emory Bankr Dev J at 337. The author notes that the Code provides inadequate protection to patent licenses as well.
41 Ibid at 338. See the discussion on particular issues in the biotechnology industry.
43 See Rachel Abrams, Matthew Goldstein & Hirotaka Tabuchi, “Erosion of Faith Was Death Knell for Mt. Gox”, DealBook (28 February 2014), online: <http://dealbook.nytimes.com/2014/02/28/mt-gox-files-for-bankruptcy/?_r=0>. The authors describe U.S. prosecutors’ lack of understanding of the way money was digitally transferred via Mt. Gox and noting that Bitcoin investors “are left to figure out what, if any, recourse they have against their losses at Mt. Gox”.
44 For example, digital assets may be said to be located on a physical server, or in the owner’s principal place of business, or elsewhere. The answer may depend on the type of data at issue—it is arguable, for instance, that bitcoins in digital form may not even constitute “data” in the typical sense of the word. See Paul Gil, “What Are Bitcoins? How Do Bitcoins Work?” (26 May
In addition to fostering the creation of new debtor entities, technology has enabled companies to combine and expand at an astonishing rate, resulting in a greater number of companies with a global reach. Faster, better means of communication and connection have made it easier than ever for companies to interact with businesses outside of their home country. Large, multinational companies now dominate the business landscape. Yet, these companies pose a challenge to the Bankruptcy Code, which has typically been focused on assisting companies based in the United States. Although Chapter 15 of the Bankruptcy Code is designed to allow US courts to work in conjunction with foreign courts for certain cross-border insolvency proceedings, observers have still noted that the Bankruptcy Code remains ill-equipped to address the problems faced by these new types of firms. Specifically, these companies, by their nature, have complex corporate structures and look quite different, on paper and in practice, from businesses with only a US presence. Thus, the Bankruptcy Code clearly must continue to adapt if it is to meet the needs of debtors with locations both within and outside of the United States.

**ii. Assets**

The location and composition of a debtor’s property can matter a great deal in bankruptcy. Although the Bankruptcy Code is federal law, bankruptcy courts apply state law with respect to a debtor’s property rights. Additionally, it is important for courts to be able to trace the debtor’s property to determine whether it has been transferred to another party. When a debtor’s assets consisted only of physical property, it was much easier to determine whether property had been transferred. Today, however, when

---

2017), Lifewire, online: <https://www.lifewire.com/what-are-bitcoins-2483146>. The author describes a bitcoin as a “simple data ledger file”.
46 “First Quarter Review”, (29 July 2015), Weil Bankruptcy Blog, online: <http://business-finance-restructuring.weil.com/quarterly-reviews/2015-first-quarter-review/>. The review describes changes that have occurred since the enactment of the 1978 Bankruptcy Code, including the rise of intangible assets, more multinational companies, and greater complexity in corporate structures.
47 *Butner* supra note 18.
48 See 11 U.S.C. § 547 (covering preferential transfers of debtor’s property), § 548 (covering fraudulent transfers).
49 See *Twyne’s Case* (1601) 76 ER 809, which discusses a fraudulent transfer of sheep.
debtors can possess property that lacks a definite geographic base, it can be more difficult to determine who has the property, or even who owns it in the first place.

Even companies that have been around for decades may now possess new types of assets, thanks to technological advances. These assets can include everything from social media accounts to intellectual property rights to domain names, and they may be stored or located all over the world. In addition, companies can now collect and store consumer information, including financial information, on a much broader scale than ever before. When the Bankruptcy Code was created in 1978, these types of assets and enhanced collection abilities were simply unimaginable. Consequently, although the Bankruptcy Code has established some mechanisms for dealing with more traditional intangible assets, such as intellectual property, it is ill-equipped to address the treatment of many of these newer assets, particularly those that lack a recognized registry system or geographic base.

For example, in the case of In re CTLI LLC, the bankruptcy court for the Southern District of Texas had to determine whether social media accounts created by the debtor company's former owner were property of the debtor's bankruptcy estate. No clear treatment for social media accounts currently exists in the Bankruptcy Code. The former owner claimed that the social media accounts belonged to him and alleged that he had privacy interests in the accounts that ought to be protected. Because of this, he argued, it would be impossible for him to share control of the accounts with the company without violating these privacy interests. The court acknowledged that social media accounts created by a person for a company can contain a mix of personal and commercial information but noted that the accounts in this case were created in the business’s name. Therefore, the court determined that the accounts belonged to the debtor company, rather than to the former owner.

---

53 In re CTLI Inc, 528 BR 359 (Bankr SD Tex 2015) No 14-33564.
54 Ibid at 363.
55 Ibid.
56 Ibid at 367-68.
Although the court in CTLI was able to reach a final determination as to the accounts, one can imagine endless variations of the same scenario, each presenting its own unique challenge. For example, what if the company had created an account for the owner’s personal use? Or what if the owner had created the account in his own name but used it to promote the company extensively? These questions remain unresolved for the moment, but will likely arise in the future.

The existence of new asset types demonstrates that technology has outpaced the Bankruptcy Code in many ways. No clear rules exist to guide judges on the treatment of digital goods, electronic data, or domain names. Yet, it is often critical to define, categorize, locate, and value these assets in a bankruptcy case because without an understanding of these assets and their values, a judge has no way to gauge a company’s ability to successfully reorganize or pay creditors.

iii. Case Participation

The previous subsections discussed specific challenges that technological change and globalization have brought to bankruptcy practice. Yet, these changes have presented new opportunities as well. Nowhere are these opportunities more evident than in the realm of case participation. It is now easier than ever for parties to access and participate in bankruptcy proceedings, regardless of where they take place. Advances in communication technology have also made it easier for debtors to disseminate information about the existence and progress of their bankruptcy cases. Parties now can communicate with each other and the court through

57 Ibid at 372.
61 This is consistent with one of the goals of bankruptcy in general and the venue rules in particular: facilitating creditor participation. See Bill Rochelle, “Texas Judge Rails Against Big Ch. 11s in NY, Delaware” (6 August 2015), Bloomberg BNA Bankruptcy Law Resource Center, online: <http://bit.ly/2upuCTT>.
62 See Pamela Foohey, “Notifying Potential Claimants in Diocese Chapter 11 Cases” (30 June 2015), Credit Slips, online: <http://bit.ly/2upFGA1> (describing the creditors’ committee motion in the Archdiocese of St. Paul and Minneapolis bankruptcy, which sought to have the
the existence of dedicated web pages and email addresses, and parties who are not physically close to where a case is taking place can participate directly in court proceedings using telephone or video chat. Thus, as debtors and assets become more dispersed, technology provides mechanisms to bring parties together in the courtroom.

Unfortunately, technological advances are not without flaws. For example, in recent years, scholars have begun to study court participation via videoconference, with some concluding that even videoconferencing does not equate to the benefits of being physically present.\(^{63}\)

Technology also has the potential to make the bankruptcy filing process easier for certain low-income and individual debtors; however, a recent court decision illustrates the need to proceed cautiously in this area. In \textit{In re Reynoso},\(^{64}\) parties used web-based software to prepare and file a bankruptcy petition, a task usually given to a lawyer. When the petition was challenged in bankruptcy court, the Ninth Circuit Court of Appeals found that the software was performing the functions of a bankruptcy petition preparer\(^ {65}\) and held that the software provider had thereby engaged in the unauthorized practice of law.\(^ {66}\)

Other concerns have arisen in connection with electronic filing (“e-filing”) of cases. Although many praise e-filing software for cutting costs in the long run, the initial learning curve for such software can be steep, even for experienced attorneys, and e-filing programs themselves are expensive to purchase, necessitating high upfront costs.\(^ {67}\) Many small and individual debtors will therefore not be able to meet the initial demands of time, effort, and money to master e-filing. This could leave these debtors at a disadvantage or even discourage them from filing completely.

bankruptcy court order all parishes within the diocese to play a video in which abuse claimants discuss the necessity of filing a claim by the bar date).

\(^{63}\) See Erich P Schellhammer, “A Technology Opportunity for Court Modernization: Remote Appearances” (January 2013), CCTJ, online: <http://bit.ly/2vOS4ZX> [Schellhammer]. The authors highlight some concerns with using technology in the courtroom, including the concern that participation via teleconference does not allow for the observation of a speaker’s non-verbal cues.


\(^{65}\) \textit{Ibid} at 1123-24.

\(^{66}\) \textit{Ibid} at 1125.

Thus, although technology can be a valuable tool for creating access to bankruptcy courts, it also creates division, as not everyone has equal access to technology. Therefore, as technological advances become integrated into bankruptcy processes, a risk arises that those without access to technology may similarly lack access to these bankruptcy procedures. In combination with the new types of debtors and assets described above, these technological changes may result in the bankruptcy system becoming more opaque and inaccessible to increasing numbers of parties.

iv. Changes to Bankruptcy Law

In response to technological changes, bankruptcy scholars, judges, and practitioners have begun to work on reforms to the Bankruptcy Code. These parties, firsthand observers of the bankruptcy system at work, have not been ignorant of the need to change the Bankruptcy Code and Rules in order to accommodate the new business models and challenges that have emerged. As previously discussed, the ABI created the Commission to Study the Reform of Chapter 11 and issued a series of proposals in December of 2014. One of the biggest changes the Commission recommended was essentially splitting the Chapter 11 system into two tracks in order to better accommodate the needs of both large, multinational firms and small- to mid-sized businesses. Chapter 11 currently serves as the primary reorganization mechanism for all of these business types, but the Commission recognized that the needs of different entities have changed so much that vast reform was needed. In light of the Commission’s sweeping proposed changes to much of Chapter 11, it is disappointing that some venue reform mechanism was not concurrently proposed. Just as technology necessitates changes to many aspects of bankruptcy law, it also affects bankruptcy procedures, including those related to venue.

Indeed, venue reform is desperately needed for Chapter 11 cases. Although scholars have debated about whether bankruptcy forum shopping is a problem for years, technological advances and globalization make it even easier for parties to manipulate the venue rules. For example, large,

---

68 For example, the needs of a large company like Enron, a business whose collapse affected parties around the world, may be quite different than those of a smaller company, which has only one location and a discrete set of assets. For further discussion on the challenges of the Enron case, see *Re Enron Corp* (2002), No 01-16034 (AJG), [2002] 274 BR 327 (Bankr SDNY).

multinational companies have a seemingly endless array of venue options available to them, while companies that exist almost entirely in the digital realm may be forced to rely solely on their state of incorporation when choosing a bankruptcy venue. Recent debates in the scholarly literature demonstrate concerns over forum shopping in the transnational context and the extent to which US bankruptcy law ought to discourage such practices. Additionally, technology makes it easier to dismiss objections to venue from smaller stakeholders. These parties often argue that bankruptcies should take place closer to the main locus of the company or closer to key employee groups. Yet, many debtors are easily able to dismiss these arguments by countering that technology allows these parties adequate remote access to the courtroom.

This manipulation need not come solely from the debtor. Lenders, particularly those who hold senior positions as well as those with security interests in most of the debtor’s assets, can also influence where the debtor files for bankruptcy. Lenders can take a company’s venue options for bankruptcy into account at the time of financing, and lenders do exercise influence over the debtor’s venue decision at the time of filing. Thus, the manipulation enabled by technology and globalization can work to the senior

case of forum shopping; Adrian Walters, “United States’ Bankruptcy Jurisdiction Over Foreign Entities: Exorbitant or Congruent?” (2017) J Corp Studies [Walters]. The author states that the “experience in Europe suggests that debtors can easily manipulate COMI for bankruptcy forum shopping purposes.”

70 Lynn M LoPucki, Courting Failure: How Competition for Big Cases is Corrupting the Bankruptcy Courts (Ann Arbor: University of Michigan Press, 2005); Pottow, ibid at 786. Here, Pottow challenges LoPucki’s contention that universalism will lead to “rampant” forum shopping; see also Honorable Samuel L Bufford, “Global Venue Controls are Coming: A Reply to Professor LoPucki” (2005) 79 Am Bankr LJ 105 [Bufford]. Bufford disagrees with LoPucki and supports modified universalism for international insolvency proceedings; Walters, supra note 69. Walters states that “by fostering swift and unconstrained access and eschewing barriers to entry that are common in many other countries...the US system is highly conducive to bankruptcy forum shopping.”

71 These casual dismissals come at a distinct cost to small stakeholders. See Patrick Cormier, “The Opportunities and Challenges of Court Remote Appearances”, Slaw (2013), online: <http://www.slaw.ca/2013/11/06/the-opportunities-and-challenges-of-court-remote-appearances/>. Cormier describes the costs and challenges of remote appearances; Schellhammer, supra note 63, who observes that technology does not definitively resolve problems of court access.

72 Susan Mathews, “Corporate Chapter 11 Bankruptcies: The Case for Venue Reform” (October 2014), ABF Journal, online: <http://www.abfjournal.com/articles/corporate-chapter-11-bankruptcies-the-case-for-venue-reform/>. Mathews notes that “the decision to file in [popular] venues seems to be driven by lenders that find these venues more favorable”.
lenders’ benefit as well, possibly to the detriment of more junior lenders or smaller stakeholders whose interests may be opposed to those of the senior lenders.

These arguments raise two critical questions, to be explored further in the next Part. First, does technology truly enable all parties to participate in a case regardless of where it is filed? And second, why does it matter where a bankruptcy case is filed at all?

III. The Interests of Justice and Process Fairness

The previous Parts illustrate that the current US bankruptcy venue rules, and perhaps much of US bankruptcy law in general, are ill-equipped to keep pace with a changing world. This Part discusses some of the key questions raised by the changes previously discussed, argues that venue remains a relevant concern in light of the interests of justice and process fairness, and proposes a conception of bankruptcy venue that is flexible and responsive to change.

i. Fundamental Questions

Changes wrought by technology and globalization have brought to the forefront a question that cuts to the heart of bankruptcy’s purpose. As previously discussed, new entities, new assets, and new opportunities for case participation pose both opportunities and challenges to current bankruptcy laws. At the core of these challenges lie questions about the entities the bankruptcy system is designed to serve. If bankruptcy is to continue to meet the needs of companies large and small, virtual and physical, and all of their numerous creditors, employees, and stakeholders, how are the bankruptcy laws to achieve this goal?

The ABI Commission was correct to recognize that bankruptcy laws need to change if the bankruptcy system is to meet the challenges of the digital age. This article posits that reform to the venue rules is a critical part of this change, just as critical as changes to the rest of the Chapter 11 system. New bankruptcy proposals have the potential to embrace technology’s advantages while minimizing its limitations and these ideas must be extensively deliberated and debated.\textsuperscript{73} Bankruptcy law must rise to meet the challenge of

\textsuperscript{73} Such laws could include everything from provision of a national bankruptcy court to the establishment of local liaisons for large bankruptcy cases.
accommodating new entity and asset types while continuing to facilitate stakeholder participation in cases where the location of a case can make a difference.

Although bankruptcy was initially designed to be a collective process, one where all parties could participate to facilitate a resolution, the debtor today (or a senior lender) may easily select a case venue that serves its own interests while casually dismissing the interests of the rest of the group. Thanks in part to technological advances, debtors and senior lenders are able to dismiss objections from parties who cannot easily participate in a case by pointing to technology that allows for remote access and participation. Manipulation of bankruptcy venue rules contributes to a perception that in many bankruptcy cases, the outcomes are predetermined, and parties are helpless to reverse the tide of decisions made by the very largest players.74

ii. Location in Bankruptcy

As described in Part I, location has always played an important role in bankruptcy policy and practice. Recently, technological innovation and increased globalization have forced the question of whether this role remains relevant.

Local court rules continue to ensure that location still matters, at least to some degree. Certain jurisdictions require local counsel to participate in all bankruptcy cases. Included among these jurisdictions is Delaware, the location of many large corporate bankruptcy cases.75 Even in jurisdictions that do not require local counsel to be present, many parties prefer that their lawyers be physically present in the courtroom, particularly if they are arguing a motion that may be difficult or challenging.76 Other parties may simply

---

74 Coordes, supra note 8 at 387. See specifically how “large bankruptcies now cater almost exclusively to the wishes of power players.”

75 See Francis Pileggi, “Delaware Practice and Procedure for Non-Delaware Lawyers”, (31 July 2012), Delaware Corporate & Commercial Litigation Blog, online: <http://www.delawarelitigation.com/2012/07/articles/commentary/delaware-practice-and-procedure-for-non-delaware-lawyers/>. Pileggi notes that Delaware courts emphasize “that a Delaware attorney of record is responsible for every action taken by his or her client... [T]he Delaware lawyer who appears in an action always remains responsible to the Court for the case and its presentation”.

76 In re Dunmore Homes, Inc (2008), No 07-13533 (MG), [2008] 380 BR 663 (Bankr SDNY). In contrast, some courts, such as the Bankruptcy Court for the Southern District of New York, have permissive pro hac vice provisions that allow “any lawyer admitted to practice in any state or federal court to be admitted pro hac vice without requiring local counsel.” Ibid at 675.
value their lawyers’ presence in court for the benefits of firsthand observation of the court, witnesses, and other parties in interest.

Although many scholars and commentators have criticized the continuing need for local counsel in a globalized, tech-friendly world, local counsel can provide distinct benefits, particularly in complex or international proceedings.\textsuperscript{77} In bankruptcy court, local rules and unwritten customs abound, creating the potential for embarrassment for parties unfamiliar with these norms.\textsuperscript{78} But perhaps the key benefit of having a local lawyer in a case is having someone familiar with the people involved.\textsuperscript{79} Ultimately, bankruptcy, and indeed the practice of all law, is about people.\textsuperscript{80} Having a person familiar with the people affected by a case can provide significant benefits. These benefits are compounded when a case crosses borders, exposing litigants to increased cultural differences, language barriers, and unfamiliar processes.\textsuperscript{81}

As telephonic and video access to court become increasingly available to parties and their lawyers, questions have arisen as to whether it is necessary for parties to be physically present in the courtroom. Yet, telephonic and video representation carries its own set of expenses, which lawyers pass on to their clients.\textsuperscript{82} Thus, representation in this manner may become more difficult to attain for smaller parties who lack the means to pay for the technology to participate in the case. Furthermore, even with all of the advances of technology, the quality of participation may not be the same as it would be if the parties were physically present in the courtroom. Thus, although technology may ameliorate some aspects of forum shopping, it does not entirely resolve the problems associated with situating a bankruptcy case far from many of the key stakeholders.

Allowing companies to file in the jurisdiction where they are incorporated may initially seem to resolve many of the issues associated with these “global” and “digital” companies. Yet, the so-called “state of incorporation” prong of the bankruptcy venue statute has been maligned by

\textsuperscript{77} Timothy Miltenberger, “The Indispensable Local Counsel”, TYL Magazine (2017) 8.
\textsuperscript{78} Ibid.
\textsuperscript{79} Ibid at 9.
\textsuperscript{80} Ibid.
\textsuperscript{81} Ibid.
\textsuperscript{82} For example, Court Call, a service that provides telephonic and video remote appearances, charges a fee for each party appearing by telephone. See Court Call, “Court Call Remote Court Appearances”, online: <https://courtcall.com/what-is-courtcall/>.
scholars and policymakers as being the source of harmful forum shopping. Indeed, it is this prong of the venue statute that has faced the most criticism, and many proposals for venue reform have called for its removal. If scholars are correct that the state of incorporation prong promotes harmful forum shopping and should be eliminated from the bankruptcy venue statute, the other prongs of the statute provide little help to these global and digital companies. For example, trying to locate a “principal” place of business or a single location for the debtor’s assets may be difficult for companies with assets located all over the world or for companies that exist almost entirely in the digital realm.

Although technology has arguably reduced the number of locations that can be considered unfair or inconvenient to non-debtor parties, opportunities for harmful forum shopping, or placing a case essentially out of the reach of many smaller stakeholders, still exist. Under the current venue rules, large debtors can now file in virtually any judicial district in the country, even if they lack a physical presence or meaningful connection to that location. Venue law arose in part out of the need to protect parties from unfair or inconvenient locations, and until this need is eliminated entirely, it remains critical for scholars and policymakers to focus on venue and case placement. This need is reflected in the two considerations a court undertakes when deciding whether to change the venue of a bankruptcy case: the court must consider the interest of justice and the convenience of the parties. Furthermore, the US Supreme Court recently affirmed the importance of situating a case in a meaningful location when it ruled that out-of-state plaintiffs could not sue in Montana based solely on the ground that the defendant railroad company did business in the state. Although this was not a bankruptcy case, the Supreme Court’s decision that parties who do not control where a case is placed must be protected has important implications for bankruptcy cases as well.

83 HR 2533, supra at 31. This proposes to eliminate state of incorporation as a venue option; Samir D Parikh, “Modern Forum Shopping in Bankruptcy” (2013) 46 Conn L Rev 159 at 200. The author advocates restricting venue in corporate cases to the location of the debtor’s principal place of business or principal assets or, in some cases, to the location of the debtor’s affiliate.
84 Patriot Coal, supra note 23 at 738. Specifically see the statement that “local federal courts were established to ensure fairness and convenience to defendants.”
85 Judiciary and Judicial Procedure, 28 USC 1984, § 1412.
86 BNSF Railway Company v Tyrrell (2017), No 16-405, at 4-9 (S Ct 2017).
Judicial legitimacy concerns also demand careful consideration of location in the bankruptcy context.\textsuperscript{87} Professor Lynn LoPucki has argued that judges compete for cases in their home jurisdiction, calling into question the neutrality of bankruptcy judges when it comes to venue procedures.\textsuperscript{88} If the current venue rules do not give clear guidance on where to situate a case, they arguably make it even easier for judges to use what influence they have to obtain the case. Thus, to the extent that the venue rules can be adjusted to discourage this type of judicial behavior, they should be reformed.

Of course, not all scholars agree with Professor LoPucki’s characterization of judicial behavior. In fact, some evidence indicates that judges can, and do regulate forum shopping themselves.\textsuperscript{89} If this is the case, then perhaps judicial mechanisms that constrain forum shopping should be strengthened.\textsuperscript{90} More research is needed to determine whether judges are effectively regulating forum shopping; if they are, allowing for gradual change through the process of judicial review may be an option to explore.

Globalization and technological change have arguably obscured the lines of what is in fact “local” for a particular debtor. Still, there remains significant interest in having bankruptcies adjudicated “at home”—near the company’s principal stakeholders and creditors, even though it is now significantly harder to determine the location of the debtor’s principal place of business or principal assets.\textsuperscript{91} As a result, the “fallback” rule becomes siting a case in the debtor’s state of incorporation. Given that the state of incorporation venue prong is already the most ubiquitous prong used for the bankruptcies of large companies, it is critical to ensure that technological change does not further exacerbate this trend. Scholars have acknowledged the disadvantages that can result from a large debtor choosing to file in its state of incorporation, particularly if that debtor has no other meaningful connection to the state.\textsuperscript{92}

\textsuperscript{87} See Coordes, \textit{supra} note 8 at 395-96 (discussing how judges may be reluctant to transfer cases).
\textsuperscript{88} LoPucki, \textit{supra} note 70.
\textsuperscript{89} Walters, \textit{supra} note 69; David L. Lawton, “If It Ain’t Broke, Don’t Fix It: Justifying Forum Shopping: An Argument in Favor of Expansive Chapter 11 Jurisdiction, \textit{INSOL International News Update} (April 2017) (noting that courts often dismiss non-US cases, at least partially because they have determined that the US was not the most appropriate venue in light of local interests).
\textsuperscript{90} Walters, \textit{supra} note 69.
\textsuperscript{91} Coordes, \textit{supra} note 8 at 390-91 (describing a continued desire among members of Congress to place bankruptcy cases near the debtor’s operational center).
\textsuperscript{92} See Parikh, \textit{supra} note 10 at 200.
Finally, concerns over process fairness suggest that companies should pay more attention to the needs of smaller stakeholders when undertaking a major restructuring.\(^\text{93}\) Process fairness is the idea that stakeholder input matters to the ultimate success or failure of any venture. Whether stakeholders consider a process to be fair or not can directly affect the amount and quality of their support in a case.\(^\text{94}\) Case studies have shown that if management is transparent about its decision-making process, this transparency can make a qualitative difference in terms of both employee morale and ultimate company success.\(^\text{95}\)

This means that companies must be cognizant of the views of their stakeholders and work to include them whenever the company undergoes a massive change such as bankruptcy reorganization. When employees feel as though they understand what is going on in a case and the reasons behind the decisions that are being made, they are more likely to support those decisions.\(^\text{96}\) Situating a bankruptcy case near a company’s employees and other key stakeholders, such as suppliers, can help these parties to feel more involved in the process and may in turn make these stakeholders more likely to support, rather than oppose, a company’s bankruptcy plan and exit strategy.

Process fairness studies also show that when employees feel betrayed by their companies, they tend to retaliate. These retaliations can, in turn, translate to serious consequences for the company’s bottom line.\(^\text{97}\) In contrast, when employees feel that they have a voice in the decision-making process, they are more likely to actively support the company’s decisions, their bosses, and the direction of the organization as a whole.\(^\text{98}\) Thus, the decision over where to situate a bankruptcy case remains significant, as it will undoubtedly be easier for a company to inform and include its stakeholders if the case is taking place in a location where these stakeholders can observe the proceedings and have a chance to become involved in them.


\(^{94}\) Stephanos Bibas, “Transparency and Participation in Criminal Procedure” (2006) 81 NYU L Rev 911 at 949. Bibas states that “people respect the law more when it is visibly fair and when they have some voice or control over its procedures.”

\(^{95}\) Brockner, *supra* note 93.

\(^{96}\) Ibid.

\(^{97}\) Ibid.

\(^{98}\) Ibid.
If venue rules become so manipulable, or so meaningless as to become obsolete, this will have a negative impact on the interest of justice, process fairness, and the overall perception of the bankruptcy system. Although the ABI Commission was unable to reach an agreement on how to reform the bankruptcy venue statute,\textsuperscript{99} others should pick up where the Commission has left off and continue the conversation about bankruptcy venue reform before it becomes even harder to implement viable reforms.

\textit{iii. Ideas for Bankruptcy Venue Reform}

Advances in technology threaten a location’s relevance to bankruptcy law. Yet, as this article has shown, a case’s location can still affect the outcome and concerns about process fairness and judicial integrity still dictate a need for consideration of location.\textsuperscript{100} The preceding subsection has illustrated the importance of continuing to discuss venue reform in bankruptcy. Although technology has arguably made it somewhat easier for parties to remotely participate in bankruptcy cases, it has also made it more difficult for new types of companies to discern where to situate their cases. As previously discussed, a popular proposal would eliminate the state of incorporation venue prong and require all bankruptcy cases to be filed in a corporation’s principal place of business. Technological advances and globalization, however, pose challenges to this proposal. Companies with assets all over the country or globe may have multiple locations that could arguably constitute their “principal place of business.” In contrast, companies that exist entirely online may not be able to identify with certainty any physical location that could serve as their principal place of business. This uncertainty creates confusion in the law and may encourage companies to try to devise a way to situate a case in a location that is convenient only to them or their most important lenders.

One potential solution for bankruptcy venue reform is to extend the Commission’s recommendation about splitting Chapter 11 into two tracks to the venue rules as well. What this would mean is that there would essentially be two sets of venue rules, one for large, multinational firms (and perhaps digital companies as well), and one for small- and medium-sized businesses.\textsuperscript{101} The venue rules could remain largely unchanged for the latter category, but

\textsuperscript{99} Commission Report, \textit{supra} note 13 at 310-14.
\textsuperscript{100} Brockner, \textit{supra} note 93.
\textsuperscript{101} Commission Report, \textit{supra} note 13 at 310-14.
change dramatically for the former. For example, to address the unique challenges of companies in the former category, some scholars have proposed creating a national bankruptcy court. This court would have ample resources to accommodate claims and interests arising from a variety of locations and would use the best technology. To the extent practical, the court could be staffed by judges who are willing to travel to different parts of the country for hearings.

Of course, the development of a national court does not mean that local courts have no role to play in the new bankruptcy system. These courts would be particularly valuable to smaller debtors and would remain the primary venues for small- and medium-sized bankruptcy cases.

Even if a national bankruptcy court is adopted, provision must still be made to guarantee process fairness and to ensure the integrity of the bankruptcy system. Developing procedural safeguards will naturally invite inquiries into the location of the relevant parties in a bankruptcy case meaning that, although perceptions may have changed about what is “local,” a bankruptcy case’s location will still be given serious consideration.

With advances in technology and globalization, a bankruptcy case’s location at first glance seems to matter less—parties who want to participate in a case can do so via Skype, telephone, or the internet. But digging deeper reveals that many of the problems associated with case participation are not eliminated by technology, and technology in fact can sometimes make the question of where to situate a bankruptcy case much harder. The idea that bankruptcy is a system designed to protect the rights of all parties remains a core principle. This means that as the Bankruptcy Code undergoes revisions, it will be critical to focus on the quality of access to bankruptcy court through the development and revision of venue rules. Even if the nature of the bankruptcy system changes, bankruptcy cases must continue to take place in venues that can meet the needs of all affected parties.

Just because the ABI Commission failed to reach a consensus on how to revise the bankruptcy venue rules does not mean that others cannot pick up

---


103 For examples of procedural safeguards that might be considered, see Coordes, supra note 8 (arguing for a hearing on venue at which the United States Trustee is heard); Bufford, supra note 70 at 107 (arguing for the separation of the decision to open a case from the decision of where a case takes place and for notice to all parties in interest with respect to the latter decision).
where the Commission has left off. The venue rules and procedures clearly need an update. Current proposals calling for a debtor’s choice of venue to be cabined to the location of the debtor’s principal place of business may make little sense in light of how technology and globalization have enabled the creation of debtor companies with no clear or primary physical location. Fresh ideas are needed to ensure the consistent treatment of debtors and those who interact with them while in bankruptcy.

Scholars have proposed many ideas for reforming the bankruptcy venue rules and procedures over the years. These proposals should not fall by the wayside simply because the ABI Commission was unable to determine the way forward. Instead, it is more critical than ever to discuss and debate ideas for venue reform before new advances make the current venue rules even more manipulable and obsolete. Although forward-looking legal reforms may be difficult to implement, encouraging proposals to be brought forward will make it more likely that viable changes can be put into place.

Even as technology advances and the world grows more interconnected, physical location will remain relevant in bankruptcy law. A case’s physical location still affects who may come to court and participate in the case, the number of witnesses or experts that may gather or interact, and the type and quality of interactions that parties can have with each other. A bankruptcy case can also provide a physical meeting space for companies with an otherwise virtual existence. If bankruptcy venue reform is truly stymied, at a minimum we must find ways to keep local parties involved.

IV. Conclusion: Ideas and Further Discussion

As discussed, there are many possible avenues for pursuing bankruptcy venue reform. Congress could make changes directly to the statute to reduce the number of venue options or to make the existing options less manipulable. Reforms might also come through creating procedures for more parties to be heard and involved in a venue decision. Reform could even come from the bankruptcy courts themselves, in the form of new local rules to govern case practices, the establishment of a national court for large cases,

---

104 Commission Report, supra note 13 at 310-14.
105 In fact, in a previous article, I have advocated for venue reform in the form of new venue procedures, with the United States Trustee taking a more active role in venue discussions and proceedings, see Coordes, supra note 8. Other proposals have been mentioned throughout this article.
or judges simply cracking down on filings that seem particularly concerning for smaller stakeholders.

Regardless of the particular avenue that is ultimately chosen, technology and globalization are making the status quo particularly unacceptable. Thus, the ABI Commission’s lack of consensus on venue reform should not end the debate. In an effort to move the debate forward, this article has shown that problems with venue and case placement will only become more complicated as companies expand globally and virtually.

At the heart of the question about venue reform is a question about the integrity of the bankruptcy system. If bankruptcy is to remain an accessible system to all parties, changes to bankruptcy venue are as necessary as any of the other changes recommended by the ABI Commission.

Technological change and globalization pose great challenges and present great opportunities for bankruptcy practitioners and theorists. As we focus on ways to adapt bankruptcy rules and procedures to the needs of a changing world, it is critical to recognize the role technology and globalization can play in keeping the bankruptcy system accessible to many, as well as the harm they can pose in granting that accessibility only at a high cost. Bankruptcy venue rules and procedures must adapt to address the threat posed to the bankruptcy system’s perceived integrity by these external forces.
Rationalizing the Defences to Enforcement under the New York Convention 1958

Dr. Chrispas Nyombi & Dr. Konstantinos Siliafis

SYNOPSIS

Since the mid-twentieth century, the New York Convention has been a central pillar of international arbitration by facilitating the global recognition and enforcement of international arbitral awards. However, despite half a century of development, there remains a significant gap in local courts’ interpretations of the exceptions to recognition and enforcement found under Article V of the Convention. This paper advocates for adopting a narrower and more unified approach to these exceptions, the rationale being that the objective of the exceptions is to protect the attractiveness of the institution of international arbitration as an alternative dispute resolution mechanism independent from national courts. The paper does not delve into the debate on delocalisation; rather, it critically evaluates the grounds available under the Convention for challenging recognition and enforcement of international awards and the issues they create for prospective participants.

Dr. Chrispas Nyombi and Dr. Konstantinos Siliafis are Senior Lecturers in Law at Canterbury Christ Church University.
INTRODUCTION

The law governing the enforcement and recognition of international arbitration was originally vested in the Geneva Protocol of 1923 and the Geneva Convention of 1927. In 1958, the Convention on the Recognition and Enforcement of Foreign Arbitral Awards (“New York Convention”) was introduced by twenty-four signatories, superseding previous international instruments and ushering in a new era of transnational commercial arbitration. The New York Convention entered into force in order to encourage the resolution of international disputes through arbitration. It applies to international awards and expressly covers awards made in a State other than the one where enforcement is sought. Thus, recognition and enforcement of international arbitral awards among member States is a central pillar of the New York Convention. As a result, the New York Convention has been described as “the single most important pillar on which the edifice of international arbitration rests”, which “perhaps could lay claim to be the most effective instance of international legislation in the entire history of commercial law.” Furthermore, the New York Convention is highly influential internationally, as ninety-eight per cent of international arbitral decisions made under the New York Convention are recognised and ninety-five per cent of arbitral awards are enforced by local courts. The high success rate is buttressed by the fact that

5 New York Convention, supra note 3, art I(1).
6 Ibid, art III.
defences to enforcement under the New York Convention are construed narrowly “to encourage the recognition and enforcement of commercial arbitration agreements in international contracts.”

The aim of this paper is threefold. First, it will explore the essential characteristics of international commercial arbitration with particular emphasis on the conditionality for recognition and enforcement of arbitral awards under the New York Convention. Secondly, it will critically examine defences to enforcement found under Article V of the New York Convention. Particular emphasis will be given to the approaches taken by national courts with regards to the Article V grounds. Last but not least, a circumspect conclusion will be reached on the direction of future development in this area.

Requirements for enforcement under the New York Convention

Enforcement of international arbitral awards is governed by Article III of the New York Convention, which indicates that contracting states must recognize arbitral awards made under the New York Convention as binding. Without this recognition, arbitral awards would become unpredictable and could have the effect of deterring persons from using arbitration as an alternative dispute resolution mechanism. Furthermore, without the New York Convention’s enforcement procedure, arbitration would not have evolved as rapidly as it has in the last thirty years. This argument is supported by an empirical study conducted by Loukas Mistelis, who found that although difficulties arise when the place where the enforcement is sought is different from the place of arbitration (seat); the most cited reason for using arbitration nonetheless is because of its excellent enforcement mechanism. None of the applicants in the study cited the New York Convention’s enforcement procedure as a major concern.

However, enforcement of arbitral awards poses a number of challenges. These challenges arise in the context of domestic laws influencing the provisions of the New York Convention at issue. Article I(1) of the New York Convention permits the enforcement of awards considered nondomestic by the

---

12 Ibid at 545.
enforcing State; however, in the absence of a unilateral definition as to what constitutes “nondomestic”, each contracting State must determine its own definition, which creates inconsistencies. These inconsistencies are clear when comparing the United States [US] and China, for example. As such, a dispute between US citizens with respect to property located abroad or involving the legal system of a foreign state is nondomestic pursuant to the US Federal Arbitration Act of 1925. In contrast, the Chinese law views as nondomestic only an award made in a country that is a contracting state of the New York Convention and with which China has a bilateral treaty on provision of judicial assistance. This limitation is as a result of the reservations that China made on its ratification of the New York Convention. China’s reservations were largely dictated by their general distrust over international arbitration, which is evident from the fact that Chinese parties are reluctant to elect arbitral seats outside China.

Enforcement under the New York Convention could also be compromised by reciprocity and commercial reservations provided for under Article I(3). Although these reservations limit the New York Convention, they nevertheless cannot be construed to act as a barrier to enforcement of arbitral awards. This is because half of the contracting states have selected the reciprocity reservation to deny enforcement of arbitral awards of non-contracting states. Secondly, the New York Convention allows contracting states to decide on whether a contract is commercial or not, which widens the scope of the Convention. Without a set criterion, commercial reservations will remain a contentious issue.

Furthermore, the New York Convention allows a party to vacate or annul an award only in the state where the award was rendered. Consequently, these states have “primary jurisdiction” over the arbitral award. The

---

13 New York Convention, supra note 3, art I(1).
16 Michael Moser, Dispute Resolution in China (Juris Publishing, 2012) at 337.
17 New York Convention, supra note 3, arts V(I)(e), VI.
18 The concept that the States in which the award was rendered have primary jurisdiction over the arbitral award was reinforced in M&C Corp, infra note 19.
decision *M&C Corp v Erwin Behr GmbH & Co* (1996) reinforced the rule that a motion to vacate may be heard “only in the courts of the country where the arbitration occurred or in the courts of any country whose procedural law was specifically invoked in the contract calling for arbitration of contractual disputes.”\(^{19}\) In addition, every state has its own statute of limitations. For instance, the UK allows a limitation period of twenty-eight days to appeal the decision of an arbitral body, \(^{20}\) whereas one month is permitted in France.\(^{21}\) There are also a number of other requirements that must be taken into account before an award can be enforced, such as the dispute being of a legal nature.\(^{22}\)

As aforementioned, the *New York Convention* limits the grounds that a party can plead by providing an exhaustive list of defences. Courts are very cautious about overturning international arbitration awards on appeal or setting aside decisions, except in the clearest of cases.\(^{23}\) This is rooted in the objectives of the *New York Convention* and the courts’ respect for parties’ contractual autonomy; thus, courts exercise their discretion when faced with arbitral appeal.\(^{24}\) The grounds for challenging an arbitral award mainly fall into two broad categories: jurisdiction and procedure. An appeal on jurisdictional grounds raises questions of whether the arbitral body has the authority to hear a particular matter. Typically, a unilateral decision to use a different arbitral institution will be challenged on jurisdiction where a party

\(^{19}\) *M&C Corp v Erwin Behr GmbH & Co KG*, 87 F (3d) 844 at 847-49 (6th Cir 1996) [*M&C Corp*].

\(^{20}\) *Arbitration Act 1996* (UK), c 23, s 70(3).

\(^{21}\) Art 1519 NC proc civ. In the UK, an action to enforce an award must be brought within six years from the date the cause of action occurred: see *Limitation Act 1980* (UK), c 58, s 7. In India the limitation period is three years: see *The Limitation Act*, No. 36, Acts of Parliament, 1963, s 101.


opts to go elsewhere.\textsuperscript{25} The body named in the arbitration agreement entered into by the parties remains the only authority over the dispute, assuming the dispute comes within the scope of the arbitration clause/agreement. If one party decides to go to another forum without the consent of its counterpart to the arbitration clause, this will give rise to the counterpart’s challenge of this other forum’s jurisdiction based on the parties’ agreement regarding forum in the arbitration clause.\textsuperscript{26}

The right to appeal is also provided for where there is evidence of procedural impropriety. This right is reflected in the seven enumerated defences to enforcement provided under Article V of the New York Convention.\textsuperscript{27} The burden of proving that the ground for appeal falls within the seven defences listed under Article V is always on the party who resists the recognition and enforcement of the award. The party opposing the request for recognition and enforcement has additional grounds to challenge the award under Article V(2).\textsuperscript{28} These grounds are public policy and the ‘non-arbitrability’ of the dispute, which means that local courts will refuse recognition and enforcement if such recognition and enforcement contravenes the public policy of the state where the recognition and enforcement is sought, or if the subject-matter of the dispute could not have been referred to arbitration pursuant to the laws of the enforcing state. These challenges to the recognition and enforcement of international arbitration awards are viewed on a case by case basis.\textsuperscript{29}

\textbf{ARTICLE V DEFENCES TO ENFORCEMENT UNDER THE NEW YORK CONVENTION}

\textbf{Incapacity / Formation of Arbitration Agreement}

First and foremost, if the parties were under incapacity at the time the arbitration agreement was made, or the arbitration agreement is not valid under the laws of the country to which the parties have subjected it (or failing such an indication, under the laws of the country where the award was made),

\textsuperscript{25} Schwebel, \textit{supra} note 22 at 83-87; Pieter Sanders, “A Twenty Years’ Review of the Convention on the Recognition and Enforcement of Foreign Arbitral Awards” (1979) 13:2 Intl Lawyer 269 at 276.

\textsuperscript{26}Schwebel, \textit{ibid} at 83-86.

\textsuperscript{27}\textit{New York Convention}, \textit{supra} note 3, art V.

\textsuperscript{28}\textit{Ibid}, art V(2).

\textsuperscript{29} Schwebel, \textit{supra} note 22 at 83-88
then enforcement of the award can be denied under Article V(1)(a).\(^{30}\) An individual lacks capacity if they are unable to enter into a contract due to factors that might affect their judgment, such as intoxication at the time of signing the contract.\(^{31}\) Furthermore, if the agreement is invalid under the contract laws of the country, then the arbitration agreement and any award originating from such an agreement is not enforceable. For example, in *Dallah v Pakistan*,\(^ {32}\) the arbitration agreement was unenforceable as the government was found to not have agreed to it. Thus, at the time of drafting the arbitration clause or agreement, parties must take into account the applicable law that controls the formation of their contract. For example, a contract that provides for the sale of cannabis will be permitted in Amsterdam but will not be enforceable in the UK.\(^ {33}\) Although an arbitration agreement may not be valid under the laws of the state that it is subject to, the arbitration agreement may still be enforceable, as the defence of invalidity under state laws only acts as a persuasive mechanism to annul the award.\(^ {34}\)

**Notice / Ability to Present a Case**

The second ground stipulated by the *New York Convention* under which enforcement of an arbitral award may be denied is when a party is given insufficient notice of the arbitration or was otherwise not able to present its case.\(^ {35}\) This was reaffirmed in *Iran Aircraft Industries v Avco Corp*,\(^ {36}\) where it was held that Avco was not denied the opportunity to present their full case, and thus the arbitral award was enforceable. Typically, there is a very narrow threshold which has to be met in order to rely on this defence. In *Parsons &

\(^{30}\) *New York Convention*, *supra* note 3, art V(1)(a).

\(^{31}\) van den Berg, “Setting Aside”, *supra* note 24 at 267.

\(^{32}\) *Dallah Real Estate and Tourism Holding Company v The Ministry of Religious Affairs, Government of Pakistan*, [2010] UKSC 46 [*Dallah v Pakistan*].


\(^{34}\) According to the language used in Article V, if a defence is breached under Article V the award “may be refused”, rather than “will be refused”. *New York Convention*, *supra* note 3, art V.


\(^{36}\) *Iran Aircraft Industries v Avco Corp*, 980 F (2d) 141 at 143 (2d Cir 1992).
Whittemore Overseas Co,\textsuperscript{37} the court held that the petitioner was not prohibited from presenting his case where a tribunal refused to accommodate the timing of the presentation of a key witness. The court held that the arbitration tribunal acted within its discretion in declining to reschedule a hearing despite the inconvenience caused to the petitioner’s witness, and therefore the award was enforceable. Furthermore, in the case of Goetech-Lizenz-AG \textit{v} Evergreen,\textsuperscript{38} the defendant failed to present witnesses at a scheduled arbitration. It was thus argued by the defendants that the arbitrator’s award was not valid pursuant to Article V(1)(b); however, it was held that the defendant was not denied the opportunity to present his defence, because he had notice of the arbitration but chose not to respond.

**Jurisdiction**

Thirdly, under Article V(1)(c), arbitration agreements can be denied enforcement if there are jurisdictional issues present. The \textit{New York Convention} stipulates that the award must deal with issues within the scope of the agreement.\textsuperscript{39} For example, in \textit{Fertilizer Corp of India v IDI Management},\textsuperscript{40} the arbitrator awarded consequential damages despite the contract expressly prohibiting such damages. The losing party took the matter to court on the basis that the arbitrator had acted \textit{ultra vires}.\textsuperscript{41} However, a US court ruled that the arbitrators acted \textit{intra vires} and that the award was enforceable under the \textit{New York Convention}. Similarly, in \textit{Encyclopaedia Universalis SA v Encyclopaedia Britannica, Inc}, it was initially held that the arbitrator’s award was unenforceable because the “arbitrator exceeded their powers” because a third arbitrator was appointed when the agreement expressly provided for only two.\textsuperscript{42}

On appeal, the Second Circuit Court of Appeals affirmed the district court’s ruling denying confirmation of the award because the composition of the board of arbitrators was not in accordance with the parties’ agreement. However, the court reversed the district court’s ruling that the arbitrators

\textsuperscript{37} Parsons \& Whittemore Overseas Co \textit{v} Societe Generale De L’Industrie Du Papier, 508 F (2d) 969 at 975-76 (2d Cir 1974) [Parsons \& Whittemore].

\textsuperscript{38} Goetech-Lizenz-AG \textit{v} Evergreen, 697 F Supp 1248 at 1253 (EDNY 1988).

\textsuperscript{39} \textit{New York Convention}, \textit{supra} note 3, art V(1)(c).

\textsuperscript{40} \textit{Fertilizer Corp of India v IDI Management Inc}, 517 F Supp 948 at 958-60 (SD Ohio 1981).

\textsuperscript{41} RW Fleming, “Arbitrators and the remedy power” (1962) 47:6 Va L Rev 1199 at 1199-1225.

\textsuperscript{42} \textit{Encyclopaedia Universalis SA v Encyclopaedia Britannica Inc}, 403 F (3d) 85 (2d Cir 2005) at 92 [\textit{Encyclopaedia Universalis}].
exceeded their powers because arbitrators acting *ultra vires* is not one of the seven exclusive grounds for denying enforcement of an arbitration award under the *New York Convention*. This exemplifies the uncertainties and unpredictability of the Article V(1)(c) *ultra vires* argument, which contributes to the inconsistency and lack of clarity surrounding the defences under the *New York Convention*. What is clear, however, is that courts are likely to refuse granting the defence, except in the clearest of circumstances, such as when a reward is specifically prohibited by a contract, thus demonstrating their narrow construal of the *ultra vires* defence.

This position is unlikely to change due to the “competence-competence doctrine” that allows arbitrators to determine their own jurisdiction, thus making it difficult to challenge the award on jurisdictional grounds. However, the problem is more likely to arise where the point of law is manifestly wrong, or the question is of general public importance. The US courts allow a mistake in law to be challenged under Article V of the *New York Convention* if there has been a manifest disregard of the law, or if it can be proved that the arbitrator purposely misapplied the law. For example, in *Libyan American Oil Co (LIAMCO) v Socialist People’s Libyan Arab Jamahirya*, formerly Libyan Arab Republic, an attempt to elevate the issue of awarding consequential damages to an issue of law was rejected despite that the arbitration clause and the underlying commercial contract excluded any award of consequential damages. The court did not want to second guess the arbitrator’s construction of the parties’ agreement, nor act in the arbitrator’s role and emphasized that “the standard of review of an arbitration award by an American court is extremely narrow”. Therefore, had this been judged in another country where judicial scrutiny is higher, the outcome might have been different. However, given the influential role

---

44 However, the award from the case of *Encyclopaedia Universalis* was vacated under Article V(1)(d) defence.
47 *Bowen v Amoco Pipeline Co*, 254 F (3d) 925 at 932 (10th Cir 2001).
50 *Ibid* at 388.
played by US arbitration, it is likely this defence would be construed narrowly in most other countries. For instance, a German court agreed with the US arbitral tribunal’s award of higher interest than was claimed, and found no breach of jurisdiction.\textsuperscript{51}

**Procedure**

The next item in Article V, Article V(1)(d), states that if the composition of the arbitral body or the arbitral procedure was not in accordance with the agreement, or was not in accordance with the law of the seat of arbitration, then the award can be invalidated; thus, procedural impropriety could render an award unenforceable.\textsuperscript{52} For instance, in *Encyclopaedia Universalis*,\textsuperscript{53} the award was vacated because a third arbitrator was appointed in contradiction to the agreement. The decision was reaffirmed in *Indus Risk Insurers v MAN Gutehoffnungshütte GmbH*\textsuperscript{54} where, in contrast to *Encyclopaedia Universalis*, the court held that the panel’s incorporation of a late-filed report did not constitute a circumstance where “the arbitral procedure was not in accordance with the agreement of the parties.”\textsuperscript{55} Similarly, in *Satyam Computer Services, Ltd v Venture Global Engineering, LLC*,\textsuperscript{56} the respondent argued that the arbitrator violated agreed-upon terms by applying New York law when the agreement contained a provision stating that Michigan law should govern. It was held that there was no procedural impropriety because both laws were similar. The aforementioned court decisions illustrates the very narrow construal of the procedural impropriety defence.\textsuperscript{57}

**Award not yet binding**

Further, under Article V(1)(e), an arbitration award can be denied enforcement if the award had been suspended or set aside. However, this is

\textsuperscript{51}``Aloe Vera of America Inc (US) v Asianic Food (S) Pte Ltd (Singapore) and Another” (2007) XXXII YBCA 489.
\textsuperscript{52} David Lipton, “The Standards on Which Arbitrators Base Their Decisions: The SRO’s Must Decide” (1988) 16 Sec Reg LJ 3.
\textsuperscript{53} *Encyclopaedia Universalis*, supra note 42.
\textsuperscript{54} *Industrial Risk Insurers v MAN Gutehoffnungshütte GmbH*, 141 F (3d) 1434 at 1442 (11th Cir 1998).
\textsuperscript{55} Ibid.
\textsuperscript{56} *Satyam Computer Services, Ltd v Venture Global Engineering LLC*, not reported in F Supp (2d) (6th Cir 2007).
\textsuperscript{57} Junita, supra note 35 at 140-64.
Defences to the New York Convention 1958

controversial because although Article 34(2) of the 2010 UNCITRAL Arbitration Rules indicates that arbitral tribunal awards are binding, nonetheless, some national courts, such as the French court in *Putrabali v Rena*,

58 investigate whether awards are permissible under their law.

59 The inclusion of local requirements gives courts at the seat of arbitration the ability to find fraud or corruption in an arbitral tribunal. In *Chromalloy v Egypt*,

60 a US court addressed an arbitration agreement between the Egyptian Air Force and an American company in which the parties had expressly agreed that the losing party would not seek review of the arbitral award. While the American company’s petition for enforcement of its award was pending before the district court, Egyptian Air filed an appeal with the Egyptian Court of Appeal to nullify the award. The district court refused to recognise the decision of the Egyptian court, finding that to do so would reward or be in support of Egyptian Air’s breach of the express agreement. As a result, the US court enforced the award (that was already set aside in Egypt) by contrasting the permissive nature of Article V with the mandatory nature of Article VII that entitles the use of local laws if they are more favourable for enforcement of arbitral awards than the New York Convention itself would be.

However, national courts generally decline to enforce annulled awards, and therefore construe the defence narrowly. For example, in *Baker v Chevron*,

62 and *TermoRío v Electrificadora*,

63 the US courts did not enforce the awards that had been set aside by the courts at the seat of arbitration. Furthermore, in one case a German court refused to enforce an award that had been set aside in Russia*

64 and in *EDF v Endesa*,

65 a Chilean court refused to enforce an award set aside in Argentina. While courts generally refuse enforcement of awards set aside in other States, they may be swayed to

---


61 Ibid at 909-10.


63 *TermoRío SA ESP v Electrificadora Del Atlanticoc SA ESP*, 421 F Supp (2d) 87 (DDC 2006).


65 “EDF Internacional SA v Endesa International SA and YPF SA, Supreme Court of Chile, 8 September 2011” (2012) 5 Arbitraje: Revista de Arbitraje Comercial y de Inversiones 915.
consider such matters where basic notions of justice are challenged. Nonetheless, the ability to set aside an award, which in theory offers little benefit to either party as it costs time and money, is another example of the uncertainties surrounding the use of this defence. The mere fact that speed and low cost are two of the fundamental advantages of arbitration could mean this defence is contrary to these principles. Nevertheless, courts have tried to infuse some degree of clarity in this area, as evidenced by the approach taken in Continental Transfer Technique Ltd v Federal Government of Nigeria, where it was held that a critical element is the seat of arbitration: “If that place [where the award is sought to be enforced] is in the territory of a party to the Convention, all other Convention States are required to recognise and enforce the award, regardless of the citizenship or domicile of the notice.”

The factors courts take into account when making a determination on whether to stay the request for recognition and enforcement include the general objectives of arbitration, which are the expeditious resolution of disputes and the avoidance of lengthy and expensive litigation. In addition, courts consider whether the award sought to be enforced will receive greater scrutiny in the foreign proceedings under a less deferential standard of review. Furthermore, courts consider the characteristics of the foreign proceedings, such as whether they were brought to enforce an award. As previously discussed, primary jurisdiction is possessed by courts in the country in which or under whose law the award was made, and these courts can apply domestic law in scrutinizing the award. This means that if a losing party has petitioned a court of the country with primary jurisdiction to vacate an award, a court with secondary jurisdiction (i.e. a country under which law the award was not

---

66 Corporación Mexicana de Mantenimiento Integral, S De RL de CV v Pemex-Exploración y Producción, 962 F Supp (2d) 642 (SDNY 2013).
67 If an award is set aside, the court automatically gains jurisdiction of the case as the parties cannot re-appeal the decision to arbitration. For further reading, see Jie Lin, “The application of the delocalisation theory in current international commercial arbitration” (2011) 22:12 Intl Co & Com L Rev 383.
68 Creighton Limited v Government of the State of Qatar 181 F (3d) 118 (DC Cir 1999) citing Restatement of the Law (Third), vol 2, Foreign Relations Law of the United States (St Paul, Minn, 1987) at comment b. See also Continental Transfer Technique Ltd v Federal Government of Nigeria, 697 F Supp (2d) 46 (DDC 2010); see also Termorio SA ESP v Electranta SP, 487 F 3d 928 (DC Cir 2007).
rendered) “may, if it considers it proper, adjourn the decision on enforcement of the award and may also, on application of the party claiming enforcement of the award, order the other party to give suitable security”, leaving the final decision of whether to vacate the award to the court of primary jurisdiction.

Arbitrability of Dispute

Under Article V(2)(a), an arbitration agreement can be denied enforcement if the subject-matter of the dispute is not arbitrable under the law of the country where enforcement is sought. This concept is known as the ‘arbitrability’ concept. It means that in contrast to commercial contracts, fair adjudication of public interest matters - such as family, criminal or bankruptcy disputes - cannot be done in a private arbitration forum, where the ability of courts to control the process, the decision making power, and the application of the law is strictly limited. These public law issues have to be raised and tried before an open/public court. On these grounds, for example, a Russian court declined to enforce an arbitral award holding that corporate governance matters are not arbitrable. The court stated that some matters concerning public interest must be resolved through a public domain.

Like other defences, and despite the existence of this defence, it is rarely successful in practice.

Public Policy

Finally, the New York Convention provides that if recognition or enforcement of an award would be contrary to the public policy of the country where enforcement is sought, then the award can be set aside. However, the New York Convention does not define “public policy,” leaving the definition broad and open to interpretation. For example, in Parsons & Whittemore Overseas Co Inc v Societe Generale de L’Industrie du Papier, public policy was held to be applicable only where enforcement would violate the

71 New York Convention, supra note 3, art VI.
74 New York Convention, supra note 3, art V(2)(b).
75 Parsons & Whittemore, supra note 37.
State’s most basic notions of morality and justice. A basic notion of Saudi Arabian public policy is that the award of monetary interest, *riba*, is forbidden, thus enforcement of an interest award would be contrary to Saudi public policy.\footnote{Interest means money paid regularly at a particular rate for the use of money lent, or for delaying the repayment of a debt. See Kristin T Ray, “The New York Convention and Saudi Arabia: Can a Country Use The Public Policy Defence To Refuse Enforcement of Non Domestic Arbitral Awards?” (1995) 18 Fordham Int'l LJ 920.} Furthermore, in *Karaha Bodas Co v Perusahaan Pertambangan*, District Judge Rosenthal stressed that “the general pro-enforcement bias informing the Convention points to a narrow reading of the public policy defence.”\footnote{*Karaha Bodas Co v Perusahaan Pertambangan Minyak Dan Gas Bumi Negara*, 364 F (3d) 274 at 288 (5th Cir 2004). On Chinese public policy, see *Hemofarm et al v Yongning*, Min Si Ta Zi No. 11, People's Court Press, 2009 (Supreme People's Court of China, June 2008).} This case is significant because it provides a clear example where a court is directly endorsing the narrow approach.

English courts have also construed this defence narrowly, with one commentator remarking that the defence may never be applied.\footnote{*The English Arbitration Act provides that a party may challenge an award based on “serious irregularity affecting the tribunal, the proceedings or the award”. *English Arbitration Act*, supra note 46, s 68.} However, in *Soleimany v Soleimany*,\footnote{*Soleimany v Soleimany* (1998), [1999] QB 785 (CA Eng).} the English courts applied the public policy defence broadly to prevent arbitration facilitating illegal contracts.\footnote{*Ibid.*} Even though the public policy defence was successful in *Soleimany*, it was nevertheless construed narrowly in *Westacre Investment Inc v Jugomport-SPDR Holding Co Ltd*\footnote{*Westacre Investment Inc v Jugomport-SPDR Holding Co Ltd* (1999), [1999] CLC 1176 (CA Eng) [Westacre].} where the court rejected the allegation that the award could not be enforced due to fraud and corruption as this issue was already *res judicata*.\footnote{A matter that has already been adjudicated and can no longer be pursued by the parties involved.} The court suggested that they were in support of arbitration and less inclined to venture into public policy issues pertaining to corruption.\footnote{Jan Paulsson, “Delocalisation of International Commercial Arbitration: When and Why it Matters” (1983) 32 ICLQ 53.} Furthermore, the US court in *Sonatrach (Algeria) v Distrigas (United States District Council)*\footnote{*Sonatrach (Algeria) v Distrigas Corporation*, 80 BR 606 (Massachusetts 1987).} also suggested that public policy defences should be construed narrowly, because in an increasingly interconnected world parties should be able to
freely negotiate internationally recognized and binding agreements.\textsuperscript{85} Thus, the US endorses the view that a foreign award would have to violate a State’s most basic concepts of morals and justice in order to deny enforcement.\textsuperscript{86} Similarly, Indian courts have endorsed this approach by emphasizing that public policy defences should only apply to matters of national interest or justice and morality.\textsuperscript{87}

It is, however, difficult to make a determination on whether this defence is uniformly construed narrowly, as States have differing interpretations of public policy and not all States have construed the public policy defence narrowly. For instance, the UAE has extended the public policy defence to ownership rights over property.\textsuperscript{88} Russian courts have used public policy to decline enforcement of awards where damages are disproportionate to the breach.\textsuperscript{89} Under the Chinese law, public policy includes social interests in addition to legal principles.\textsuperscript{90} This approach is reinforced by an obligation on the Chinese courts to consider public policy factors, which are wide, in every case.\textsuperscript{91} For example, in the famous \textit{Heavy Metal} case,\textsuperscript{92} a Chinese court rejected enforcement of an arbitral award because the performance of heavy metal music went contrary to national sentiments as well as societal interests. It follows that an award could be denied enforcement where one State does

\textsuperscript{85}Ibid at 612; \textit{Federal Arbitration Act}, USC 9 (1925), ss 10(a) outlines criteria as where: (1) “the award was procured by corruption, fraud or undue means (2) there was evident partiality or corruption in the arbitrators, (3) the arbitrators were guilty of misconduct ...or...other misbehaviour by which the rights of any party have been prejudiced.”; Roberts S Matlin, “The Federal Courts and the Enforcement of Foreign Arbitral Awards” (1984) 5 Pace L Rev 215 at 179.

\textsuperscript{86}Parsons & Whittemore, supra note 37 at 974.


\textsuperscript{88}Batti Real Estate Development v Dynasty Zarooni Inc, Petition No. 14 (Dubai Court of Cassation 2012) (United Arab Emirates).

\textsuperscript{89}Presidium of Supreme Commercial Arbitrazh Court, Information Letter No 156, 26 February 2013, s 6 (Russia).

\textsuperscript{90}Lanfang Fei, “Public Policy as a bar to enforcement of international arbitral awards: A review of the Chinese approach” (2010) 26 Arb Intl 301 at 303 [Fei].

\textsuperscript{91}Randal Peerenboom, “Seek truth from the facts: An empirical study of enforcement of arbitral awards in the PRC” (2001) 49 Am J Comp L 39; Explanations on and Answers to Practical Questions in the Trial of Foreign-Related Commercial and Maritime Cases No 1, (Issued by the Supreme People’s Court, April 8 2004), art 43 (China).

\textsuperscript{92}Fei, supra note 90 at 307, citing \textit{American Production Co and Tom Flight Co v Chinese Women’s Travel Agency Ta}, 1997 No. 35 (Sup. People’s Ct. reply, issued on 26 December 1997) [\textit{Heavy Metal}].
not tolerate a behaviour that it considers contrary to societal interest while such an award remains enforced in another State despite allegations of fraud and impropriety, as in Westacre. Although respect for a State’s culture is understandable, public policy should not be used as a barrier to enforcement of arbitral awards for seemingly archaic matters. A line must be drawn between social factors and threats to a nation’s justice and security.

**Proper Forum for Collecting under the Award**

Upon securing an award, a winning party will seek to collect. Commonly, the award is enforced under the New York Convention provided that the debtor has assets within the State where the enforcement is sought. However, in an event where a State has no jurisdiction over the debtor, it would not be appropriate under the principle of *forum non conveniens* ('not a convenient or appropriate forum') to grant enforcement where an adequate alternative forum exists. 93 Although the *forum non conveniens* principle is construed narrowly for the purposes of recognition and enforcement of international awards, it can nonetheless be problematic and lead to uncertainty. For example, in *Base Metal Trading Ltd v OJSC Novokuznetsky Aluminum Factory*, it was held that the property against which the collection under the award was sought had to relate to the claimants’ cause of action, 94 while in *Glencore Grain Rotterdam BV v Shivnath Rai Harnarain* it did not have to be so related. 95 Thus, the principle of *forum non conveniens* adds another domestic law consideration for the enforceability of awards.

**Estoppel**

Estoppel has been raised as an additional ground for refusing recognition and enforcement of an award. A party may be found to have waived its right to oppose a recognition and enforcement request if it did not act on an

---


94 *Base Metal Trading, Ltd v OJSC Novokuznetsky Aluminum Factory*, 283 F (3d) 208 (CA 4th Cir 2002).

95 *Glencore Grain Rotterdam BV v Shivnath Rai Harnarain Co*, 284 F (3d) 1114 (CA 9th Cir 2002); see also *Monegasque de Reassurances v Nak Naftogaz of Ukraine and State of Ukraine*, 311 F (3d) 488 (CA 2nd Cir 2002). In this case there was uncertainty as to whether the correct party was before the court and since it was determined that Ukrainian law could resolve the issue, the matter was resolved in a Ukraine court.
alleged impropriety despite becoming aware of it during the arbitral process. Estoppel was argued, for example, in AAOT v Trade Services Inc, where the party did not raise the alleged procedural impropriety until after the award was rendered. The court found that the party’s failure to act at the time was fatal to the argument.

Furthermore, in La Societe Nationale Pour la Recherche La Production, Le Transport, La Transformation et La Commercialisation des Hydrocarbures v Shaheen Natural Resources Company Inc, Shaheen lost its objection over the “alleged untimeliness of the award” because it was not raised before the tribunal in a timely manner. The court concluded in this case that refusal of the arbitral award at this stage would go contrary to the goal and the purpose of the New York Convention.

THE WAY FORWARD

The above analysis has shown that grounds for refusing recognition and enforcement under the New York Convention are numerous, and can be subject to different interpretation by the courts of the country where an award is sought to be enforced. The differences in interpretation contribute to the absence of a uniform international approach to Article V.

A number of academic commentators have made suggestions on ways of reforming enforcement under the New York Convention in order to create a uniform approach. Lu made a similar observation by arguing that arbitration panels are free to choose how they conduct hearings because the New York Convention is too vague. Lu’s criticism gives rise to a proposal for a non-exhaustive international list outlining minimum standards to be met by international arbitration panels. Developing uniform standards will prevent recurring inconsistencies generated by domestic interpretation. It will provide guidelines for parties to follow, leading to predictability and fair outcomes.

---

96 AAOT Foreign Econ Association (VO) Technostroyexport v International Development & Trade Services Inc, 139 F (3d) 980 (CA 2nd Cir 1998).
Wheeless further argued that in order to encourage the use of arbitration, arbitral tribunals should adhere to procedural rules. Quigley agrees with Wheeless, that justice is at the heart of the grounds for refusing recognition and enforcement under the New York Convention because they embody the basic notion of due process. This is reasonable, and ought to be adopted as a universal modus operandi. Furthermore, Mehren argues that the grounds for refusing enforcement are wide and cover all eventual irregularities, thus promoting fairness and equality to those who choose arbitration. The privilege of taking the matter to court helps to overcome the involvement of the arbitration panel and provides the parties to arbitration with confidence of achieving justice from an impartial body. However, the whole purpose of arbitration is to avoid court intervention; therefore, broad application of the defences could undermine the whole process.

Paulson does not support the arbitration process being scrutinized with court intervention, but concedes that this is unavoidable because when an arbitration award requires enforcement or appeal, the parties have no choice but to resort to courts. Parties can limit courts’ intervention by relying on limitation or reservation clauses in the arbitration agreement. A limitation clause would provide parties the opportunity to decide on a course of action in the event of an appeal or any challenge with respect to the award. For example, on the occurrence of such a challenge, parties can agree in advance to take the matter to another panel of arbitrators. This option will enable parties to avoid their matters automatically going to a local court.

Furthermore, with the continuing growth of arbitration comes an influx of complex cases, which involve multiple contracts with multiple parties from various States. Although domestic courts are competent and authoritative adjudicators, the logistics and complexity of local procedures can be an

impediment to handling a process that was designed to be streamlined and responsive to the commercial reality. Moreover, participation of a State as a party opposing the recognition and enforcement process in its local court adds an extra layer of complexity and may pose a challenge to the adjudicator, especially in countries where the rule of law is less developed. It would therefore be in the interests of justice if adjudication of challenges to arbitral awards were done in an international court forum. As a way forward, the creation of a world court of arbitration that has exclusive jurisdiction to assess whether the grounds under Article V are properly invoked is desirable and merits serious consideration.

On that analysis, it is clear that grounds to refuse recognition and enforcement under the New York Convention are constructed narrowly within its parameters. In fact, the reasons to refuse enforcement should be grounded in the principles of fundamental justice as the only persuasive grounds for challenging arbitral awards.\(^{103}\) Although the majority of local courts appear to agree with this concept, some States use public policy as a means of limiting enforcement of arbitral awards. On balance, however, it is reasonable to conclude that the New York Convention is internationally respected and adhered to given that courts endorse ninety-five per cent of international arbitral awards.

CONCLUSION

In conclusion, there seems to be a consensus amongst member States on the approach that should be used to interpret the enumerated grounds invoked to deny recognition and enforcement of international arbitral awards under the New York Convention. Courts encourage narrow application of the defences in order to preserve the autonomy of arbitration and promote it as an alternative dispute resolution mechanism. However, a lack of uniformity and consistency in arbitral practice remains a major challenge. This calls for greater international cooperation and consultation among judiciaries and international arbitral tribunals in order to achieve uniformity and consistency in international arbitration.

\(^{103}\) van den Berg, “Uniform Interpretation”, supra note 23 at 268.
Faltering Blocks in the Arguments against Unitary Taxation and the Formulary Apportionment Approach to Income Allocation

ALEXANDER EZENAGU *

SYNOPSIS

Tax treaties represent a highly developed area of international cooperation. They are the primary means by which countries cooperate to avoid double taxation and remove barriers to international business, but they have also become tools used by multinationals to avoid or evade taxation on a global scale. Integral to the present day model of tax treaties is the separate accounting and arm’s length standard of income allocation (SA-AL), adjudged to be fraught with challenges and accountable largely for the widespread tax evasion and avoidance practices of the modern era.

While this existing method for measuring and taxing the income of multinationals is increasingly unsatisfactory, viable alternatives have traditionally been rejected as equally unworkable. This paper will make the case that the alternatives are actually becoming increasingly viable and should be adopted going forward, despite the

* Alexander Ezenagu is a PhD Candidate in International Tax Law at McGill University and holds a Master of Law degree (LLM) from the University of Cambridge. His doctoral research focuses on national and international tax law and policy issues, with emphasis on the relationship between taxation and economic development and the role of government and non-government institutions and actors in the creation of tax policy norms.
inevitable transition costs that will arise in the short term. First, this paper will set out the two distinct approaches to taxing multinational: the treatment of companies as separate entities with separate accounting and an arm’s length standard of transfer pricing or rather as integrated entities with total net income being apportioned among integrated entities by means of a uniform apportionment formula. Second, it will explain why one method has been chosen instead of the other, further explaining why the chosen method is failing and why a viable alternative is necessary. Third, it will explore why the alternative is simultaneously becoming more viable despite past resistance. The paper will conclude that the alternative is now possible and should be adopted before the current method fails completely.

INTRODUCTION: CONCEPTS AND HISTORY

This paper focuses on limitations with regards to the tax treatment of multinational entities (MNEs) under the existing global system, which treats related entities as separate and further demands they act at arm’s length when dealing with each other. It further seeks to advance an alternative global tax system where MNEs are treated as single entities for tax purposes with their global profits apportioned among jurisdictions based on the value of the economic activities taking place in those jurisdictions. It advocates for a unitary taxation and formulary apportionment approach to income allocation arising from the cross-border economic activities of MNEs.

The paper opens by defining two concepts: the separate entity and arm’s length standard (SE-ALS) of income allocation and the unitary taxation and formulary apportionment approach (UT-FA) of income allocation. It provides a history of the existing system, its limitations and efforts by supranational bodies to improve it. It considers the arguments against the alternatives to UT-FA, why those arguments are neither as relevant nor potent today as they were decades ago, and ultimately makes recommendations. It concludes by arguing that the transitional cost is not as costly as the negative effects of continuing and defending the existing system in the long run.

Concepts

Arm’s Length Standard

The global consensus to treat multinational entities (MNEs) as separate entities requires that in the transfer of goods and services among entities within the MNE, especially with regards to setting the prices of such transfers, they act at arm’s length. This implies that while they are integrated entities
under the ownership and control of a parent company (a central management and vertical organizational structure) they must act as independent entities would. This treatment of subsidiaries by a parent company is contained in Article 9 of both the OECD Model Treaty and the UN Model Treaty. It provides that:

Where:

a) An enterprise of a contracting state participates directly or indirectly in the management, control or capital of an enterprise of the other contracting state, or

b) The same persons participate directly or indirectly in the management, control or capital of an enterprise of a contracting state and an enterprise of the other contracting state,

and in either case, conditions are made or imposed between the two enterprises in their commercial or financial relations which differ from those which would be made between independent enterprises, then any profits which would, but for those conditions, have accrued to one of the enterprises, but, by reason of those conditions, have not so accrued, may be included in the profits of that enterprise and taxed accordingly.¹

This treatment is equally applied to the treatment of branches or permanent establishments (PEs) of companies resident in other jurisdictions. Article 7(2) of the model convention provides that:

...the profits that are attributable in each Contracting State to the permanent establishment referred to in paragraph 1 are the profits it might be expected to make, in particular in its dealings with other parts of the enterprise, if it were a separate and independent enterprise engaged in the same or similar activities under the same or similar conditions, taking into account the functions performed, assets used and risks assumed by the enterprise through the permanent establishment and through the other parts of the enterprise.²

To achieve the arm’s length treatment of MNEs in the transfer of goods and services among them, methodologies have been recommended by supranational bodies such as the OECD and the UN to guide taxpayers and tax authorities in agreeing on a price for the transfer of goods and services that would be deemed to be arm’s length. Three of these methods are aimed at the transaction or the price charged for the goods and services, and the other two are aimed at the profit declared by the taxpayer for the overall transaction.

² Ibid, art 7(2).
The traditional transactional methods are the Comparable Uncontrolled Price Method (CUP), the Resale Price Margin (RPM) and the Cost Plus Method (CPM). The traditional profit methods are the Transactional Net Margin Method (TNMM) and the Profit Split Method (PSM). For all five methods, the taxpayer must include comparables in justifying the arm’s length standard, that is, the taxpayer must show that the price fixed or the profit declared is the same as would be achieved by unrelated parties dealing with each other under the same conditions and on the same terms. There is a provision for adjustment by the tax authority where this may not completely be the case.

The taxpayer needs to convince the tax authority by preparing transfer pricing (TP) documentation – a process that is resource-intensive, demands expertise, incurs great costs and takes time. Where the tax authority is not convinced, it may re-characterize the transaction and adjust the profits.

**Unitary Taxation and Formulary Apportionment**

Unitary taxation (UT) operates from the understanding that the profits generated by an integrated firm arise from the integration of its activities. Under this approach, a consolidated account for the integrated firm is furnished and profits apportioned based on factors that reflecting the economic activities in each jurisdiction in relation to the functions performed, assets used and risks assumed. Cobham and Loretz describe UT as:

...treating each multinational group of companies as a unit, regardless of the geographical and jurisdictional location of the individual subsidiaries; calculating profit and loss on a group-wide basis; and then allocating the taxing rights on this consolidated profit between the jurisdictions with which the group has a nexus, according to the extent of the economic activity.

---

3 Prem Sikka, “How to Take a Serious Bite out of Corporate Tax Avoidance”, *The Guardian* (24 May 2013), online: <https://www.theguardian.com/commentisfree/2013/may/24/corporate-tax-avoidance-unitary-taxation-g8>. Here, the author argues that though unitary taxation is not a magical solution to the deep-seated problems of capitalism, its strong points call for its adoption.

This approach starts with the basic idea that both the residence state and the source state have a co-existing interest in the combined income.\(^5\) It further considers an MNE as a single business, which for the sake of convenience, is divided into purely formal, separately-incorporated subsidiaries. Under this approach, “the global income of the MNE needs to be computed, then such income is apportioned between the various component parts of the enterprise by way of a formula which reflects the economic contribution of each part to the derivation of profits”\(^6\). Consequently, a state may subject that portion of the multi-national corporate group’s income deemed attributable to the state to taxation based upon the group’s presence or activities within the state.

In the United States, where the unitary taxation and formulary apportionment approach of income allocation has been used for decades when allocating profits among taxing jurisdictions within the US, the Supreme Court established the pillars of a unitary business to be: functional integration, centralized management and economies of scale.\(^7\) Segal further opines that the Unitary Taxation and Formulary Apportionment (UT-FA) approach differs from the Separate Accounting and Arm’s Length Standard (SA-ALS) in two main aspects: the fact that intra-group transfers negate each other as the tax is based on the worldwide income of the group rather than the individual income of some particular unit of the group; and the unitary method takes into account the income of the unitary group as a whole rather than merely transactions undergone directly by the unit within the taxing jurisdiction.\(^8\)

Notwithstanding the presence and use of the UT-FA approach for over a century, the OECD has consistently refused to consider its application on a global scale. In the 2017, transfer pricing guidelines published by the OECD stated that “...no legitimate alternative to the arm’s length principle has emerged. Global formulary apportionment, sometimes mentioned as a possible alternative, would not be acceptable in theory, implementation, or

---


Asper Review

It further stated that the most significant concern with global formulary apportionment is the difficulty in implementing the system in a manner that both protects against double taxation and ensures single taxation. Julie Roin holds the view that formulary apportionment will be vulnerable to the well-known avoidance techniques present in a separate-entity/transactional environment, further suggesting that the perceived benefits associated with formulary apportionment may not exceed the associated costs, calling on national tax policy makers to more effectively focus on ways to buttress the separate entity/transactional approach.

History: The Journey to the Existing System

The end of World War I saw the immediate commencement of a second war, a resource war. International trade was gaining steam in the 19th century and trade, shipping, and aviation agreements expanded rapidly as countries were committed to rebuilding their economies and territories after the damage caused by war. These expansionary plans and the associated cross-border trade meant that states were further determined to secure and increase their revenue streams. The MNEs, which had become the vehicles for international trade, were worried about double taxation of their income arising from trade engagements with more than one country, for example their home state and the host state. These concerns, fears, and worries brought countries to agree on the prevention of double taxation; hence, the birth of international tax treaties on the prevention of double taxation. All stakeholders agreed that taxation, and the risk of double taxation, presented hindrances to economic relations between countries.

In 1920, the International Chamber of Commerce (ICC) advocated for a profit-split method of income allocation, and stated that profits should be

---


10 Ibid.


taxed in each country in proportion to the profit realized therein.\textsuperscript{13} The ICC further suggested that where countries could not agree, allocation would be presumed to be proportional to sales (turn-over), effectively advocating for the adoption of the formulary apportionment mechanism in income allocation.\textsuperscript{14} The ICC’s work in addressing double taxation of global income was transferred to the League of Nations in 1923,\textsuperscript{15} which favored the primacy of residence in the allocation of taxing rights.

This meant that residual income was allocated to residence countries while source countries were left to collect withholding taxes on the categories of income where permanent establishment was decided by the state. Experts also recommended that companies be treated as separate entities rather than a global company for the purpose of taxation. The implication of this is that companies treated transactions with related companies as independent, fixing prices that reflect market in the transfer of prices amongst them.\textsuperscript{16}

Wells and Lowell have argued that the:

...framework of taxation that evolved in the 1920s was based on the mercantilist belief that imperial countries were the source of capital and know-how while the colonies were passive suppliers of goods or services with little value added functionality. As a result, the right to tax residual income belonged to the residence countries of the imperial companies (England in this example). Source countries (India in the example) were allowed to tax only routine profits deemed earned therein and impose withholding taxes on certain types of outbound payments.\textsuperscript{17}

They further argue that the League of Nations model was constructed from distinct policy judgments, such as:

(i) source country should tax local operations;
(ii) residual income should be earned by the residence country;
(iii) the presence of an interim holding company in a country should cause that country to be treated as a residence country;
(iv) subsidiaries, by themselves, should not be treated as permanent establishments of the offshore parent company; and
(v) transfer pricing is to be applied on a separate account basis (collectively, the “foundational premise”).\textsuperscript{18}

\textsuperscript{13} Wells, supra note 5.
\textsuperscript{14} Ibid.
\textsuperscript{15} The League of Nations was established by the Treaty of Versailles on June 28, 1919, with the aim of promoting international cooperation and to secure international peace and security.
\textsuperscript{17} Wells, supra note 5 at 10.
\textsuperscript{18} Ibid at 27.
These five principles would come to define tax relations between countries and influence both the current OECD and UN Tax Treaty.

It is important to note that the economic experts at the League of Nations framed and stated the issues as involving a conflict of interest between debtor countries who import capital, and creditor countries who export capital. They further posited that double taxation should be avoided by vesting the primary taxing jurisdiction in the country to which the taxpayer owed its “economic allegiance”, relying on Schanz’s economic allegiance theory. In the end, a classification-and-assignment approach was adopted, which classified income into various categories and assigned taxing jurisdiction based on the classification.

Schanz’s economic allegiance theory provided that revenue should be apportioned according to contribution, and recommended a 75-25 sharing formula between source and residence states. Relying on Schanz’s theory, it is difficult to appreciate the classification-and-assignment approach used by the League of Nations. From a logical standpoint, the proper choice should have been option 3 (the proportional allocation of income) as recommended by Schanz in his economic allegiance theory.

This model of income allocation was subsequently accepted by the Organisation for European Economic Co-operation (OEEC) and the Organisation for Economic Cooperation and Development (OECD). It has become the basis for treaty negotiations amongst states. Thus, the ICC model approach of treating global companies as unitary for tax purposes and apportioning the residual income of a global company between the residence and source countries based on economic contributions conducted in each country (formulary apportionment) was rejected in favour of treating entities as separate and taxing based on classification and assignment of income, which favored the country of residence.

Arguments against the UT-FA approach to income allocation include the following: the jurisdiction of a state to tax income should be limited to

19 Ibid at 15.
21 The OECD is an advisory organization for economic cooperation established in 1961, which currently has 35 member countries.
23 See generally Picciotto, supra note 16.
income derived from its territory; only a state of residence may tax the worldwide income of a taxpayer; tax authorities do not have the right to demand information about an enterprise’s business establishments in other states; the exercise of cross-border tax controls could contribute to animosity; global calculations of income according to the different principles of the various states would impose heavy administrative burdens on enterprises; differences in accounting principles, languages, currencies, etc. would create practical problems; profit potentials differed from country to country; and that it would be difficult to reach agreement on common rules for the calculation of the worldwide income.

Inhibitory factors in the election of the UT-FA approach include: on ‘consolidating or combining’ separate entities, the question is whether to delineate the group in terms of legal or economic relationships; do we apply the formulary apportionment to worldwide income or a “water-edge” limitation; what is an acceptable formulary apportionment formula; what are the factors to include in such formula and the definition of the factors such as payroll, property and sales; how do we reconcile the challenge posed by the divergence between financial and tax accounting in adopting formulaic apportionment approach. Further, what income is to be apportioned, should it include dividends, interest, patent and copyright royalties, rents and royalties from real estate and income from the purchase and sale of real property or capital assets? How do we treat intangibles and the advent and prominence of e-commerce/digital economy? 24 Other arguments regularly advanced against the election of the UT-FA approach are: fiscal sovereignty of states; hurdles to exchange of information among states; administrative burdens of transition to a new global tax system; and political will of nations to agree on formula, factors, accounting standards, language of reporting, profit determination and delineation of integrated entities. 25

In summary, issues of sovereignty, exchange of information, agreement on appropriate factors and formulae, and implementation hurdles, have all


influenced the decision of the experts. I set out to address these concerns below.

A FAILING SYSTEM

Litany of Woes

The League of Nation’s election to use a classification-and-assignment approach has created a platform for multinational companies to earn a material portion of their income in low or no tax jurisdictions through base erosion and profit shifting. The argument quickly moved from double taxation to double non-taxation. It became obvious to the international community that companies manipulated the transfer price amongst themselves to pay low or no tax in jurisdictions where they operated from, and equally shifted profit to low tax jurisdictions or eroded tax bases in countries with high tax rates. The League of Nations Model presented the opportunity for rent-seeking and manipulation of the system. It equally provided substantial advantage to MNEs over non-MNEs, thereby raising horizontal equity issues. A consequence of transfer pricing manipulation by MNEs is that national tax jurisdictions suffer economic stress and may have to bridge the revenue gap by taxing middle or low income earners at a higher rate than they are able to pay.

In more concrete terms, the 2015 UN Conference on Trade and Development publication claims that developing countries lose $100 billion per year due to tax avoidance by multinational companies and as much as $300 billion in total lost development finance. Between 1970 and 2010, the

---


27 This is the unit price assigned to goods and services between the parent company and subsidiaries or between divisions within the same firm.

28 The “Ability to Pay” principle of taxation provides that taxation ought to be progressive, taxing those with the ability to pay more than those without the ability, however, the rent-seeking activities of MNEs have resulted in a regressive tax system in most tax jurisdictions and this trend is unjust. See Eduardo Baistrocchi, “The Transfer Pricing Problem: A Global Proposal for Simplification” (2006) 59:4 The Tax Lawyer 941 [Baistrocchi].

29 Tax Justice Network, “UNCTAD: multinational tax avoidance costs developing countries $100 billion+” (16 March 2015), online: <http://www.taxjustice.net/2015/03/26/unctad-multinational-tax-avoidance-costs-developing-countries-100-billion/>.
capital flight through tax evasion and avoidance schemes stood at $814 billion, exceeding the official development aid of $659 billion and foreign direct investment of $306 billion for the same period.\(^{30}\) Developed countries are not spared. In 2008, the US alone lost an estimated $90 billion to the practice of shifting profits overseas, which was about 30% corporate income tax revenues or three times the federal budget for foreign aid.\(^{31}\) In 2009, Barclays Bank declared global profits of £4.6 billion but paid only £113 million in UK corporation tax, an effective rate of 2.4%.\(^{32}\) Google was accused of generating £11.5 billion from the UK between 2006 and 2011, but paid just £10 million in UK corporation tax.\(^{33}\) Amazon paid £11.9 million in corporation tax in 2014 even though it made £5.3 billion in sales from the UK.\(^{34}\) Using tax-planning schemes, such as the ‘double Irish, Dutch sandwich’, companies are able to move the bulk of their revenues to low or no tax jurisdictions like Ireland, Luxembourg and Bermuda. From 2009 to 2012, Apple got away with sending $74 billion in profits to its Irish subsidiaries, even though Apple products were designed in the United States, assembled mostly in China, and sold in Europe, Africa, Asia, and the Middle East, with relatively few sales in Ireland. Apple was able to assign $74 billion to Ireland, by taking advantage of loopholes in tax treaties modeled after the League of Nations’ Model, alongside a secret tax deal with Ireland, which enabled Apple to pay a total effective tax rate of 1% in Ireland. Though Apple had three subsidiaries in Ireland, each claimed to have tax residency nowhere, which effectively led to tax dodging.\(^{35}\)


However, developing countries are worse hit by tax avoidance schemes because, unlike developed countries, they rely heavily on corporate tax revenue. Durst aptly describes the situation thus:

Countries with highly developed economies often place limited reliance on revenue raised from corporate taxation because other sources of revenue, such as individual income taxes and consumption taxes, are arguably sufficient to meet national needs. In many developing countries, however, much domestic economic activity occurs informally, with few if any books and records maintained, so the government’s ability to raise revenue from individual income taxes and consumption taxes is severely limited. For these countries, corporate taxation, and especially taxation of income from cross-border operations, represents a substantial portion of the potentially available revenue base. These countries cannot afford to sacrifice a large proportion of their corporate tax bases, and the perpetuation of BEPS therefore poses a significant national hardship.  

The SE-ALS standard of income allocation is fraught with challenges. The standard ignores that MNEs are an amalgamation of branches and subsidiaries and sometimes arise specifically to take advantage of the profit potential in internalizing transactions within a group. With globalization, multinational corporations can now have their production in one state, supply of technical support and marketing in another state and their main center of management in yet another. Technological advancements and e-commerce further internationalize the business activities of MNEs thereby revealing the inadequacy of tax principles that are nearly a century old. Furthermore, the rise of MNEs can be attributed to the minimization of cross border transaction costs. SE-ALS is unable to provide taxpayers with a clear sense of how they are expected to behave in the legal system in which they operate.

The myth that members of a group of companies are separate entities from their parent company (and capable of independently dealing with the parent company or their subsidiaries) is problematic. It poses a challenge to properly determining an arm’s length price in a situation where companies transfer unique products with no existing market. This includes, for instance,

39 Baistrocchi, supra note 28.
40 Celestin, supra note 6.
intangibles like unique patents. Also, it can be time-consuming for tax authorities to resolve conflicts regarding whether the correct taxable income has been reported in a particular jurisdiction. This can lead to significant adjustments and reassessments against the relevant entity.

The arm’s length standard asserts that associated enterprises should always deal with each other on the same business terms as fully independent entities. To arrive at an arm’s length for the transfer of goods and services among related entities, methodologies have been recommended by supranational bodies such as the OECD and the UN in order to guide tax authorities in arriving at a price for the transfer of goods and services that would be deemed ‘arm’s length’. Three of these methods are aimed at the transaction or the price charged for the goods and services, the other two are aimed at the profit declared by the taxpayer for the overall transaction. The traditional transactional methods are the Comparable Uncontrolled Price Method (CUP), the Resale Price Margin (RPM) and the Cost Plus Method (CPM). The traditional profit methods are the Transactional Net Margin Method (TNMM) and the Profit Split Method (PSM). For all five methods, the taxpayer must include comparables in justifying the arm’s length standard, that is, that the price fixed or the profit declared is the same as would be achieved by unrelated parties dealing with each other under the same conditions and on the same terms. There is provision for adjustment by the tax authority where this may not completely be the case.

Having arrived at an arm’s length price, the taxpayer must convince the tax authority by preparing transfer pricing (TP) documentation. This is time and resource-intensive and demands expertise. Where the tax authority is not convinced, it may re-characterize the transaction and adjust the profits. This could lead to double taxation where the corresponding tax jurisdiction or authority does not do same. Also, in developing countries with limited capacity, taxpayers employ tax firms who have better knowledge and can effectively engage in tax planning. This could either shift profits to low tax jurisdiction or erode the bases of the host jurisdiction.

These methodologies are limited by the incapacity of tax authorities expected to properly implement them. Insufficient comparables of similar

---


The complexity of applying transfer pricing methodologies to the transfer of goods and services among related entities, and the exchange of goods and services among entities of the same MNE create opportunities to undercharge or overcharge the transfer prices, especially in developing countries. Where the transfer price does not reflect market value (due to an over or undercharge resulting in loss or profit on the part of the buyer or the seller) what arises is the opportunity to manipulate where the profit was realized.\footnote{See generally Peter Dietsch, \textit{Catching Capital: The Ethics of Tax Competition} (Oxford: Oxford University Press, 2015); L Eden, \textit{Taxes, Transfer Pricing, and the Multinational Enterprise}, in AM Rugman & TL Brewer, eds, \textit{The Oxford Handbook of International Business} (Oxford: Oxford University Press, 2001) at 591-619} This profit is often shifted to a low-or-no-tax jurisdiction, thus depriving the host or home state of revenue which would have otherwise accrued to them.\footnote{See generally Richard Murphy, \textit{The Joy of Tax: How a Fair Tax System Can Create a Better Society} (London: Bantam Press, 2015); Avi-Yona, \textit{supra} note 26.} Kale and Mahoney write that the SA-ALS approach “proved to be an administrative headache due to the accounting process of arm’s length pricing and consequently inconsistent with corporate interests”\footnote{Tracy Kaye & Michael Mahoney, “Various Approaches to Sourcing Multijurisdictional Values: Sourcing Options Available to Tax Policy Makers” (2009) State & Local Tax Lawyer: Symposium Edition 57 at 61.} These loopholes allow for base erosion and profit shifting and make the existing system indefensible.

\textbf{Rescuing a Failing System}

Attempts have been made at domestic and international levels to address the base erosion and profit shifting of income occurring in both residence and source countries through regimes such as controlled foreign corporations (CFC) rules, anti-avoidance principles, transfer pricing guidelines, penalties, exchange of information agreements, general anti-avoidance rules (GAAR) and stringent disclosure requirements.\footnote{Cooper et al, \textit{supra} note 42.}

At the domestic level, countries insert safe harbor provisions in their tax regime to limit the arm’s length standard of income allocation and the application of the transfer pricing methodologies. For instance, Regulation 15
of the *Transfer Pricing Regulations of Nigeria* 2012, provides that a connected taxable person may be exempted from the requirements of TP documentation where

a) the controlled transactions are priced in accordance with the requirement of Nigerian statutory provisions; or

b) the prices of connected transactions have been approved by other Government regulatory agencies or authorities established under Nigerian law and satisfactory to the Service to be at arm’s length.47

The implications of this provision are to provide reliance on statutory prices quoted in a market, and agreements reached by the taxpayer with other government agencies, such as the National Office for Technology Acquisition and Promotion (NOTAP), in the case of Nigeria.

Apart from promulgating safe harbor provisions, some jurisdictions have amended the existing transfer pricing methodologies to reflect economic realities or achieve expected outcomes. For instance, in Brazil, the CUP, RPM and CPM transfer pricing methodologies have been amended by introducing acceptable fixed margin and mark up ratios. In 2013, the Brazilian government further amended the transfer price methodologies to provide greater certainty to taxpayers. The new amendments cover sectors trading in pharmaceutical products, tobacco-related products (40% profit margin), chemical products, glass-related products, cellulose, paper or paper-related products and metallurgy (30% profit margin) and all other sectors (20% profit margin). For companies engaged in import and export of goods, predetermined gross profit mark-up may be used by the taxpayer (15% for exports and 20% for imports) in arriving at an acceptable price for the Brazilian tax authority.48

These adjustments set out to reduce the need for extensive TP documentation and provide price or profit benchmarks upon which parties

---


may agree, thus eliminating the search for comparables and guarantees certainty in tax revenue to the tax jurisdiction.

At the global level, the OECD’s Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations (the “OECD Guidelines”), which were first issued in 1979 and most recently revised in July 2017, provide guidance for application of the arm’s length principle. The guidelines detail a nine-step process for conducting an arm’s length analysis, which includes conducting a comparability study using one of five comparability methods (three so-called traditional transactions methods and two transactional profit methods) to determine what an arm’s length price would be. The steps culminate in a comparability adjustment, and a determination of an arm’s length price.49

Also, the OECD, in conjunction with G20 members50 initiated the Base Erosion and Profit Shifting (BEPS) Project51 to address the risk of transfer price abuse. A key aspect of the BEPS Action Plan is the country-by-country reporting (CBCR). The OECD hopes that, with CBCR, tax administrations where a company operates will get aggregate information annually, starting with 2016 accounts, relating to the global allocation of income and taxes paid, together with other indicators of the location of economic activity within the MNE group. It will also cover information on which entities do business in a particular jurisdiction and the business activities each entity engages in. It urges countries across the world to require all multi-national enterprises resident within their jurisdiction to file reports that detail how their global operations are structured. The project also incorporates rules that will oblige companies based in a non-compliant country to file the same

49OECD Transfer Pricing Guidelines, supra note 9.
50 The G20 members in their St. Petersburg Declaration (2013) called for changes to the international tax laws in order to ensure that profits are taxed where economic activities occur and value is created. The Declaration can be accessed through the G20 Research Group website: G20 Information Centre, G20 Leader’s Declaration, September 6, 2013 Summit in St. Petersburg, Russia (University of Toronto G20 Research Group website, 2017), online: <http://www.g20.utoronto.ca/2013/2013-0906-declaration.html>.
51 This is particularly harmful to developing countries where tax revenue as a percentage of GDP is around half of that in OECD countries. Governments across the world rely on five primary sources of tax revenue: personal income taxes, corporate taxes, sales and excise taxes, property and wealth taxes, and payroll taxes.
information elsewhere. The expectation is that tax authorities will thereby find it easier to recognize and tackle transfer pricing abuses.\textsuperscript{52}

Other attempts at fixing the broken system by the OECD, include the Multilateral Competent Authority Agreement (MCAA) and the Global Forum on Transparency and Exchange of Information for Tax Purposes (Global Forum). The MCAA, a multilateral framework agreement, provides a standardized and efficient mechanism to facilitate the automatic exchange of information in accordance with the Standard for Automatic Exchange of Financial Information in Tax Matters (the Standard\textsuperscript{53}), thus avoiding the need for several bilateral agreements. This agreement enables joint audit and automatic exchange of information, thereby promoting transparency and accessibility of information. The Global Forum on the other hand, already entered into by 135 jurisdictions, utilizes robust peer review that ensures that standards are adhered to on transparency and information sharing. The OECD seeks to level the playing field for developing countries as these countries can now get relevant information from all other signatories to the agreements.\textsuperscript{54}

However, these remedies introduced by the OECD and other stakeholders have failed to address the biting tax evasion and avoidance practices of multinational companies. From a developing country perspective, they are deemed non-inclusive and fail to address some of the pressing tax issues faced by developing countries, such as tax incentives. The current international tax system appears to have created a monster. It birthed a global tax war, and has led to direct tax competition among developing countries.

\textsuperscript{52} For further discussion on the BEPS project, see Organization for Economic Cooperation and Development (OECD), OECD Library (website), online: <http://www.oecd-ilibrary.org/taxation/oecd-g20-base-erosion-and-profit-shifting-project_23132612>.


THE RISING CLAMOUR FOR AN ALTERNATIVE

At the center of global tax discourse is the question of ‘the right to tax’. While the Global North advocates taxation based on the residence of the MNE, developing countries would rather base their revenues on the “source income” as this allows them to obtain tax revenues from the local operations of foreign companies. The OECD BEPS project shies away from addressing the elephant in the room (the inherent right to tax). Any meaningful progress in addressing the failing system of international taxation will necessitate a return to the negotiation table, as was done under the auspices of the League of Nations, to consider viable alternatives to the present system. Overhauling the global tax system and its practices is fundamental if we are to deliver stronger growth for a post-crisis world. This next section considers the alternative of the unitary taxation and formulary apportionment methodology of income allocation. First, we start by addressing some of the criticisms of the UT-FA approach in the past, and whether such challenges are present today and are still relevant in today’s course. Secondly, we analyze the UT-FA approach on its merits.

DEBUNKING EXISTING MYTHS

Below, we highlight some of the arguments frequently canvassed against the election or transition to a UT-FA approach of income allocation.

Sovereignty

The primordial politico-economic territorial structure and division of the world is under threat. Technological advancement, the expansion of e-commerce, and the birth of the internet economy, coupled with the gradual erosion of the physical-presence business concept, necessitate that we re-think the bonds that tie us together and the international laws by which we play the game of peaceful and meaningful co-existence. International trade, investment, and finance have become the hallmarks of economic globalization. As we inevitably transcend to a “stateless world” it becomes important that we negotiate the terms of the sovereignty states claim to inalienably possess.

The claim that a country’s unlimited right to tax income derived from it, or accruing to it, might no longer be true (if not in theory at least in practice). Physical markets are becoming obsolete, and virtual marketplaces (networks of computers and computer terminals), are emerging as the site for transactions. Enforcement of the tax liability of a corporation by a state may not be achievable without the cooperation of other states, thus giving rise to a conflict of interest between the host state and the home state. Equally, the tax policies pursued by one state can impact the economies of other state-termed “fiscal externalities”. A case in support of this point is the recent European Commission decision involving Apple and the Irish government. The European Commission ruled that Apple owed the Irish government €13 billion in tax non-remittance attributable to Ireland’s state aid program. The interesting perspective in this case is the potential threat to a country’s claim of fiscal sovereignty, power to attract FDI and control its market, though one may argue that Ireland gave up this right and power by joining the EU. Thus, we are faced with a global issue: the balance of the fiscal sovereignty of states, and the need to have a fair allocation of income where taxes are paid in jurisdictions where they are actually realized.

Also, the world has moved from an aggregation of different national economies to reliant and interdependent economies with integrated factors of production which produce income for all. Under the present global tax system, this integration and inter-dependence is not accounted for nor rewarded. Formulary apportionment allocates income to the states who contribute to income generation. The global consensus is that income should

---


58 For example, the Luxleaks case and the state aid cases in the EU reveal the limitations of economic determinism and fiscal sovereignty as we know them.


be taxed where the economic activities take place and where value is created.\textsuperscript{61} Furthermore, tax competition among countries only leads to a broken international tax system. This is favorable to MNEs who take advantage of the loopholes in the international tax system to shift profits to low (or zero) tax jurisdictions, and avoid paying their share of taxes in the jurisdictions where the real economic activities occur.\textsuperscript{62}

This implies that there is an urgent need to move beyond the narrow definition of sovereignty and the separate entity approach of MNEs to a pragmatic approach embracing the integrated and reliant nature of MNEs. In line with globalization and the current international nature of business, taxation must shift from national/domestic control to international control. Roger Wesley puts it thus:

As the world slowly emerges from its worst recession in thirty years, it has become painfully evident that no nation-state rich or strong is insulated from the actions of its neighbors. Ours is a global economy. “Recessions, inflation, trade relations, monetary stability, gluts and scarcities of products and materials...are international phenomena” affecting all national participants. The realities of economic life push toward global interdependence, discrediting in the process outgrown concepts of economic determinism.\textsuperscript{63}

Mintz argues that coordination of tax policies can ameliorate the impact of fiscal externalities and improve economic performance in coordinating countries, and that coordination facilitates the free flow of business inputs across national boundaries, minimizes fiscal externalities, and reduces the cost of compliance and administration.\textsuperscript{64} I tend to agree.\textsuperscript{65}

---


\textsuperscript{63} Roger Wesley, “Problems in Regulating the Multinational Enterprise- An Overview” (1976) 10:4 The Intl Lawyer at 613-22.


\textsuperscript{65} Important to note that the global impacts of recessions are attributable to the presence of MNEs abroad and the increase in FDI by firms.
Information Exchange and the Finance Sector

An offshoot of the exercise of the sovereignty of a state is the state’s exercise of sovereign power to exchange information with other states as it wills as well as the unfettered control of its finance sector. The events of the last decade prove that these powers are constantly slipping from the state’s grip. In 2010, the US enacted the Foreign Account Tax Compliance Act (FATCA), requiring financial institutions around the world to disclose to the IRS large accounts of US persons, or pay a 30% withholding tax on the institution’s US earnings.66 Countries, following FATCA, are setting up similar bank account disclosure systems with global reach. Central to the OECD’s BEPS project is the access to information. The OECD, with the backing of the G-20 leaders, has initiated and successfully procured countries to sign MCAA for the automatic exchange of Country-by-Country reports (CBCR).67 The OECD hopes that:

...with country-by-country reporting, tax administrations where a company operates will get aggregate information annually, starting with 2016 accounts, relating to the global allocation of income and taxes paid, together with other indicators of the location of economic activity within the MNE group. It will also cover information about which entities do business in a particular jurisdiction and the business activities each entity engages in. The information will be collected by the country of residence of the MNE group, and will then be exchanged through exchange of information supported by such agreements as signed today.68

Based on information gathered in 2016, it is expected by the OECD that the first exchanges will start in 2017-2018. Global expectation is that the Country-by-Country reports will result in the exchange of information between tax authorities relating to financial accounts in each jurisdiction operated by MNEs and residents of the other jurisdictions. Account information to be reported on includes account balances, interest, dividends, and sale and redemption proceeds from financial assets. The financial institutions here include deposit-taking banks, custodial institutions, investment entities and insurance companies. For developing countries struggling with tax evasion and illicit capital flight, the OECD’s Country-by-Country report model is a viable tool for addressing tax evasion considering that these developing countries will be armed with relevant information in the fight against tax evasion and avoidance. The OECD, in conjunction with the Global Forum, is working to enhance the capacity of developing countries to exchange information with other jurisdictions so that they can take advantage of these tools. The reciprocal nature of the Country-by-Country reports facilitates this endeavor.

The EU has accepted the recommendations to introduce: country-by-country reporting of multinationals’ activities; common consolidated tax base (CCTB); better protection of whistleblowers; extension of automatic exchange of information on tax rulings to all tax rulings and available to the public; countermeasures towards companies that make use of tax havens; changes to the EU state aid regime as it relates to tax through binding guidelines; etc.69

The US FATCA, OECD’s BEPS project and the MCAA and the Country-by-Country reports models, EU’s CCCTB and general proposals, underscore the commitment of stakeholders to address the ills of tax evasion and tax avoidance, and the importance of information in the fight. To ensure the accessibility of information, bank and tax secrecy would necessarily have to give way. A country’s power over this information will necessarily be lessened or redefined in today’s globalized world.

Implementation Hurdles

Admittedly, moving from the SA-ALS approach to the UT-FA approach of income allocation presents massive implementation issues, justifying the present fears, skepticism and reluctance amongst stakeholders. One such hurdle is the accounting standard to be adopted. Conventional accounting standards are not suitable for tax purposes as taxable profit or chargeable income differs from accounting profits of a corporation. Secondly, the accounting world is split between choice of IFRS (used mostly in the EU), and the GAAP (the system used in the US). From an accounting perspective, these are two significant challenges that will hinder the development of a global tax system. How do we achieve convergence between accounting standards and tax standards, and how do we achieve convergence between the IFRS or the GAAP? Do we pick one over the other? How do we reconcile the fundamental differences between the US GAAP and IFRS? How do we define the tax base? Or is mutual recognition the passport to having agreeable tax accounting standards? Can one standard be deemed equivalent to the other? Or is compliance with one standard a sufficient substitute for the other? How do we ensure countries comply with any chosen standard? Will this be a multilateral or bilateral treaty agreement? Are there lessons to learn from the EU, with the adoption of the IAS/IFRS Regulation (EC) 1606/2002? Also, the verification of a group’s financial accounts raises technical and administrative challenges which require time and effort to resolve.

Prem Sikka and Richard Murphy argue that the relationship between tax and financial reporting is now remote, and that no jurisdiction we can identify relies upon unadjusted traditional accounting profit as a basis for the taxation of corporate income. [They] recognize the divergence between the bases on which expenses, revenue and profits are recognized in corporate financial statements and those allowed for tax purposes. They further opine that “since consolidated group accounting is now the normal arrangement for most multinational group entities, it might seem that their accounts prepared on this method could form the basis for tax assessment under a unitary taxation system.” I agree, however, the tasks of

---

70 See Avi-Yonah, supra note 24 at 26.
71 GAAPs of US, Japan, China, Canada, South Korea and India are found to be equivalent to International Financial Reporting Standards (IFRS) as adopted by the EU.
72 The Indian tax authority rejects the use of the IFRS system for tax purposes.
73 Sikka & Murphy, supra note 24 at 3.
74 Ibid at 9.
converting accounting profit to tax profit and verification of that conversion poses the daunting task. Sikka and Murphy further recommend starting with the worldwide consolidated financial accounts of the corporate group and adjusting them to the appropriate tax accounting standards, and providing a trail from financial accounts representing actual economic activity undertaken to tax accounts.\(^{75}\)

It must be said that cooperation among countries, including cooperation in achieving greater convergence between book and tax accounting in order to ease the problems in the translation of accounts, is a *sine qua non* to the adoption of the UT-FA approach.\(^{76}\) Cooperation is also needed in the areas of enforcement and examination of tax standards.\(^{77}\)

Also, defining the tax base presents opportunities for tax avoidance. Siu et al. caution that “for any instrument assessed on a base that incorporates expenses (such as overheads, transportation or intra-firm services), or capital costs (such as depreciation, inventory cost or risk), the scope of aggregation will have a significant effect on determination of the taxable base.”\(^{78}\) Canada uses a harmonized tax base where similar allocation formulae is used by all provinces; it is federally-defined although provinces are left to determine the tax rates and reliefs to levy and grant on the apportioned income. The proposed EU’s CCCTB approach provides for a high degree of tax base harmonization, including the definitions of profit, loss, revenue, expenses, deductible items and depreciation framework. Tax base under the EU’s CCCTB is defined as revenue less any exempt revenue, deductible expenses and other deductible items.\(^{79}\)

Language and currency impediments are surmountable in today’s 21st century world and have been addressed by different trade and international agreements and practices. The relative success of the OECD’s BEPS Project, and the political strength shown by the OECD in orchestrating the passage and ratification of the *Multilateral Convention to Implement Tax Treaty Related

\(^{75}\) *Ibid.*
\(^{76}\) The EU’s proposed CCCTB offers a conceptual taxation framework, which can be improved upon for global or regional adoption.
\(^{77}\) Durst, *supra* note 36.
Measures to Prevent BEPS (MLI) by over 70 countries, makes the claim of transitional cost or difficulty fall on its face.

**Determination of the Factors and the Formula**

No literature has succeeded in proposing a formula that agreeably captures the location of economic activity; however, countries have experimented with different factors, formulae and combinations. Canada allocates corporate income through a formula of two factors: sales by destination and payroll.\(^8^0\) The Massachusetts equally-weighted formula of three factors – property, payroll and sales – was the prevailing formula in the U.S. but has only been implemented by twelve US states.\(^8^1\) Alaska uses an extraction factor in addition to the factors of property and sales, in the oil, gas and pipeline sector when a corporation is involved in all three businesses and reverts to the two-factor approach of property and sales by destination. For mineral revenue dependent source states, an origin-based sales factor is believed to capture returns from extraction.\(^8^2\)

Payroll captures the wages and salaries paid to employees. The argument for a payroll factor is that it recognizes the source state as an integral aspect of income generation. It also offers administrative advantages. Arguments against payroll are the treatment of fringe benefits and payment to independent contractors. Should payroll be on the cost or turnover of employees or should it be on the number of employees? Apportionment by cost or turnover of employees may favour developed countries, given the high wage rates and per capita income, while apportionment by number of employees would have a strong redistributive effect, favoring developing countries. The EU CCCTB proposes an equally weighted turnover of employees and number of employees in the payroll factor.\(^8^3\)


\(^8^1\) The single-sales by destination formula is now being implemented in sixteen U.S. states. See generally Siu et al, “Unitary Taxation”, supra note 25.

\(^8^2\) Ibid.

\(^8^3\) See generally Cobham, supra note 4; see also Sol Picciotto, “Towards Unitary Taxation of Transnational Corporations” (2012) Tax Justice Network, online: <https://www.taxjustice.net/cms/upload/pdf/Towards_Unitary_Taxation_1-1.pdf> [Picciotto, “Towards Unitary”].
In the proposed EU CCCTB the asset factor will consist of all fixed tangible assets. Intangibles and financial assets will be excluded from the formula due to their mobile nature and the risk of circumventing the system. In the US, the property factor considers only tangible property, and excludes intangible property. The value of property is based on historical cost rather than its actual value, with no adjustment for either depreciation or inflation. This use of historical cost provides a poor approximation of both the value of the asset and the user cost of capital. Ignoring intangible assets is one of the failings of the US formulary apportionment model. Recognizing it presents transfer pricing problems similar to those currently faced.

The proposed EU CCCTB addresses the sales factor by including the proceeds of sales and of any other transaction, net of value added tax and other taxes and duties collected on behalf of government agencies.\footnote{CCCTB, supra note 79, art 4.} Sales factor in the US, in addition to capturing the sale of goods, also captures receipts from the provision of services, rentals, and royalties. Sale of tangible property is generally attributed to the state of destination, while intangible properties are commonly attributed to the state where income-producing activities are performed, or the state of the market for such services or intangibles (the destination). Some US states deploy a “throwback” provision. These provisions enable the ‘state of origin’ to tax sales made to the federal government, or sales made to a state who lacks jurisdiction to tax the vendor’s income. Canada applies a destination-based rule that attributes gross revenue to the permanent establishment where the customer is located. The average of the share of income deemed earned in each province is the average of the share of gross revenue attributed to the permanent establishment in the province, and salaries and wages paid in the year by the corporation to employees of the permanent establishment. Canada also applies the throwback rule.\footnote{Stefan Mayer, \textit{Formulary Apportionment for the Internal Market} 17 IBFD Doctoral Series, Vol. 17 (IBFD Academic Council, 2009) at 61, online: <https://www.ibfd.org/IBFD-Products/Formulary-Apportionment-Internal-Market-Review>.

Determining the appropriate formula and factors presents the biggest challenge to adopting the UT-FA approach. Avi-Yonah and Clausing propose a sales-based formula to be determined on the basis of the location of the customer rather than the location of production.\footnote{Kimberly Clausing & Reuven Avi-Yonah, “Reforming Corporate Taxation in a Global Economy: A Proposal to Adopt Formulary Apportionment” (2008) The Brookings Institution (website), online: <https://www.brookings.edu/research/reforming-corporate-taxation-in-a-}}
a sales-based only formula is that it caters to the demand side of the market without providing for the supply side. This formula is inappropriate for a global adoption of the UT-FA approach to income allocation as many developing countries are suppliers, rather than purchasers, of goods and services. This exclusive formula deprives them of needed and earned revenue. Also, determining the destination of sales and what constitutes sales may prove difficult for tax authorities, and even more so with e-commerce and technological advancement. For example, is the transfer of software a sale or a license? Who is the customer for the purpose of tax jurisdiction? Do intermediaries exist?

Experts have proposed replacing the property factor with the equity and quasi-equity factor.87 They claim that this formula would allocate business income among the jurisdictions that hosted production/generation of such income (including the jurisdictions whose residents contributed equity and quasi-equity that enabled the business to operate and to the jurisdictions whose residents purchase the goods or receive the services. As a commercial lawyer, it is my opinion that incorporating the financial market and the equity/quasi-equity factors in the current mix will exacerbate the current situation as both factors are easily open to manipulations. With bank secrecy still in place, it becomes easy to shift the origin of financing deals to tax havens and protective jurisdictions, notwithstanding that the real financing is in a different jurisdiction. Also, equity/quasi-equity capital may be routed through offshore tax havens as these countries provide ease of incorporation and bank secrecy.88

Another consideration is whether to apply formulary apportionment to worldwide income or a “water’s edge” limitation. Walter Hellerstein writes that the unitary taxation of corporate income by some US states was the worldwide combination of “the consolidation of the activities of related entities deemed to be engaged in a unitary business, no matter where such activities occurred and no matter where the entities or their parents were

---


88 Mauritius, British Virgin Islands and the Netherlands are some jurisdictions I have come across in practice as a commercial lawyer used for round-tripping with protective measures.
resident.” He further opines that states abandoned worldwide combination and retreated to a “water’s edge” definition of a combined (or consolidated) group limited to US domestic corporations. Certain tax haven corporations, and foreign corporations that exceed a threshold of business activity in the US. Canada adopts a “water’s edge” limitation. The proposed EU CCCTB would apply only within the EU, using a strict water’s edge approach.

Developing suitable factors and formulae demands great political will and cooperation, since the factors and formulae determine the allocation of income to any state. Cobham and Loretz warn that “the specific apportionment formula chosen is likely to have substantial redistributive consequences given that proposals for unitary taxation originate largely from the observation that profits are misaligned under the current system...” This is where the real work lies. Unlike a century ago, relevant research is ongoing to bridge this knowledge gap, and as such, a strong case for shifting to the UT-FA approach exists. The EU’s CCCTB, which addresses most of these issues, presents useful guidance for the rest of the world.

**Unitary Taxation and Formulary Apportionment Approach on its Merit**

In the assessment of the appropriate international tax policy, some criteria are widely accepted. These criteria are: internation equity, inter-taxpayer equity, neutrality, simplicity, and administrative efficiency. It is on the basis of these criteria that we assess the suitability or otherwise of the UT-FA approach to income allocation.

Advocates for the adoption of the UT-FA approach believe that the incentive to shift income to low tax jurisdiction will disappear. Sol Picciotto argues that:

> A unitary approach would replace three major elements which create fundamental problems for taxation of TNCs under the ALP: (i) the need for detailed scrutiny of internal accounts and pricing and for the negotiation of adjustments based on the ALP; (ii) the need to deal with profit-shifting within the firm, especially using tax

---

90 Ibid.
91 Cobham, supra note 4 at 7.
92 For a richer discussion on these criteria, see Jinyan Li, “Global Profit Split: An Evolutionary Approach to International Income Allocation” (2002) 50:3 Can Tax J 823.
93 For a broader discussion, see Siu et al, “Unitary Taxation”, supra note 25.
havens, by complex anti-avoidance measures, such as rules against thin capitalization, controlled foreign corporations, and abuse of treaty benefits; and (iii) source and residence attribution rules.  

Michael Durst opines that, “the adoption by a country of a formulary approach to income apportionment would appear to offer a more reliable means of curtailing base erosion, particularly over the long term, than attempting to apply a mixture of politically vulnerable, and often only partially effective, anti-avoidance measures.” This does not necessarily mean the end of tax havens, however, and tax havens would have to be engaged in real economic activities to be allotted income. Alex Cobham and Loretz, argue that a shift to unitary taxation will shift the corporate tax base away from countries with ‘favorable’ tax regimes, for example, Luxembourg, the Netherlands and Ireland, to countries where real economic activities occur. This ensures inter-taxpayer equity, as companies are now forced to pay taxes where the economic activities occur, ending the shift of the tax burden to individual taxpayers.

The requirement for MNEs to prepare a combined report covering the whole of the corporate group engaged in a unitary business ensures administrative efficiency and simplicity of applying the tax system for tax authorities, given that the information becomes available to administer effectively. Consolidated or combined reporting is essential to an effective UT-FA approach. Derek Devgun writes that, “under worldwide combine reporting, local taxable income is recalculated by multiplying the worldwide combined income of related or controlled parties (i.e. the unitary business) by a fraction based on the proportion of payroll, property and sales within the

95 Ibid at 10.
96 Durst, supra note 36.
97 Kerrie Sadiq, “Unitary Taxation of the Finance Sector” (2014) International Centre for Tax and Development Working Paper 25. Sadiq further proposes that unitary taxation based on formulary apportionment be implemented in the financial sector and recommends an equally weighted two-factor formula of labour and sales, where labour reflects both remuneration and number of staff, and the formula should be applied to all the income of a multinational financial institution (MNFI) on a combined income basis. In his paper, he specifically argues that the nature of the business of MNFI, the integrated nature and inter-connectedness demands the UT-FA approach, as a contrary approach will be fertile ground for transfer pricing abuses.
98 Cobham, supra note 4.
jurisdiction imposing the tax".\textsuperscript{100} The proposed EU CCCTB condemns the current system as “a cumbersome process, both timing-wise and economically, which diverts the effort out of the main thrust of doing business.”\textsuperscript{101} The EU equally seeks to reduce firms’ compliance cost, and to achieve simplicity and neutrality through harmonization of the tax base, a consolidated account and a pre-determined formula.\textsuperscript{102} Siu et al., in their study of the application of UT-FA in the extractive industry (EI), recommend a shift from the current system to the UT-FA approach.\textsuperscript{103} They claim this would assist governments not only to improve general CIT design but also to develop better rent/profit-related levies. Also, the use of UT based on a common global corporate group tax base for TNCs in the EI sector could reduce administration and compliance costs associated with both CIT and rent/profit-related levies. They surmised that a UT approach, which requires country-by-country reporting on revenue, costs and tax payments, and the apportionment of global revenue based on that information, could assist in the development of better EI levy policies that are sensitive to the risks and costs incurred by private companies when extracting resources.\textsuperscript{104}

Inter-nation equity is achieved through global redistribution of income. This ensures that countries earn a fair share of return on their factors of production and do not suffer from base erosion and profit shifting that currently occur where profit declaration is disproportionate to real economic activities taking place in some jurisdictions, and willful tax competition aimed to attract return but do not necessarily result in substantial investment in their country. The LuxLeaks scandal provides a suitable example. Runkel and Schjelderup argue that the switch from Separate Accounting to Formulary Apportionment substantially increases average tax revenue and welfare of the

\begin{footnotesize}
\begin{enumerate}
\item The proposed EU CCCTB condemns the current system as “a cumbersome process, both timing-wise and economically, which diverts the effort out of the main thrust of doing business.”
\item Siu et al., in their study of the application of UT-FA in the extractive industry (EI), recommend a shift from the current system to the UT-FA approach.
\item They claim this would assist governments not only to improve general CIT design but also to develop better rent/profit-related levies.
\item Inter-nation equity is achieved through global redistribution of income. This ensures that countries earn a fair share of return on their factors of production and do not suffer from base erosion and profit shifting that currently occur where profit declaration is disproportionate to real economic activities taking place in some jurisdictions, and willful tax competition aimed to attract return but do not necessarily result in substantial investment in their country. The LuxLeaks scandal provides a suitable example. Runkel and Schjelderup argue that the switch from Separate Accounting to Formulary Apportionment substantially increases average tax revenue and welfare of the
\end{enumerate}
\end{footnotesize}
member states.\textsuperscript{105} It has been argued that a switch to the UT-FA approach would subject MNEs to a two-digit tax rate (in comparison with the single digit that multinational corporations are currently effectively taxed) and would allocate income among states more justly.

Avi-Yonah and Clausing propose a system of formulary apportionment for taxing the corporate income of multinational firms. They hold the same view that in an increasingly global world economy it is difficult to assign profits to individual countries, and attempts to do so create opportunities for tax avoidance.\textsuperscript{106} Avi-Yonah and Clausing argue that under their proposed formulary apportionment system, firms would no longer have an artificial tax incentive to shift income to low-tax locations. The complexity and administrative burden of the current system would also be reduced, and the proposed system would be better suited to an integrated world economy and more compatible with the tax policy goals of efficiency, equity and simplicity.\textsuperscript{107} They argue that the present US system provides an artificial tax incentive to earn income in low-tax jurisdictions, rewards aggressive tax planning, and is not compatible with any common metrics of efficiency.\textsuperscript{108} I conclude that a shift to the UT-FA approach of income allocation would achieve inter-nation equity.

\textbf{RECOMMENDATIONS AND CONCLUSION}

Admittedly, the transition to the UT-FA approach has a high cost. The inhibitory factors are still present; however, the biggest hurdle is that of the political will on the part of supranational bodies like the OECD to accept its strength and relevance in today’s global economy. Achieving global consensus on the factors and formula to be applied presents a daunting task.

While this is a subject that calls for more research and focus, I recommend a system that incorporates the sales factor, asset and payroll. The sales factor will be equally weighted between an origin-based sales factor and a destination-based sales factor, with provisions for throwbacks. The origin-based sales factor captures source states and the destination-based sales factor accommodates the market for the goods. Throwbacks should be extended to states that refuse to tax in spite of having jurisdiction, thus eliminating any

\textsuperscript{106} Avi-Yonah, supra note 26
\textsuperscript{107} Ibid.
\textsuperscript{108} Ibid.
tax sparing provisions. I admit that special formulae may have to be created for specific industries and goods such as intangibles and e-commerce. Just as the EU, payroll should be equally weighted between number of employees and turnover of employees. Assets present a more daunting task and demand further reflection.

Seeking a one-size-fits-all UT-FA approach is impractical and will hinder any real actualization of the UT-FA approach of income allocation. Some industries, goods and services are distinct from others and must be accorded such distinct treatment. The valuation of intangibles such as intellectual property rights (IPRs) cannot be treated in the same way as tangible products such as shoes.

Finally, while there are many uncertainties in the design and implementation of the UT-FA approach, this paper argues that the setbacks and challenges, hitherto claimed, are being remedied and may not pose as great of a challenge as they did in the past. It further argues that given the failing system of income allocation that currently exists, a change to a viable alternative, as offered by the UT-FA approach, is worthwhile and appropriate in today’s further globalizing and integrating economies.
INTRODUCTION

It is not surprising that academic and legislative debates on corporate crime in the United Kingdom (UK) and United States (US) are still dealing with fundamental issues on whether and how corporations should be subject to criminal law. In January 2017, the UK issued a call for evidence to change the standard of liability for ‘economic crimes’, highlighting failure of the common law identification doctrine to achieve its twin goals of deterrence and retribution. Similarly, there have been continuous debates in the US.
over whether there should be changes to federal criminal code to include more specific language about criminal intent.\(^2\) 

Today, describing the scope of corporate crime, whether it even fulfils the criteria of a ‘crime’, and the behaviors it should encompass, is difficult. The topic remains highly controversial in today’s legislative debates throughout various jurisdictions. Corporate criminal liability can be defined as the mechanism through which corporations can be found to be responsible under the law for the criminal acts of employees that are undertaken during the normal course of business operations to further the goals of that entity.\(^3\) There remains a divergence of views with regards to (a) the theoretical justifications for holding corporations liable under criminal law; (b) the circumstances in which the criminal acts of certain employees should be implicated to the corporation; (c) whether it is theoretically sensible and practical to hold corporations liable for crimes other than strict liability offences that require intent, recklessness and negligence; and (d) the appropriate mechanisms for sentencing.\(^4\)

This article argues that uncertainty surrounding these key inquiries have stunted the development of corporate criminal law. It is apparent that modern criminal law has been insufficient in achieving the goals of deterrence and retribution, a failure that can be traced back to two fundamental reasons. First, the first cases recognizing corporate criminal law have not coherently assessed the objectives of criminalizing corporations and addressed how such objectives can be achieved.\(^5\) Second, the globalization of businesses presents a significant challenge to regulating corporate misconduct across jurisdictions. As a result, the current standards of liability are incompatible with achieving the general goals of sentencing. Focusing on inquiries (b) and (c) from the paragraph above, the paper proposes modifications to the standard of liability for holding corporations liable under criminal law for criminal conduct requiring proof of mens rea, drawing


\(^5\) Ibid.
from insights taken from Canadian law. The Canadian standard of liability provides a balanced approach that overcomes some of the flaws found in the UK and US models.

Section 1 discusses the US federal standard of respondeat superior. Section 2 evaluates the identification doctrine under UK law and the Model Penal Code adopted by many States in the US, and critiques the recent movement in the UK towards piecemeal reforms focused on adopting different standards of liability for different misconducts. Section 3 proposes reforms to the current liability standards by exploring Canada’s approach to liability, which adopts an identification model that takes into accounts the differences between corporate structures and types of corporate offenders.

1. THE RESPONDEAT SUPERIOR STANDARD

Corporations are held to the respondeat superior standard under US federal law for crimes requiring a culpable mind. The US Constitution provides that States have the right to set their own criminal laws, except for any offences that have a federal interest. In this sense, the federal code overlays the codes of all States and the District of Columbia. Federal criminal laws apply to corporations regardless of whether those laws explicitly mention corporations or not. The United States Code states that “unless the context indicates otherwise... the words ‘person’ and ‘whoever’ include corporations, companies, associations, firms, partnerships, societies, and joint stock companies as well as individuals.”

New York Central & Hudson River Railroad Company v United States7 (New York Central) established the basis for corporate criminal liability in the United States and contains the current standard of corporate liability under US federal law.8 It specified that, for the purposes of public policy, the tort principle of respondeat superior should extend to criminal law also. Respondeat superior holds principals liable for the wrongful actions of their agents if the actions were completed for the benefit of the principal and were within the express or apparent scope of their duties. The Court explained that if officers

---

6 Words Denoting Number, Gender and So Forth, USC tit 1 §1 (2017).
7 New York Central R Co v United States 212 U.S. 481 (New York Central 1909) [New York Central].
and agents of corporations make decisions in accordance with the corporations’ purposes, motivations, and goals, then the corporation itself can be seen to have intended the moral and/or immoral acts, and should be correspondingly praised and/or held accountable for such acts.\(^9\) In *New York Central*, a railroad company and two of its agents (the general traffic manager and assistant traffic manager) were held liable to pay rebates to sugar refining companies on shipments of goods between states, as they were in violation of the *Elkins Act*.\(^10\) The court considered the constitutional validity of the *Elkins Act*, which stated:

(a) Anything done or omitted to be done by a corporation common carrier subject to the act to regulate commerce...which, if done or omitted to be done by any director or officer thereof, or any receiver, trustee lessee, agent, or person acting for or employed by such corporation, would constitute a misdemeanour under said acts, or under this act, shall also be held to be a misdemeanour committed by such corporation; and

(b) In construing and enforcing the provisions of this section, the act, omission, or failure of any officer, agent, or other person acting for or employed by any common carrier, acting within the scope of his employment, shall, in every case, be also deemed to be the act, omission, or failure of such carrier, as well as that of the person.\(^11\)

The Court held that corporations should be held criminally liable for the acts of their agents that are in violation of the *Elkins Act*. In particular, agents are assumed to have had the power to perform acts within the scope of their employment for the benefit of the principal without the principal having necessarily participated in or expressly authorized that conduct.\(^12\) Because corporations accept the benefits of the agents’ conduct through profits derived, it was held to not be contrary to public policy to hold the corporation accountable for unauthorized acts by its agents if they had accepted the benefits of such acts. The Court further explained that since the corporation itself cannot be arrested or imprisoned, its property could be taken for the purposes of punishment or compensation.\(^13\)

Today, the *respondeat superior* doctrine, established through *New York Central*,\(^14\) and the corresponding aggregation theory, reflect the current

---

9 *New York Central*, supra note 7 at 495.
11 Ibid.
12 *New York Central*, supra note 7 at 495.
13 Ibid at 493.
14 Ibid at 481.
standard of liability for crimes requiring proof of mens rea. A corporation can be prosecuted for intent, recklessness or knowledge if an employee commits a crime, as “the only way in which a corporation could act is through the individuals who act on its behalf.”

Under the respondeat superior doctrine, the criminal act of any employee acting within the scope of his or her employment and on behalf of the corporation, regardless of status within the corporate structure or his or her capability to represent the corporation, can be implicated to the corporation. Further, the aggregation theory adds to this and provides that “corporations could be held criminally liable on the act of one employee and on the culpability of one or more other employees who cumulatively, but not individually, met the requirements of actus reus and mens rea of the crime”.

CHALLENGING THE RESPONDEAT SUPERIOR STANDARD

Looking at the early cases establishing corporate criminal liability in the United States reveals fundamental issues with adopting the respondeat superior standard. New York Central failed to appropriately address the theoretical coherence of adopting the standard of liability, as the standard of liability was too broad and does not account for the current variety of corporate governance structures and types of corporations.

First, New York Central did not provide a reason beyond ‘public policy’ for extending respondeat superior (a tort principle) to criminal law. It cited The Queen v The Great North of England Railway Co, an English case which held that a corporation was criminally liable for illegally cutting through and obstructing a highway. It specified that for the purposes of public policy, respondeat superior should extend to criminal law. This evidences that the standard of liability arose not necessarily as a reasoned policy choice but rather due to “shifting trends in legal formalisms” and that the growing size and power of corporations had enabled them to cause harm to society in

---

15 Ibid.
18 The Queen v Great North of England Railway 9 Q B 315, 319 (1846).
19 Ibid.
pursuit of their economic goals without being held accountable. Regulating corporations using criminal law was justified by its expediency and effectiveness in responding to these problems that could not be solved through the private remedies that existed.

Hasnas states, “New York Central was a mistake when it was decided, remains a mistake today, and should be explicitly overruled.” Hasnas rejects the extension of the tort respondeat superior theory to criminal law because of its distinct purposes. Criminal liability is imposed for the purpose of punishment, whereas the purpose of imposing tort law is to achieve ‘corrective justice’, which would require an individual even without personal fault to pay compensation to restore an injured party for an act they caused or benefited from. Using public policy as a reason to extend this principle to criminal law authorizes a form of collective punishment that directly contradicts the fundamental principles of a liberal society historically advocated by Anglo-American criminal law. Additionally, Hasnas argues that corporations cannot possess a particular state of mind when committing an act; hence ‘mens rea’ cannot be satisfied by a corporation.

One central premise of this article centers around the failure of early cases to explain the adoption of a certain standard of liability and how such a standard would fulfill the objectives of the law, often defined as retribution and deterrence under current sentencing policies. Although this article does not deal with theoretical arguments as to whether corporations should be held morally responsible, it is generally believed that eliminating ‘morality’ out of the discussion over whether criminal law should be applied would open the doors to justifying corporate criminal liability on many grounds. By focusing on this traditional standard of liability, vicarious liability in practice fails to take into account the differences between corporations. In sum, corporations should not be held responsible for unauthorized decisions made by employees.

---

21 Thomas, supra note 20; Baer, supra note 8.
22 Hasnas, supra note 17 at 1329.
23 Ibid at 1349.
24 Ibid at 1338.
25 Sentencing Council 1, supra note 1; Sentencing Council 2, supra note 1; Desio, supra note 1.
26 Baer, supra note 8.
Respondeat superior implies that the acts of low level employees in an international firm can be implicated to the corporation, even if committed without the authorization of the firm. Such a standard does not necessarily align with the goal of deterring corporations from committing corporate wrongs if they can still be held liable for criminal acts committed by employees who do not follow otherwise compliant corporate policies. Although prosecution does take into account whether the specific employee made efforts to hide his unauthorized conduct, the corporation will often still be held liable. The corporation would still be required to “enter a deferred or non-prosecution agreement, whereby the company agrees to pay extensive fines, cease certain activities, agree to an outsider monitor, and engage in additional reforms set out by the relevant prosecutor.” However, the Sentencing Commission Guidelines state that the fines can be reduced by up to 95% if the corporation can show evidence of an effective compliance sentence, but that would ultimately be contingent upon reporting to authorities and that the conduct was not committed by high level personnel.

2. The Identification Doctrine

The identification doctrine reflects the corporate criminal liability standard in UK common law and in several States in the US who have adopted the Model Penal Code.

First, UK law holds that the type of crime determines the parameters of corporate criminal liability. Corporations can be held liable for any crimes committed by individuals under the separate legal personality principle. The ‘alter ego principle’, otherwise known as the ‘directing mind’ or identification doctrine, applies to criminal offences requiring mens rea. In Tesco v Nattrass, one of the leading cases explaining this standard of liability, Lord Diplock stated:

---

27 Desio, supra note 1.
28 Baer, supra note 8.
29 Desio, supra note 1.
32 Wells, supra note 30 at 42.
What natural persons are to be treated in law as being the company for the purpose of acts done in the course of its business, including the taking of precautions and the exercise or due diligence to avoid the commission of a criminal offence, is to be found by identifying those natural persons who by the memorandum and articles of association or as a result of action taken by the directors, of by the company in general meeting pursuant to the articles, are entrusted with the exercise of the powers of the company.34

Under the doctrine, only agents in high positions capable of making decisions on behalf of the company or who play a key function within the corporation’s decisional structure who, acting within the scope of their employment, can replicate the corporation’s intention are held to represent the corporation (even if they acted against corporate policy). This is because they are considered to have “[directed] the mind of the company towards commission of an offence.”35 Therefore, a corporation would not be liable unless these identified agents are evidenced to have had the requisite actus reus and mens rea for the relevant offence.36

Furthermore, the UK Parliament has passed a few acts that set different standards of liability for different types of corporate misconduct. Notably, the Corporate Manslaughter and Corporate Homicide Act, 2007 abolished the common law manslaughter offence that followed the alter ego principle.37 Under the Act, an organization can be found criminally liable “if the way its activities are managed or organized- (a) causes a person’s death, and (b) amounts to a gross breach of a relevant duty of care owed by the organization to the deceased.”38 The corporation would only be found liable if the activities were managed or organized by one or more members of ‘senior management’, which are defined as those who have:

significant roles in (i) the making of decisions about how the whole or substantial part of its activities are to be manage or organized or (ii) the actual managing or organizing of the whole or a substantial part of those roles.39

Therefore, the Act takes a broader approach than the common law approach because it does not require identification of the culpable act or mind of a specific individual within the corporation, but instead puts the onus on the prosecution to prove that the criminal act was associated with the

34 Ibid at 200.
35 Wells, supra note 30 at 42.
36 Ibid at 43.
37 Corporate Manslaughter and Corporate Homicide Act 2007 (UK) c 19, s 20.
38 Ibid, s 1.
39 Ibid, s 1(4).
way senior management conducted its affairs before liability can be established.

Moreover, under the Bribery Act, 2010 an organization can be held liable if a person associated with the organization bribed a person with the intention of “[obtaining or retaining] business for the organization, or [obtaining or retaining] an advantage in the conduct of business for the organisation”. 40 Unless the organization proves they had “adequate procedures designed to prevent persons associated with the organisation from undertaking such conduct,” 41 a person is considered to be associated with the organization if they perform services for or on behalf of the organization, regardless of their capacity (including employees, agents or subsidiary). Additionally, employees are assumed to perform services for, or on behalf of, the organization unless it can be proven otherwise. 42

Second, the American Law Institute’s Model Penal Code (MPC) codified substantive criminal laws and is often relied upon by courts to interpret and apply the law. 43 A number of US States have adopted some or revised provisions of the MPC. Robinson and Dubber state: “if there can be said to be an “American criminal code,” the Model Penal Code is it...even within the minority of states without a modern code, the Model Penal Code has great influence, as courts regularly rely upon it to fashion the law that the state’s criminal code fails to provide.” 44

The MPC sets the circumstances in which corporations can be held criminally liable. This evidences an approach that is more restrictive, and is more similar to the English identification model approach, with limited exceptions. 45 Section 2.07 states that strict liability statutes apply to corporations unless indicated otherwise by legislation, and specifies three situations for non-strict liability offences where a corporation can be found to be criminally liable:

(1) The first base of liability is when an agent acting on behalf of the corporation within the scope of his employment, commits a misconduct that violates

40 Bribery Act 2010 (UK) c 23, s 71.
41 Ibid.
42 Ibid, s 8.
44 Ibid at 319.
45 Sara Sun Baele, “The Development and Evolution of the U.S. Law of Corporate Criminal Liability” (German Conference on Comparative Law, September 2013) at 10, online: <http://scholarship.law.duke.edu/faculty_scholarship/3205>.
statutes outside the Code and plainly state that there is a legislative purpose for imposition of the liability.

(2) The second base of liability applies when corporation fails to meet a duty of affirmative performance imposed by law.46

(3) The third base of liability applies to all situations where the board of directors or a high managerial agent acting on behalf of the corporation within the scope of their employment “authorised, requested, commanded, performed or recklessly tolerated” an offense.47 A high managerial agent is defined as an officer of a corporation that has “duties of such responsibility that his conduct may fairly be assumed to represent the policy of the corporation or association.”48

The MPC has a more restrictive approach because the third base of liability adds an approach similar to the identification doctrine and a reserves a ‘due diligence’ defence for the first base of liability that is governed by the respondeat superior doctrine and aggregation theory.

CHALLENGING THE IDENTIFICATION DOCTRINE STANDARD

The common law identification doctrine is a deficient tool for the effective enforcement of criminal law against large modern corporations as it fails to address the complexity of current corporate structures. It is often difficult to identify the specific senior officer(s) responsible for a violation of a certain criminal law. Limiting liability to ‘high managerial agents’ fails to recognize that those who have managerial positions can, in practice, delegate tasks to those below who then commit the offence.49 Given the complexity of corporate structures, it is also difficult to identify specific agents within the corporation to fulfill both the actus reus and mens rea components.50 Tariq states that “critically, the larger the company is, the harder the prosecutor’s task is to locate first, the culpable alter ego of that corporation.”51 In the context of large modern international corporations it is difficult to attribute responsibility to directors and individuals who have official duties alone.

47 Ibid, s 2.07(c).
48 Ibid, s 2.07(a)(b).
The identification standard also encourages corporations to decentralize responsibilities to avoid liability by making it difficult to identify any one senior individual in charge of a particular operation. In practice, senior officers can delegate their tasks to low level employees to avoid criminal liability, thus making it difficult to identify the specific agents that fulfill the *actus reus* and *mens rea* components of many criminal offences. It is often hard to prove whether the delegation actually took place at all, as the corporation can always argue that the lower level employee committed the act without any authority from the corporation itself. The UK Ministry of Justice has stated that the identification doctrine may encourage bad corporate culture and practices, such as manipulating meeting minutes through not recording those present in order to conceal the presence of board members, or creating and using zero-asset companies to handle negotiations with third party agents.\(^52\)

In the United Kingdom, the approach by the courts has not always been consistent, particularly with regards to ‘hybrid’ regulatory offences that have defences based on due diligence or lack of knowledge, or where the definition of the offence requires constructive knowledge.\(^53\) The approach of the courts in classifying these offences has been to consider the specific circumstances of each case. In *Tesco v Nattrass*, Lord Reid states that “Parliament has chosen to deal with the problem piecemeal... the main object of these provisions must have been to distinguish between those who are in some degree blameworthy and those who are not, and to enable the latter to escape from conviction if they can show that they were in no way to blame.”\(^54\)

In *Tesco v Nattrass*,\(^55\) a corporation was charged with an offence under the *Trade Descriptions Act, 1968* when a cashier employee sold a product above the advertised price, contrary to the training the employee received and the orders of the corporation. The defence to the violation invoked was mistake; that is, either reliance on flawed information supplied to the person or an accident beyond the control of a person could excuse the commission of the offence (provided the person took all reasonable precautions and exercised due diligence).\(^56\) The defence was raised based on the fact that the criminal acts of a manager in a store could not be attributed back to the corporation.

---

52 Ministry of Justice, *supra* note 3.
54 *Tesco v Nattrass*, *supra* note 33 at 169–70.
56 *Ibid* at 168.
Here, all reasonable precautions and due diligence were exercised by the corporation to avoid the commission of such an offence by himself or someone under his control.\textsuperscript{57} The Court held that vicarious liability could only apply if the criminal act was committed by an employee without the consent of the corporation and that the corporation could not be blamed for the acts of employees conducted against the procedures of the corporation.\textsuperscript{58} This approach is still followed by recent cases including \textit{Ferguson v British Gas Trading Ltd}\textsuperscript{59}, \textit{Jetivia SA and another v Bilta (UK) Limited (in liquidation) and others}\textsuperscript{60} and \textit{G-Star Raw Cv v Rhodi Ltd}.\textsuperscript{61}

On the other hand, the case was distinguished in \textit{Tesco v Brent London}.\textsuperscript{62} In that case, the court held that in certain hybrid offences the vicarious liability route towards establishing liability would be followed when a manager of a store sold a video with an age classification certificate to a person below the specified age. Section 11 of the \textit{Video Recordings Act, 1984} regulated the selling and supplying of video work with age classification certificates. It would thus be an offence to sell or supply a video to someone under the age listed in the classification certificate.\textsuperscript{63} The defences offered included a lack of knowledge or reasonable grounds to believe that the classification certificate contained the age statement, or a lack of knowledge or reasonable grounds that the person was not of the required age.\textsuperscript{64} The Court held that regulations would be followed in order to hold the corporation liable and the corporation could not rely on the defence of ‘lack of knowledge’ in that case. Importantly, the case did not have a ‘due diligence’ defence like \textit{Tesco v Nattrass}. Additionally, the act was committed by a senior manager rather than an employee. Lord Buckley stated “Can a company which contracted to supply a video rely on a defence ...on the basis that its employee’s state of mind cannot be imputed to itself? ...No.”\textsuperscript{65} The

\textsuperscript{57} \textit{Ibid} at 153, 167.
\textsuperscript{58} \textit{Ibid} at 173.
\textsuperscript{59} \textit{Ferguson v British Gas Trading Ltd.} [2010] 1 WLR 785.
\textsuperscript{60} \textit{Jetivia SA and another v Bilta (UK) Limited and others} (2015), [2015] USKC 23; \textit{Tesco Nattrass}, supra note 33.
\textsuperscript{61} \textit{G-Star Raw CV v Rhodi Ltd & Ors} (2015), [2015] EWHC 216 (Ch).
\textsuperscript{62} \textit{Tesco Stores Ltd v Brent London Borough Council} [1993] 2 All ER 718 [\textit{Tesco v Brent London}].
\textsuperscript{63} \textit{Video Recordings Act 1984} (UK), c 39, s 11(1).
\textsuperscript{64} \textit{Ibid}, s 11(2).
\textsuperscript{65} \textit{Tesco v Brent London}, supra note 62 at 1045.
Court reasoned that if the directing mind approach was followed and that defence was to be applied, no large company could ever be held liable.66

The requirement for identifying agents and obtaining sufficient evidence that they committed a criminal act and that they intended to commit that act makes it difficult to hold corporations with complex organizational structures liable. In Tesco v Nattrass, the Court held that when an agent delegates managerial functions within a corporation with the result that the commission of a criminal act occurs, this comes within the directing mind and will principle and the corporation should accordingly be held liable.67 Celia Wells rightly argues that the traditional theories of corporate liability fail to “[tackle] the question of corporate risk taking.”68 It is often still difficult, however to prove whether criminal conduct has in fact been delegated to an employee or not in the appropriate sense of the term.

The UK Parliament has attempted to overcome such difficulties by codifying new standards of liability for different types of corporate misconduct. The Bribery Act carves out a due diligence defence and the Corporate Manslaughter and Corporate Homicide Act introduces a holistic approach to identifying whether management has failed to meet the standard required under the law. The UK Ministry of Justice’s recent call to change the standard of liability limits some reforms to the law for economic crimes, which are limited to money laundering, fraud, and false accounting.69 The report recognizes the failure of the identification model in combatting corporate crime due to the inherent difficulties of finding corporations liable described above. The question remains: is such a move going in the right direction?

This article argues that legislative reform in the United Kingdom and United States is necessary and inevitable. Nevertheless, there should still be steps taken towards a comprehensive reform of the law rather than reform on a piecemeal basis. Economic crimes could, for instance, encompass misconduct well beyond that specified in the call for evidence; corporations can always be said to be committing criminal misconducts for economic benefit. Reforming the law narrowly would likely lead to different standards of liability for comparable misconduct, leading to a fragmented system. The next section will explore the benefits of adopting an alternative standard that

66 Ibid at 1042.
67 Tesco v Nattrass, supra note 33 at 174.
68 Wells, supra note 30 at 44.
69 Ministry of Justice, supra note 3.
overcomes the flaws discussed with regards to the \textit{respondeat superior} standard and the identification doctrine.

3. \textbf{LESSONS FROM CANADA: MOVING TOWARDS AN ALTERNATIVE ‘IDENTIFICATION DOCTRINE’ STANDARD}

The Canadian federal system may be an alternative standard which offers a multi-dimensional approach through legislation. House Government Bill C-45 was passed in 2003 to amend the Criminal Code\textsuperscript{70} and “modernise the law with respect to the criminal liability of corporations and sentencing of corporations.”\textsuperscript{71}

Corporations are covered under the Criminal Code. Section 2.1 states that ‘every one’, ‘person’, or ‘owner’ includes “public bodies, bodies corporate, societies, companies.”\textsuperscript{72} Section 22.1 governs offences where negligence must be proven for an organization to be prosecuted for a criminal offence. It states:

In respect of an offence that requires the prosecution to prove negligence, an organization is a party to the offence if:
(a) acting within the scope of their authority
   (i) one of its representatives is a party to the offence, or
   (ii) two or more of its representatives engage in conduct, whether by act of omission, such that, if it had been the conduct of only one representative, that representative would have been a party to the offence; and
(b) the senior officer who is responsible for the aspect of the organisation’s activities that is relevant to the offence departs – or the senior officers, collectively, depart markedly from the standard of care that, in the circumstances, could reasonably be expected to prevent a representative of the organisation from being a party to the offence.

In \textit{R v Metron Construction},\textsuperscript{73} an Ontario corporation was convicted of criminal negligence when an independent contractor hired by the corporation as a site supervisor departed from the standard of care expected of a reasonably prudent person. He was hired to work as a site supervisor to manage a project restoring concrete balconies in Toronto. The way the

\textsuperscript{72} Ibid.
\textsuperscript{73} R v Metron Construction Corporation, 2013 ONCA 541, OJ No 3909 (QL).
operations were conducted resulted in the death of four workers on the site.\textsuperscript{74} He failed to ensure that the workers received written instructions in their respective languages for the use of fall protection systems and did not instruct the workers to use the swing stage in accordance with safety practices, nor did he prevent bodily harm and death through ensuring that lifelines were used during work hours. The Court held that the independent contractor came within the definition of a senior officer as defined by section 2 of the Criminal Code.

Section 22.2 further governs conduct that is based on a mens rea standard, excluding negligence. It states:

In respect of an offence that requires the prosecution to prove fault – other than negligence – an organisation is a party to the offence if, with the intent at least in part to benefit the organisation, one of its senior officers

(a) acting within the scope of their authority, is party to the offence;
(b) having the mental state required to be party to the offence and acting within the scope of their authority directs the work of other representatives of the organisation so that they do the act or make the omission specified in the offence; or
(c) knowing that a representative of the organisation is or is about to be party to offence, does not take all reasonable measures to stop them from being a party to offence.

Importantly, Section 217.1 concerning the Duties Tending to Preservation of Life states that:

everyone who undertakes, or had the authority, to direct how another person does work or performs a task is under a legal duty to take reasonable steps to prevent bodily harm to that person, or any other person, arising from that work or task.

Clarifying the meaning of these provisions, Section 21(1) states that parties to offences include those who “actually commit the crime, does or omits to do anything for the purpose of aiding any person to commit it; or abets any person in committing it.” Section 2 defines a senior officer as:

a representative who plays an important role in the establishment of an organization’s policies or is responsible for managing an important aspect of the organization’s activities and, in the case of a body corporate, includes a director, its chief executive officer and its chief financial officer.

In Global Fuels (R c Pétroles Global Inc),\textsuperscript{75} a corporation was found liable when a regional manager illegally participated in price fixing and, with his

\textsuperscript{74} Ibid.
\textsuperscript{75} R c Pétroles Global Inc., 2013 QCCS 4262.
knowledge, allowed his six territory managers to engage in price fixing also. The case affirmed that middle managers are included in the definition of ‘senior officer’ under Section 2 of the Criminal Code, and that their criminal conduct can indeed be attributed to the corporation.

The Canadian approach advantageously takes into account the differences between corporate structures for offences requiring mens rea. Through this approach, corporations are unable to evade liability simply by delegating duties to lower level managers, and would not be held responsible when low level employees commit criminal acts that the corporation did not authorize and took reasonable steps to avoid. It helps address the perspective that corporations are more easily held liable than the rest of society. This would not undermine the protection and restoration of rights of individuals within the society, and would serve to increase public confidence in the criminal justice system on the whole.

Having discussed the recent call for evidence by the UK Ministry of Justice for further piecemeal codification of certain types of corporate misconduct, Canada’s identification principle offers an approach that benefits from an overall examination of corporate criminal liability. Reforms that categorize corporate misconduct and adopt different standards of liability for such misconduct run the risk of miscategorizing and setting different standards of liability for criminal acts that may fall under more than one category.

Movement towards codification of the law may be interpreted as a practical impossibility given that it would involve abolishing a number of recent laws passed in recent years. Theoretically, however, there are advantages to adopting two standards of liability that categorize misconduct according the type of mens rea required. The process of changing the law must include agreement on why and when corporations should be criminally liable and theoretical approaches to retribution and deterrence need to be consider in order to justify such changes. Further, the impacts of punishment on stakeholders and the complexity and variety of corporate structures as well as the short and long-term impact of prosecution on the operations of defendant corporations and other corporations in similar industries must be factored.

CONCLUSION

Overall, this article argues for changes to the current standards of corporate criminal liability in the UK and USA in order to better combat current changes in corporate activities, corporate governance structures, and increased globalization, all of which are necessary to achieve the criminal law
goals of deterrence and retribution. The failure of early cases to examine the objectives of criminal law as applied to corporations and to address how such objectives could be achieved resulted in the adoption of standards of liability that are not aligned with the objectives of sentencing. Canada offers an alternative approach that stands as a compromise between the respondeat superior standard and the strict identification model. It also provides insight into the advantages of a comprehensive regime that categorizes the standard of liability in criminal negligence and states of mind, as opposed to various standards of liability which depend on the specific type of the corporate misconduct. Such an approach would overcome the flaws of over-deterrence and bring unity to the different standards of liability that may fall under more than one category.
Friends with Net Benefits: the Investment Canada Act and State Owned Enterprises

ROBERT SROKA*

SYNOPSIS

This paper evaluates the net benefit test of the Investment Canada Act as it pertains to the investment review of acquisitions by foreign State Owned Enterprises in Canada’s resource sector. Through a literature review, followed by a comparative analysis of Canadian and international acquisition examples and investment review frameworks, the paper leads to a law reform proposal for modifying the net benefit test.

I. INTRODUCTION

This project examines the framework surrounding foreign investment by state-owned enterprises (SOEs) in Canada’s resource sector, with a particular focus on the net benefit test (NBT) of the Investment Canada Act (ICA) and SOE investment originating from China. More specifically, the

* Robert Sroka is a PhD and LLM student at the University of Michigan. He holds a BA and JD from the University of British Columbia, as well a LLM from the University of Calgary. He is called in both British Columbia and Alberta.
research problem addresses whether the current framework best balances the inherent competing interests in foreign investment review and maximizes net benefit to Canada, with specific regard to resource sector acquisitions of SOEs, and asks whether a revised review standard would better fulfil this objective. In addition to more lengthy guidelines on SOE transactions, the basic policy considerations behind the NBT, which will be explained later in full, are:

- the effect of the investment on the level and nature of economic activity in Canada;
- the degree and participation by Canadians;
- ... productivity, efficiency, technological development, product innovation and variety;
- competition in Canada;
- the compatibility with national industrial, economic and cultural policies; [and]
- Canada's ability to compete in world markets.

Generally, complaints with the status quo concern the complete discretion possessed by the Minister of Industry (and in practice, the Prime Minister’s Office) to decide approval and attach conditions based on the above noted set of policy considerations without providing any public reasons as to why a decision was made. In turn, both investors and potential acquisition targets are left committing significant resources to unclear targets veiled by considerations of political expediency.

This lack of clarity and certainty, combined with the risks of arbitrary decision-making, may effectively depreciate Canadian assets as prospective acquirers are dissuaded by the opportunity costs of the process, which in turn reduces the range of potentially competitive bids. In the context of SOEs, these impacts are compounded by a more stringent review process with correspondingly weaker probabilities of success, and, even if success is attained, costlier conditions for approval. Accordingly, I argue that with a modified NBT Canada can introduce greater clarity and certainty to the

---

1 Investment Canada Act, RSC 1985, c C-28 (1st Supp) [ICA].
review process and remove much of the scope for arbitrariness, while still setting a high bar for review.

After this introductory section, I will move to a history of foreign investment review in Canada and an evaluation of the current the ICA and NBT in detail. This will be followed in section III by a survey of the literature on Canadian foreign investment review and select major Canadian transactions relevant to the SOE and resource context. Next in section IV, I examine the broader literature focused on foreign investment and SOEs, including sovereign wealth funds, with a particular focus on Chinese SOEs and the corresponding sub-literature. From there, I move in section V to an overview of and comparison with the American and Australian review regimes, as well as select major transaction reviews from both countries. Finally, I offer a reform proposal prior to a short conclusion.

II. FOREIGN INVESTMENT REVIEW IN CANADA: THE FOREIGN INVESTMENT REVIEW ACT

Out of concern over foreign ownership of major businesses in Canada and increasing nationalist sentiment, the early 1970s saw the passing of the Foreign Investment Review Act (FIRA). Prior to the FIRA, no large scale regulation of foreign investment existed in Canada and estimates pegged over one-third of enterprise in the country as being controlled by foreign, mostly American, entities. This led to a perception of Canada becoming a “branch plant” economy where Canadian businesses did not evolve beyond subsidiary status, as well as less favourable tax revenue outcomes for government and forced capital flight out of Canada.

The FIRA regulated three investment types made by a “non-eligible” person: acquisition of control, expansion of existing “non-eligible” investment, and new “non-eligible” business. These investment forms triggered the requirement for the proponent to demonstrate that the investment would be of significant benefit to Canada. The predecessor of the modern NBT, like the current NBT, placed the affirmative burden on the applicant to show significant benefit and provided a similar laundry list of factors that could be considered in the decision, including:

5 Foreign Investment Review Act, RSC 1973, c C-46.
7 Ibid at 178-79.
8 Ibid at 181.
(a) the effect of the acquisition or establishment on the level and nature of economic activity in Canada, including ... the effect on employment, on resource processing, on the utilization of parts, components and services produced in Canada, and on exports from Canada;
(b) the degree and significance of participation by Canadians in the business enterprise . . . and in any industry or industries in Canada of which the business enterprise . . . would form a part;
(c) the effect of the acquisition or establishment on productivity, industrial efficiency, technological development, product innovation and product variety in Canada;
(d) the effect of the acquisition or establishment on competition within any . . . industries in Canada; and
(e) the compatibility of the acquisition or establishment with national industrial and economic policies, taking into consideration industrial and economic policy objectives enunciated by . . . any province likely to be significantly affected by the acquisition or establishment."

If the similarities were not already obvious, in writing on the FIRA prior to the ICA’s creation, O’Sullivan commented that:

the test of significant benefit is so vague that decisions may be made on almost any basis whatsoever - including a political basis. Since the test is not a clear-cut one, the review process is turned into more of a negotiating session between the government and the investor, rather than a rigid application of rules. Under the FIRA, then, Canada has a chance to shape what benefits it can expect from a given investment proposal.10

Unlike its successor however, the FIRA came with an administrative review agency. Although the Foreign Investment Review Agency was intended to be an independent tribunal, its recommendations were subject to final ministerial approval.11 This veto power made outcomes potentially hostage to political incentives where the review agency’s decision failed to provide political cover, even if this was rare in practice. Further, the body of decisions surrounding the FIRA is only of limited use in the SOE context. With the FIRA being a creature of concerns over American investment, the outward oriented SOEs of modern times were absent from the equation.

The Investment Canada Act

In 1985, the FIRA was replaced with the ICA by the Mulroney government and the ICA has since been altered at its fringes on numerous occasions. With non-cultural properties, the most recent amendment to the

9 Ibid at 187.
10 Ibid at 187-88.
11 Ibid at 188.
ICA increased the review threshold to $1 billion for direct acquisition of a Canadian enterprise by an investor from a WTO member party, up from $600 million and $369 million previously. This amount is calculated on new enterprise value (as opposed to book value) basis, with different formulas for targets based on their status of being publicly traded or privately held companies, or mere asset collections. The enterprise value can be deceptive as it is possible for transactions that would not have been reviewable under the previous rules to trigger review under the new formula because of significant liabilities or market capitalization.

**Administrative Procedures**

In response to a proposed investment, the Minister has 45 days to make a decision, with the unilateral option for a 30 day extension and the ability to secure a further extension with the proponent’s consent. The absence of a decision or extension notice within the prescribed timeframes means that the proposal has been approved. The Minister also has the option to reject the investment but offer the proponent 30 days within which to resubmit, with revised undertakings and representations. In making a decision on net benefit, the Minister will on paper consult with the provinces and other federal ministries, although the unadvertised and perhaps crucial decision-making centre is the Prime Minister’s Office. With strong trends in recent decades for increased size, stature and centralized control within the Prime Minister’s Office, the unofficial reality is that decisions are filtering down through staff members often primarily concerned with political optics.

Eighteen months after a reviewed and implemented investment is approved, the government will conduct a general performance evaluation. This evaluation does not reveal any hard criteria beyond that “investment performance will be judged in the context of the overall results” and that

---

15 Ibid.
16 Ibid.
substantial consistency with expectations, taking into account “subsequent economic circumstances”, will mean no further grounds for monitoring.\textsuperscript{17} Most interesting is that the government is provided an easy escape from scrutiny beyond initial performance monitoring through an explicit outlining that an investor will not be held responsible for a failure to meet undertakings that are (very generally) beyond their control.

The Net Benefit Test

The NBT is not particularly complex. For proposed investment from SOEs and non-SOEs, regardless of WTO membership status of the originating party’s territory, the core of the test is the same, with the Minister of Industry needing to be satisfied that the transaction is likely to be of net benefit to Canada. The Minister may consider the following factors:

(a) the effect of the investment on the level and nature of economic activity in Canada, including, without limiting the generality of the foregoing, the effect on employment, on resource processing, on the utilization of parts, components and services produced in Canada and on exports from Canada;
(b) the degree and significance of participation by Canadians in the Canadian business or new Canadian business and in any industry or industries in Canada of which the Canadian business or new Canadian business forms or would form a part;
(c) the effect of the investment on productivity, industrial efficiency, technological development, product innovation and product variety in Canada;
(d) the effect of the investment on competition within any industry or industries in Canada;
(e) the compatibility of the investment with national industrial, economic and cultural policies, taking into consideration industrial, economic and cultural policy objectives enunciated by the government or legislature of any province likely to be significantly affected by the investment; and
(f) the contribution of the investment to Canada’s ability to compete in world markets.\textsuperscript{18}

Guidelines and Statements on Assessment of Net Benefit for SOEs

The government introduced SOE guidelines in 2007, in response to public and media attention paid to foreign investments in the resource sector,

\textsuperscript{17} ICA Guidelines, \textit{supra} note 2.
\textsuperscript{18} \textit{Ibid}, s 20.
including a bid by a Chinese SOE for mining giant Noranda,\(^\text{19}\) as well as more general concerns over potential long term detrimental effects on the Canadian economy due to the “adverse effects on the efficiency, productivity and competitiveness” of SOE acquisitions.\(^\text{20}\) These guidelines were updated in 2012 and the government provided a new statement on SOE investment. The SOE guidelines defined a SOE as “an enterprise that is owned, controlled or influenced, directly or indirectly by a foreign government.”\(^\text{21}\) Any proponent of a reviewable investment is required to disclose their controlling party and any state ownership or control, whether or not the state role is direct or indirect.\(^\text{22}\)

For all deemed SOEs, the proponent will have to address their vulnerability to state guidance in their proposal and undertakings as well as show how their operations will be both transparent and commercial in nature (as opposed to being driven by non-commercial interests).\(^\text{23}\) The guidelines include example undertakings such as placing Canadians in positions of senior management or as independent members on the board of directors, incorporation in a Canadian jurisdiction and listing of the target business on stock exchange in Canada.\(^\text{24}\)

In terms of review, the policy is for the Minister to determine net benefit in accordance with the already existing general principles under the ICA. Additionally, the review process will go into detail on projected commercial impacts, which include:

- export;
- ...process[ing];
- the participation of Canadians in its operations in Canada and elsewhere;
- the impact of the investment on productivity and industrial efficiency in Canada;
- support of on-going innovation, research and development in Canada;

---


\(^\text{21}\) ICA Guidelines, supra note 2.

\(^\text{22}\) Ibid.

\(^\text{23}\) Ibid.

\(^\text{24}\) Ibid.
and the appropriate level of capital expenditures to maintain the Canadian business in a globally competitive position.25

The December 2012 “Statement Regarding Investment by Foreign State-Owned Enterprises” (the Statement) reinforced and highlighted that the enterprise must be free from political interference, comply with the laws of Canada, and make broadly “positive contributions to the productivity and industrial efficiency of the Canadian business”.26 The Statement made clear that SOE acquisitions of Canadian holdings in oil sands would be deemed of net benefit only on an exceptional basis.27 These additional burdens for SOEs have translated into lengthier and more contentious review processes.28 The Statement must be understood in the context of it being delivered on the same date as the government’s approval of two major energy sector acquisitions by SOEs that this paper will address later: Nexen by the China National Offshore Oil Company (CNOOC) and Progress Energy by Malaysian giant Petronas.

The following spring, the government in its omnibus budget bill amended the ICA to reflect the new guidelines and the Statement. The amendments included a formal definition of SOE to the following:

State-owned enterprise means
(a) the government of a foreign state, whether federal, state or local, or an agency of such a government;
(b) an entity that is controlled or influenced, directly or indirectly, by a government or agency referred to in paragraph (a); or
(c) an individual who is acting under the direction of a government or agency referred to in paragraph (a) or who is acting under the influence, directly or indirectly, of such a government or agency.29

However, no further interpretation has been given regarding what constitutes direct or indirect control,30 seemingly leaving some uncertainty and considerable leeway for the government to make determinations as it

25 Ibid.
26 Ibid.
27 Ibid.
29 ICA, supra note 1, s 3.
pleases. This is further amplified with the amendments which implement a means for the government to make retroactive determinations of SOE control. Practitioners have noted that this could lead to the Minister deeming a company otherwise considered a “Canadian business” under the ICA as an SOE and have warned that such a step breeds major uncertainty.

Guidelines on Acquisitions of Oil and Gas Interests

The ICA guidelines have also been created for certain energy sector acquisitions. With the threat of a “natural resources grab” ever present, special guidelines are to be expected. However, as opposed to directly addressing concerns about foreign governments taking control of potentially strategic assets, the guidelines provide more practical guidance from a transactional perspective, namely how “who” holding “what” will be viewed through the ICA lens.

Of note, mere exploratory interests that are not currently capable of producing revenue are not considered businesses under the ICA and thus are not subject to notification or review. However, while proven reserves or deposits with “economically recoverable quantities” will be viewed as a business once recovery for production has started, this would still seem to leave a minor potential regulatory gap. Here, a foreign state-controlled Canadian number company could acquire exploration rights prior to

---

31 Canadian business means a business carried on in Canada that has (a) a place of business in Canada, (b) an individual or individuals in Canada who are employed or self-employed in connection with the business, and (c) assets in Canada used in carrying on the business; ICA, supra note 1, s 3.
32 McCarthy Tetrault, supra note 30.
33 The Library of Parliament’s 2011 research publication on foreign investment defines a “natural resources grab” as where “a foreign investor tries to acquire a company strictly for its natural resource assets while shedding all value-added activities associated with it.” The paper goes on to further highlight “The Special Place of Natural Resources in Foreign Investment Review”, noting that natural resources bring with them unique issues in terms of inherent supply limitation, nation-state specificity creating unusual bargaining clout, and ability to separate or tie together value-added opportunities via vertical control and integration or lack thereof. See Frigon, supra note 14 at 8.
34 “ICA Interpretation Note No. 4 – Business” highlights that the exemption for exploration and pre-production operations applies the same with mineral extraction outside of oil and gas as well. However, a mine or extraction operation that has been closed is classified as a “business”. See Industry Canada, “Investment Canada Act Interpretation Notes”, (6 May 2011), online: <http://www.ic.gc.ca/eic/site/ica-lic.nsf/eng/h_lk0067.html>.
35 Ibid.
development of production and prospectively avoid the ICA, including the amendments concerning retroactive determination of SOE control discussed above. This being said, with the provinces having significant control of resource tenure, as well as the approval process for extraction, federalism can provide a strong brake. When combined with the breadth of more proven and development ready options that outward state directed investment will have globally, any gap may simply not be that attractive to SOEs or their respective states, although this may be a future subject worthy of inquiry beyond speculation.

National Security Review

Although a somewhat separate concept from the NBT, the national security review is important to understanding the Canadian review framework. Introduced via amendment to the ICA in 2009\(^36\) and recently expanded upon with new Guidelines, national security review can be seen as a fleshing out of, and compliment to, the NBT. Whereas previous to the 2009 amendment, a supposed national security threat would be encompassed under the more general calculations of net benefit, the national security regulations formalized a process for both identification of certain “threats” and their assessment and review.\(^37\) Procedurally, a national security concern arising out of Canadian investments by foreign parties will be identified to the Minister of Industry by the security services.\(^38\) The Minister will decide whether to send the investment to cabinet for potential review, the standard being whether the investment “could be injurious to national security” and cabinet in turn will have the ability to decide on whether to order a formal review.\(^39\) Upon such order, the review will formally be conducted by the Minister of Industry, in concert with the Minister of Public Safety, and then submitted via a report with recommendations to cabinet.\(^40\)

The 2016 Guidelines detailed nine factors the Minister may take into account in assessing investments under the national security framework,

\(^{36}\) Further amendments to the national security review regulations were made in 2015, but mostly concerned procedural timelines, namely providing the government “with the flexibility to extend time periods for the review”. See Regulations Amending the National Security Review of Investments Regulations, Canada SOR/2015-311.

\(^{37}\) Ibid.

\(^{38}\) Ibid.

\(^{39}\) Ibid.

\(^{40}\) Ibid.
although criteria beyond these nine were not precluded from consideration. These factors are:

- The potential effects of the investment on Canada’s defence capabilities and interests;
- The potential effects of the investment on the transfer of sensitive technology or know-how outside of Canada;
- Involvement in the research, manufacture or sale of goods/technology identified in Section 35 of the Defence Production Act;
- The potential impact of the investment on the security of Canada’s critical infrastructure. Critical infrastructure refers to processes, systems, facilities, technologies, networks, assets and services essential to the health, safety, security or economic well-being of Canadians and the effective functioning of government;
- The potential impact of the investment on the supply of critical goods and services to Canadians, or the supply of goods and services to the Government of Canada;
- The potential of the investment to enable foreign surveillance or espionage;
- The potential of the investment to hinder current or future intelligence or law enforcement operations;
- The potential impact of the investment on Canada’s international interests, including foreign relationships; and,
- The potential of the investment to involve or facilitate the activities of illicit actors, such as terrorists, terrorist organizations or organized crime.  

These new Guidelines have provided a degree of clarity and transparency to the national security review process, but leave outstanding issues as to how impact is measured, the relative impact and valuation of concerns, and thresholds of acceptable risk.

In the first potential opportunity to evaluate the Guidelines’ operation however, the government declined to even trigger the national security review process for the takeover of Norsat, a defence technology contractor, by a large private Chinese firm, a transaction that sparked the US Department of Defense to announce a review of all Norsat contracts. Speaking on the decision to not proceed with an ICA national security review, Prime Minister Justin Trudeau outlined that “[i]n this case, our very effective national-security agencies made a professional determination that there were no significant national-security concerns about this particular transaction and it didn’t need

---

41 ICA Guidelines, supra note 2.
to go through further reviews." Thus it seems that despite issuing ICA Guidelines only months prior, in practice the government continued the tradition of operating outside of the formal framework on an ad-hoc basis at the direction of the Prime Minister’s Office, even though the public reasoning for the decision to not formally review was seemingly based upon criteria present in the ICA Guidelines.

**Canadian Investment Review Literature**

The bulk of the Canadian-focused literature, whether it comes from the worlds of academia, policy or legal practice, is favourable to both reform of the ICA and investment by SOEs in Canada. Yet the best starting point is perhaps not with the ICA or SOEs at all, but rather with an evaluation of the predecessor FIRA. In the most cited work on the topic, O’Sullivan explores the economic nationalist sentiment that led to the FIRA’s passing, as well as the mechanics of the legislation itself. He is inconclusive on the success of the FIRA regime, noting that the FIRA neither damaged Canada’s international trade relationships, nor was a definitive solution. His judgments, however, are based from the economic nationalist perspective that informed the FIRA itself, which substantially contrasts with both the present-day review framework and to a lesser extent, the perceived broader problems with Canadian foreign investment review (historical fears of being too permissive versus modern thoughts of being too restrictive).

Dawson is far more critical of the FIRA, arguing that it cast too large of a net through its low review thresholds. In turn, Dawson believes the low review thresholds caused significant wait times, which when accompanied with administrative costs and performance requirements “had a chilling effect on investment”. With regard to the ICA, Dawson takes a more positive tone towards both foreign direct investment and the NBT, but acknowledges the potential for a similar chilling effect while seeing the need for SOEs to receive

---

43 Ibid.
44 O’Sullivan, supra note 6.
45 Ibid.
46 Ibid.
additional scrutiny.\textsuperscript{48} However, this openness to differential treatment does not extend more broadly to investment in different sectors.

In June 2008 the government’s Competition Policy Review Panel report challenged the OECD’s consistent assertion that Canada was among its most investment restrictive members.\textsuperscript{49} Still, in addition to proposing raised thresholds and a national security component, the Panel recommended changes to the NBT itself, including reversing the onus whereby instead of making proponents demonstrate that the acquisition would be of net benefit, the government would have to show the investment to be “contrary to Canada’s national interest” to reject the transaction.\textsuperscript{50}

Most others take harsher positions towards the ICA and NBT however. Bergevin and Schwanen (echoed by Deveau and Calof),\textsuperscript{51} argue strongly against the NBT in its current form, citing subjectivity, lack of clarity and the inappropriateness of the test in certain situations. While the authors do not directly address SOEs, they conclude that the test needs to be reformed or replaced.\textsuperscript{52} Bradford goes somewhat further, if into the past, and highlights that the ICA has really evolved into a remake of the FIRA as a “black box of unknowns”.\textsuperscript{53} Globerman, commissioned to study the ICA by the Competition Policy Review Panel, likewise concludes that the ICA’s review process failed to enhance the net benefits of inward foreign direct investment to Canada.\textsuperscript{54} He also suggests that increased review thresholds, a general lowering of administrative costs, and a refocusing towards national security, show promise in rectifying some issues.\textsuperscript{55}

\footnotesize
\textsuperscript{48} Ibid at 6, 14.
\textsuperscript{49} Industry Canada, “Compete to Win” (Ottawa: Industry Canada, 2008) at 29 [Compete to Win].
\textsuperscript{50} Ibid at 32.
\textsuperscript{51} Denise Deveau, “Does the net benefit test make the grade?”, Financial Post (23 May 2013), online: <http://business.financialpost.com/driving-competition/does-the-net-benefit-test-make-the-grade>.
\textsuperscript{52} Phillipe Bergevin & Daniel Schwanen, “Reforming the Investment Canada Act: Walk More Softly, Carry a Bigger Stick” (2011) [unpublished, archived at CD Howe Institute] [Bergevin & Schwanen].
\textsuperscript{55} Ibid at 5, 66-67.
In a similarly conceived and directed publication, Moran uses a survey of global and Canadian acquisitions by Chinese SOEs to argue that fears of a Chinese lock-up of Canadian resources are unsubstantiated. Instead of reviewing and blocking potential Chinese SOE bids for Canadian resource assets based on a net benefit assessment, Moran advocates moving the substance of review to a national security test. Paying considerable attention to BHP Billiton’s failed bid for Potash Corporation, Moran questions the wisdom of denying takeovers in strategic resource sectors by non-SOEs where the only viable alternatives for acquisition are SOEs. The article, aimed most directly at policymakers, is especially useful for its comparison of the relative merits of an SOE review system premised on national security concerns as opposed to the economic, competition and industrial-focused NBT.

Writing from a Chinese perspective after the CNOOC acquisition of Nexen, Woo comes to some analogous conclusions to Moran, setting out that discrimination against SOE investment is not only undesirable, but threatens Canada’s appeal as a destination for foreign investment more generally. Woo further criticises moves by the Canadian government to add review barriers to SOE investment as disproportionate to the risk posed by such investments, arguing instead that since the risk factor “of SOEs boil down to security issues – however defined – the proper mechanism for assessing those risks is the national security provision of the ICA rather than a separate set of guidelines for SOEs that merely reiterate the net benefit test.” Further, while the national security provision may be the more appropriate forum for review, aggressive implementation of national security review cannot be allowed to be used as “a pretext for protectionism or as an excuse for jingoism.”

Finally, the hard political science perspective of Hale examines the federal government and media response where Canadian resource companies came under foreign control with the objective of isolating what balance and policy objectives a seemingly “new” (as of the 2000s) consensus on foreign investment represents. Hale somewhat interestingly concludes that his case

---

57 Ibid.
59 Ibid at 35.
60 Ibid.
studies represent a larger move away from industrial policy towards neoliberalism.\footnote{Ibid.}

**Major ICA Reviews of Canadian Transactions**

**BHP Billiton and Potash Corporation**

The first major red flag\footnote{The quashing of aerospace technology firm McDonald Detweiler’s takeover by Alliant Techsystems (an American company) was the first rejection under the ICA; see Oliver Borges, “Investment Canada Act: Minister of Industry Blocks Acquisition of MDA by Alliant Techsystems” (12 May 2008), online: <http://www.mccarthy.ca/article_detail.aspx?id=3996>.} that ICA approval was not a formality in the resource sector came with BHP Billiton of Australia’s bid for Potash Corporation.\footnote{Stockwell Day & Michael Yang, McMillan LLP, Media Release, “Asian Foreign Investment in Canada: Moving Forward” (June 2012), online: <http://www.mcmillan.ca/mobile/Asian-foreign-investment-in-Canada-moving-forward> [Day].} This $38 billion proposed transaction was the second rejection under the ICA and concerned a private firm from one of Canada’s closest friends, ironically enough aiming to takeover a Canadian SOE.\footnote{Ibid.} Here, the Harper government responded in large part to opposition to the deal from the Government of Saskatchewan, a political ally, despite the fact that Potash was projected to be a drag on Saskatchewan’s budget.\footnote{Ibid.} Potash was deemed to be a strategic asset that could only be entrusted to a private or foreign entity if, in the words of a prominent Harper-era cabinet minister, “it would be developed in a way that brought growth and returns to the province’s tax base.”\footnote{Ibid.} BHP’s troubles also arose from focusing its lobbying efforts on the federal government, and thus not properly negating the significant provincial dissent that came to the forefront.\footnote{Ibid.}

**SINOPEC and Daylight Energy**

Prior to the flashpoint SOE transactions concerning CNOOC and Petronas, considerable attention was paid to Sinopec’s acquisition of Daylight Energy. In addition to being one of China’s largest companies, Sinopec is found in some lists as one of the largest companies in the world. The bid for
$2 billion gained significant attention, being framed in a statement from Prime Minister Harper outlining that Canada “welcomed investment by China and other countries, so long as the acquisitions were economic in nature and don’t have other strategic and political connections.”69 Still, the relatively small nature of the assets involved lessened the prospective strategic impacts and the deal was approved.

This transaction, perhaps because of Sinopec’s place in China’s political economy, generally drew more attention than the parade of other SOE acquisitions in the energy sector in the years leading up to 2012. These included Abu Dhabi National Energy Company’s purchase of PrimeWest, IPIC (also of Abu Dhabi) and Nova, Sinopec’s subsequent investment in Syncrude by taking over ConocoPhillips’ position, Korea National Oil Corporation and Harvest Energy, Norway’s Statoil and North America Oil Sands as well as Thailand’s PTT Exploration into Statoil’s stake.70

Petronas and Progress Energy

Obviously however, the conversation in recent years about SOE investment in Canada’s resource sector is centred on the CNOOC-Nexen and Petronas-Progress Energy transactions, forever paired through their eventual approval on the same day. These two investments signalled the climax of SOE investment in Canadian energy. Beginning with the latter, in June 2012 Malaysia’s national oil and gas company Petronas announced a prospective takeover of Calgary’s Progress Energy. From its namesake twin-tower headquarters in Kuala Lumpur, Petronas at the time held energy interests in more than 30 countries. The Petronas bid, originally for $5.5 billion, was in short order increased to $5.9 billion after a rival made its own acquisition attempt.71 In reviewing the transaction, the government did not provide its decision until mere minutes before the deadline on October 19th, with the

70 Torys, supra note 28.
71 Bennett Jones LLP, Media Release, “Petronas Decision Signals Feds’ Caution on SOE Investment Reviews” (8 November 2012), online: <https://www.bennettjones.com/Publications%20Section/Updates/PETRONAS%20Decision%20Signals%20Feds%20Caution%20on%20SOE%20Investment%20Reviews> [Bennett Jones].
Minister of Industry outlining with no further detail that he was “not satisfied that the proposed investment is likely to be of net benefit to Canada.”

Despite speculation that the rejection centred on the transaction’s own merits, there were rumours that Petronas and the government had been very close to reaching a settlement on approval when Petronas attempted to force the government’s hand by not consenting to an extension of the review given that the government wished to concurrently decide the Petronas transaction with the later developing CNOOC bid for Nexen. Perhaps judging that its approval odds were superior as a Malaysian SOE in an isolated $6 billion transaction than by having its fate paired with a Chinese SOE in a $15 billion transaction, Petronas pushed for an expedient resolution. While the government may have called Petronas’ bluff at first instance, the company knew that even in the event of rejection it would have a further 30 days to provide additional undertakings or revise its deal to alter the outcome. Even at the time of its initial rejection, the deal was considered by well-informed observers to have a realistic opportunity to eventually gain approval.

However, there was some indication that Petronas had underestimated the government relations aspect of ICA review considering the wide ultimate discretion that cabinet has. More troubling may be that the procedural reality means that lobbying efforts can have a measurable impact in the assessment of a supposedly objective test early on. Although lobbying is commonplace in Canadian politics, lobbying attempts to influence a subjective policymaking process (under the guise of ICA “consultations”) as opposed to swaying the interpretation of how a set bar should be raised or lowered in a particular instance, are of questionable compatibility with an objective standard aimed at maintaining a level playing field for investment.

Petronas may seem far less potentially strategically threatening than Sinopec or CNOOC, simply because of its country of origin. It is true that Malaysia is a middling regional power in Southeast Asia with a historically pro-Western bent. However, aspects of Petronas’ operations touch upon red flags typically associated with SOEs, namely corruption.

---

73 Bennett Jones, supra note 71.
74 Torys, supra note 28.
75 Ibid.
76 Bennett Jones, supra note 71.
77 Ibid.
78 Vaughn Palmer, “A $700-Million Corruption Cloud Engulfs Malaysian Prime Minister”, The Vancouver Sun (5 August 2015), online: <http://www.vancouversun.com/news/vaughn+palmer+million+corruption+cloud+engulfs+m
the Malaysian government have been implicated in numerous corruption scandals before and since the Progress Energy transaction, despite formally substantial efforts to control the country’s rampant and ingrained corruption.\(^7^9\) As the contributor of up to 45 percent of Malaysia’s budget revenues at the time of its Progress bid,\(^8^0\) Petronas is difficult to separate from the Malaysian government, which has been effectively controlled by one party for decades.\(^8^1\) Most recently, the government has been embroiled in a scandal that saw some (US) $700 million transferred from a state-owned investment fund to the personal accounts of the Prime Minister.\(^8^2\) From a Canadian perspective, the risk of corruption, especially highlighted in this latest scheme,\(^8^3\) would be asset stripping and skimming of proceeds. Such a distortion may threaten the market orientation of the acquisition and in turn the health of the venture, along with corresponding jobs, revenues and benefits down the Canadian economic chain.

Eventually, in December 2012, approval was granted by the Ministry of Industry after Petronas made “additional representations and...further undertakings”\(^8^4\). These were rumoured to include a Canadian stock listing that would incur local regulatory oversight,\(^8^5\) and the appointment of independent investors to the Progress board.\(^8^6\) Indeed, the period subsequent
to the approval saw Progress double employment in its Calgary office, maintain its local top management and gain access to almost four times as much capital on an annual basis to make further acquisitions, with the company poised to invest many billions more in Canadian natural gas. Yet the ensuing years have only served to underline corruption concerns with the Malaysian government.

CNOOC and Nexen

The more politically contentious of the two reviews, CNOOC and Nexen, could arguably have commenced with another transaction; CNOOC’s $2.1 billion acquisition of OPTI Canada, a relatively small oil sands player with strong pre-existing ties to CNOOC, in November 2011. Nexen had already taken a 15 percent interest in its oil sands joint venture and with the company continuing to struggle, CNOOC stepped in to acquire the whole. As the bankruptcy or major restructuring of OPTI was expected, CNOOC’s OPTI acquisition was not seen in a particularly negative light and thus its review and approval under the ICA garnered far less scrutiny than CNOOC’s subsequent Nexen proposal.

The following July saw CNOOC’s bid for Nexen at a price over 60 percent of its traded value, equating to just over $15 billion, the largest ever transaction between Canadian and Chinese parties. In a relatively quiet time of the year for media, the CNOOC proposal was able to drive the following Canadian news cycle. The Canadian government meanwhile largely deferred comment with leaked reports of a cabinet split on how to handle the proposal, and public polling showing opposition to the deal was countered

---

89 “CNOOC Wraps up Opti Deal to Expand in Oil Sands”, Reuters (28 November 2011), online: <http://ca.reuters.com/article/businessNews/idCATRE7AR1B020111128>.
90 Ibid.
by support for the acquisition from Alberta’s premier and much of the country’s business establishment.\textsuperscript{93}

By the end of August, CNOOC formally made an application for ICA approval and three weeks later the deal was approved by Nexen shareholders.\textsuperscript{94} Public statements from the Prime Minister highlighted the imbalance in the Sino-Canadian trading relationship in China’s favour as well as the need for that to change.\textsuperscript{95} Still, with mixed signals prevailing, the government simply twice deferred the decision to push the limits of the ICA review period in a review process that became not so much an application package being reviewed, but an \textit{ad hoc} (and perhaps FIRA-style) negotiation.\textsuperscript{96} Finally, on December 7\textsuperscript{th}, the Minister of Industry announced the transaction’s approval.\textsuperscript{97} Prime Minister Harper’s statement on the approval was careful to underline that “[i]t is important that Canadian and also foreign investors understand that this is not the beginning of a trend but rather the end of a trend.”\textsuperscript{98}

The transaction and review process should be seen against the backdrop of both a series of acquisitions of large Canadian resource companies in the 2000s as well as already existing joint ventures between SOEs and domestic energy players, and in turn as a signal on whether these partnerships in the oil sands could evolve into controlling relationships.\textsuperscript{99} Hale describes the CNOOC bid as “deeply intertwined in broader political and economic debates over the growing power of foreign SOEs” as well as the specific debate on how to treat Chinese investment in particular.\textsuperscript{100} Here, there was a public debate between realists espousing the net benefit of Chinese investment,
moderates hedging between openness to foreign capital but wanting to limit that capital to positions short of effective control, and those simply opposed.

This latter category emphasized concerns central to investment (including absence of reciprocity for prospective Canadian investment and lack of remedies for CNOOC’s commitments) as well as those beyond the direct scope of investment (human rights, espionage, China’s politically influenced economy and reputation for corruption). Still, Hale noted that the concurrent review of the Petronas bid for Progress Energy served as a moderating force, focusing attention on an objective standard for SOEs as opposed to a standard for SOEs based on national origin. While the concurrent stream was perhaps a negative for Petronas, it was of benefit to CNOOC.

For its part, CNOOC learned lessons from previous review processes to frame its deal in the best possible terms, effectively cornering the government to either approve the transaction on the objective standards of the NBT which CNOOC, advised by leading Canadian lobbying and law firms, had gone to pains to meet, or demonstrate that a rejection was politically driven. CNOOC’s offerings included a commitment to move their head office for non-Chinese operations to Calgary, a public listing on the Toronto Stock Exchange, as well as maintaining Nexen’s management, its 3,000 employees and community contributions.

The government’s response and the split within cabinet can also be seen as being informed more by competing foreign policy approaches towards China – pro trade and investment accompanied by warming relations and engagement on one hand, versus voicing concerns about human rights combined with suspicion of Chinese ambitions, business practices and espionage. Balance was the order of the day, with the Harper government’s position attempting to show that Canada was still open to foreign investment in the energy sector, but if that investment came from a SOE and it was an

101 Ibid at 365-66.
102 Ibid.
103 Namely the failed acquisition of Potash Corporation in Saskatchewan by BHP Billiton.
106 Burton, supra note 92 at 52.
attempt to gain a controlling interest, there would be additional thresholds to meet and the investor would face correspondingly lessened prospects of approval. Yet despite considerable reservations, the government also apparently wanted to avoid perceptions that it had acted arbitrarily by approving the transaction as evaluated by the existing rules, while still effectively closing the door to further transactions.

The entirety of the CNOOC-Nexen review process highlighted the ad hoc, political and somewhat arbitrary nature the decision-making outcomes of the NBT can foster. Net benefit is clearly driven by more than an objective analysis of benefit to the national interest and extends to net benefit of political outcomes for the governing party. Likewise, instead of policy informing the baseline of an objective test, which in turn would be applied to a fact-set to determine outcomes, the government’s maneuverings saw the facts in this instance (and other instances) develop the policy outcome.

However, with the benefit now of several years’ hindsight to evaluate the transaction, the clear reality is that CNOOC has not done well out of its acquisition of Nexen and is widely viewed to have overpaid. Calgary and the oil sands have suffered from a decline in oil prices and CNOOC has shut down the Long Lake oil sands project that was the original spearhead in its Canadian investment. The time elapsed has also shown that CNOOC has not kept certain commitments it made as part of its bid and the breaking of those commitments has not met consequences. CNOOC fired much of Nexen’s top management in late 2014 and brought in replacements from China in order to “better align” Nexen with CNOOC. This was followed up with hundreds of layoffs in both front office and operational positions over 2015 and 2016. CNOOC has likely escaped greater scrutiny for the


108 Ibid.


simple reason that these could be viewed as legitimate business moves considering the underperformance of the Nexen portion of CNOOC as well as the significant turmoil experienced across the Alberta energy sector.

Yet, should further controlling acquisitions by SOEs of Canadian resource assets materialize, those bidders may have to bear the burden of CNOOC’s less than stellar track record in the event they are not already scared off by the post-Nexen SOE regulations and CNOOC’s subpar return on Nexen. This being said, the current government appears more open to Chinese investment than the previous one, leaving space for a new cost-benefit analysis of a new Prime Minister’s Office to emerge with different outcomes, albeit within the same old ad hoc, arbitrary, and politically-based decision-making process.

III. PERSPECTIVES ON FOREIGN INVESTMENT AND SOES

Much of the conversation on foreign investment review policy as it pertains to the resource sector can be viewed through the lens of tension between nationalist and liberal perspectives. Wilson explains resource nationalism as a mercantile approach on the part of a state endowed with natural resources to direct the resource development path in support of certain national objectives that it “would not otherwise obtain if exploitation was left to international market processes.” Simply, this perspective puts forth that since non-renewable resources are by their very nature scarce, the state needs to maximize domestic returns beyond what would be achieved by global markets. These efforts can include price controls or propagation of cartels with like-minded producers, the creation of national resource

---


companies and ownership controls, and activities to increase or develop new downstream sectors.\textsuperscript{115}

The liberal alternative on the other hand, allows the flowing hands of free (or mostly free) market capital to determine outcomes. Traditionally, resource sector foreign investment was viewed as providing the investment receiving country benefits in terms of capital, market access, extraction technology, and technical knowledge.\textsuperscript{116} A trend has also been noted in managing risk through vertical integration – the foreign vertically integrated firm will be able to absorb shorter term market flows that a smaller domestic and single sector exposed firm may not, providing a form of certainty and stability.\textsuperscript{117}

The perceived risks are, however, likewise present and are often viewed as shadowing the justifications for resource nationalism. These include loss of control, subsequent loss of jobs or lessened pressure to maintain jobs in a foreign territory, the loss of higher value economic activities derived from midstream, downstream and corporate activities, as well as reduced income tax revenues. The possibility that sub-optimal business outcomes will ensue relative to when the resource is locally owned and isolated from potentially subsidizing other aspects of an integrated supply chain, is also a pertinent concern.\textsuperscript{118} With the aforementioned possibility for downstream and corporate head office jobs to be consolidated offshore, a revenue regime designed for a current reality may be exposed to a shift in business model.

Still, despite the potential for negative impacts in certain instances and macroeconomic risks, the general consensus is that foreign direct investment is broadly beneficial to the recipient country.\textsuperscript{119} Since the demise of the National Energy Program, this liberal consensus (although not without notable exceptions – such as Russia and Venezuela) has more often than not

\textsuperscript{117} \textit{Ibid} at 143.
\textsuperscript{118} \textit{Ibid}.
generally extended to Canada, as well as the United States and Australia, the two extended examples that this paper will later evaluate.

**Chinese State-Owned Foreign Investment**

As this paper focuses on state-owned foreign investment, the obvious question is whether there is something unique to state-driven investment, and in particular Chinese SOE investment, that diverges from the consensus on foreign direct investment in general. Here, despite the narrowed parameters, the answers are similar. Of course, the practical reality of SOE investment is that Chinese SOEs are perceived as a more significant security threat – after all, the national conversation was not particularly bothered when Norwegian SOE Statoil purchased North American Oil Sands Corporation in 2007.\(^{120}\) Considering China’s geopolitical position, ambitions and the volume of Chinese SOE investment in the natural resources sector, the special attention is warranted – there is no other comparable source of SOE investment that can compete on both these levels.

Structurally, control of China’s SOEs is delegated by the State Assets Supervision and Administration Council (SASAC) and SOEs are separate legal entities from the state with operational independence.\(^ {121}\) On its face, this formal autonomy seems to share much in common with Canadian Crown corporations, although Chen strongly underlines perceived differences between the two, namely that Chinese SOEs are “aimed at creating a government monopoly or dominance” in certain strategic industries and “obtain a competitive edge over private rivals” with their growth often exceeding that found in the private sector.\(^ {122}\)

The SASAC is able to exert macro-control through regulation, in particular the *Interim Regulations on Supervision and Management of State-owned Assets of Enterprises*.\(^ {123}\) This includes the appointment or removal of

---


\(^{121}\) Golding, *supra* note 119 at 540.

\(^{122}\) Duanjie Chen, “China’s State-Owned Enterprises: How Much do we Know? From CNOOC to its Siblings” (2013) [unpublished, archived at the University of Calgary School of Public Policy] at 8.

management, the ability to nominate board directors, ensure how state voting rights are instructed,124 as well as have say in “major matters.”125 These major matters include foreign investments, although here any Chinese firm (public, private or a joint venture with a foreign firm) requires approval from two state bodies – the National Development and Reform Commission (NDRC) and the Ministry of Commerce (MoC).126 Both bodies evaluate the investment proposals on a number of commercial and economic levels, with the former especially focusing on the investment’s compatibility with Chinese industrial policy.127 Once the NDRC and MoC approve, a further review from the State Administration of Foreign Exchange occurs, stemming from China’s outward currency controls.128 This layering of internal bureaucracy, each representing a potentially competing interest and fiefdom, helps sustain arguments that Chinese SOE investment cannot be seen as the state acting with one strategic voice and accordingly, that SOE capital should be treated no different than other foreign capital.

On this particular subject, Du sets forth that despite some legitimate concerns, the “fear factor” on the part of countries receiving investment is considerably in excess of the appropriate grounding for such concerns.129 Expanding on this sentiment from a more case centric view, Zhang demonstrates through CNOOC’s attempted acquisition of Unocal in the US (discussed in greater detail later) and the European Union’s antitrust scrutiny of Chinese SOE transactions, how counterproductive regulatory responses can arise from overstated suspicions of Chinese SOEs.130 Zhang further outlines that many Chinese SOE investments are far more potentially detrimental to China than the countries of investment due to squandering of assets arising from the poor governance and lack of transparency that plagues many SOEs.131

This discussion can be seen as being extended in Drysdale’s analysis of Chinese flows of SOE investment into Australia’s resource sector, which

124 Ibid.
125 Ibid. “Major matters” are covered in arts 22-24.
126 Golding, supra note 119 at 541.
127 Ibid at 542.
128 Ibid.
131 Ibid.
argues that Chinese SOE investment in Australian resources is quite beneficial to Australia.\textsuperscript{132} Cornish makes similar points in the Canadian context, namely that SOEs are misunderstood and should be treated the same as any other foreign investor to the extent they act in such a manner.\textsuperscript{133} Backed by her own literature review comparing Chinese SOEs to non-state foreign investors on a number of fronts, Cornish believes that Canada will be missing a substantial economic opportunity, as well as an opportunity to have a certain degree of influence over these SOEs, should it fail to provide similar treatment.\textsuperscript{134}

\textbf{IV. SOE INVESTMENT REVIEW REGIMES}

The Organization for Economic Cooperation and Development (OECD) conducted a 2015 study on the foreign investment review regimes of its members, looking at the existence and structure of regimes, as well as more specifically examining the treatment of SOEs.\textsuperscript{135} Interestingly, most states were viewed as treating private foreign investment the same as SOEs and only seven of forty-six countries party to the OECD Declaration on International Investment and Multinational Enterprises were deemed to have “specific restrictions” in the context of the Declaration.\textsuperscript{136} These measures can be seen as respecting certain sectors or being more blanketing.\textsuperscript{137} Australia, Canada, Russia and the United States were identified as having “rules specifically addressing inward investments by foreign [SOEs]”.\textsuperscript{138} The American and Australian regimes are further detailed below.

\begin{itemize}
\item Peter Drysdale, “A New Look at Chinese FDI in Australia” (2011) 19:4 China & World Economy 54.
\item Margaret Cornish, “Behaviour of Chinese SOEs: Implications for Investment and Cooperation in Canada” (2012) [unpublished, archived at Canadian International Council].
\item Ibid at 8.
\item Ibid.
\item Ibid.
\item Ibid.
\end{itemize}
United States

Established by executive order in 1975, and with its modern jurisdiction most substantially defined by the 1988 “Exon-Florio Amendment” to the 1950 Defense Production Act, the Committee on Foreign Investment in the United States (CFIUS) process has been tweaked, expanded upon, and significantly reformed numerous times in its history. Of most recent note was the publication of formal guidance to potential acquiring parties in 2008. Unlike Australia or Canada, the American framework for foreign investment review is not strictly mandatory or automatically invoked at certain thresholds. Still, with the CFIUS possessing the ability to retroactively undo transactions, avoidance games would not seem to be an overly prudent tactic. In this light, the scope of the CFIUS review is larger than the ICA, as the CFIUS process potentially encapsulates any investment that is not new, or “greenfield”, investment.

Formally, the CFIUS is a multi-agency committee under the Department of Treasury to “review transactions that could result in control of a U.S. business by a foreign person (“covered transactions”) in order to determine the effect of such transactions on the national security of the United States.” Ultimately however, the President has the sole authority “to suspend or prohibit a covered transaction” and in this respect shares obvious parallels to the reality of ICA review in Canada, where little doubt exists that the Prime Minister holds the effective final say on the Minister of Industry’s review decision. Between 1988 and 2006, of the well over 1,500 transactions the CFIUS has been notified of, twenty-five spurred a full investigation. Thirteen of these transactions were withdrawn when the

141 James Jackson, “The Committee on Foreign Investment in the United States (CFIUS)” (2016) (Washington DC: Congressional Research Service) at 1-3 [Jackson].
142 Ibid at 5.
143 Shima, supra note 135 at 22.
145 Department of the Treasury, “The Committee on Foreign Investment in the United States (CFIUS)” (20 December 2012), online: <https://www.treasury.gov/resource-center/international/Pages/Committee-on-Foreign-Investment-in-US.aspx>.
146 Ibid.
147 Jackson, supra note 141 at 28.
CFIUS provided notice to the proponent of a second stage of review.\textsuperscript{148} Only one of the remaining twelve transactions that went to the President’s desk was not provided a green light.\textsuperscript{149} This sole rejection came in 1990 with the acquisition of Seattle-based manufacturer of aerospace parts Mamco by China National Aero Technology and Export Corporation, an SOE. Stemming from worries over the Chinese government being able to avoid export controls and gain access to technology it would otherwise be denied, then President George HW Bush ordered the divestment of Mamco after the transaction had been concluded,\textsuperscript{150} demonstrating the retroactive remedies available under the CFIUS framework.

Evoking the NBT, the CFIUS has its own laundry list to consider in national security review.\textsuperscript{151} Compared to the ICA framework, the American list focuses far more on national defense interests rather than interests stemming from industrial or economic impact. National defense interests are also not as explicitly separated from industrial or economic policy, but rather entail viewing industrial and economic concerns through a primary and overarching lens of defense. Considering the nature of the US economy, its strategic position and heightened sense of vulnerability to terrorism in the post-September 11\textsuperscript{th} world,\textsuperscript{152} this outlook is not surprising. With the world’s largest economy, diversified amongst many industrial sectors, the ability for any one transaction to impact the economy or a particular sector is minimized relative to what its impact could be in a smaller economy.

When combined with the US’ position of being the de-facto guarantor of security for much of the world and a proponent of the free flow of trade and investment, the strategic impetus seems different. In the aftermath of September 11\textsuperscript{th} which has influenced the development of current criteria,\textsuperscript{153} the US may be more conscious of maintaining its strategic military advantage. Compared to Canada, with a less diverse, highly natural resource reliant, and obviously smaller economy (and military), the manifestation of strategic position and priority seems rationally connected to the legislative outcome in the form of CFIUS.

\textsuperscript{148} Ibid.
\textsuperscript{149} Ibid.
\textsuperscript{151} \textit{Defense Production Act}, supra note 140 at §721(f), cited in \textit{Jackson}, supra note 141 at 21.
\textsuperscript{152} Ibid at 1.
\textsuperscript{153} Ibid.
As specifically applied to SOEs however, the Department of Treasury guidance document dedicates a section to “Foreign Government-Controlled Transactions” (FGCT), which explicitly includes SOEs in addition to foreign government agencies, government pension funds and SWFs. The document takes the position that a FGCT “does not, in itself, mean that it poses national security risk”, explaining that transactions are reviewed holistically on the facts, with overt consideration given to:

the extent to which the basic investment management policies of the investor require investment decisions to be based solely on commercial grounds; the degree to which, in practice, the investor's management and investment decisions are exercised independently from the controlling government, including whether governance structures are in place to ensure independence; the degree of transparency and disclosure of the purpose, investment objectives, institutional arrangements, and financial information of the investor; and the degree to which the investor complies with applicable regulatory and disclosure requirements of the countries in which they invest.154

This list bears considerably more resemblance to Canada’s guidelines on SOE investment and the NBT in general than the CFIUS review factors do to the NBT more broadly, although Canada is seemingly more focused on tangible outcomes (such as corporate board representation).

Notable Foreign Investment Reviews in the United States

In practice, the CFIUS has been involved in several notable proposed SOE transactions since the mid-2000s, and these transactions have led to the further evolution of the CFIUS structure. Two of these instances occurred during the second term of the George W. Bush Administration; the Dubai World Ports acquisition of Peninsular and Oriental Steam Navigation Company (Peninsular) and CNOOC’s proposed acquisition of oil producer Unocal. A third, CNOOC:Nexen, arose during the Obama Administration. Considering the scope and focus of this paper, this section will discuss the CNOOC involved transactions.

CNOOC and Unocal

At the time of CNOOC’s proposed acquisition of Unocal, oil prices were high, driven in large part by Chinese demand.155 Against a backdrop of concerns about China’s rise as an economic and military challenger to the US,156 as well as questions about its propensity to follow the norms of global commerce (with accusations of currency manipulation, failure to abide by the conditions of its WTO assent, and state sanctioned theft of intellectual property at the forefront),157 the bid by CNOOC quickly became a political lightning rod despite Unocal being a relatively insignificant oil company in terms of assets physically located in the US.158

Two months prior to the $18.5 billion (USD) CNOOC bid, global petroleum behemoth Chevron had actually made its own bid for Unocal for $16.5 billion (USD). Within weeks of the CNOOC bid, the House of Representatives passed a creative attack on the CFIUS process via an amendment to prohibit the use of the Treasury to approve the sale of Unocal to CNOOC.159 This amendment came two days prior to the formal CFIUS notice being filed by CNOOC.160 A bill was then introduced in the Senate two weeks later to explicitly prevent CNOOC from acquiring Unocal,161 followed closely by Chevron increasing its total bid as well as the cash component of that bid.162 These legislative events mirrored the public discourse and the advancement of HR 6/PL 109-58 (requiring the study of Chinese energy policy by the Secretaries of Defense and Homeland Security prior to CFIUS review) to final passage and presentation to the President was accompanied with the withdrawal of the CNOOC bid.163 Two days after President Bush signed the Energy Policy Act of 2005 into law, a large piece of

156 Ibid at 1-3.
157 Ibid at 10. At the time, Unocal’s US Operations Accounted for 1 Percent of American Hydrocarbon Consumption.
158 Ibid at 1, 10. Unocal’s Exploration and Drilling Activities at the time were concentrated in Southeast Asia and the Gulf of Mexico. Unocal sold its Canadian subsidiary during July 2005 in a $1.8 billion (USD) transaction.
159 Nanto, supra note 155 at 1.
160 Ibid. The amendment was ultimately struck from the final bill.
161 Ibid at 16.
162 Ibid at 1.
163 Ibid at 16.
Asper Review

[Vol. XVII]

legislation that went well beyond the Unocal issue, and Chevron’s bid was accepted by Unocal, bringing the saga to an end.

Although this outcome precluded the opportunity to see the CFIUS process further unfold in the CNOOC context, it did provide perspective into the political maneuverings that influenced the outcome. Whereas with the fused legislative and executive power found in the Canadian Westminster system, especially in instances of majority government, it is simple to see the executive driven decision-making process as the competing centres of powers in the American system make cause and effect determinations less clear. In the CNOOC-Unocal instance, President Bush and the executive branch were basically silent on the issue. The eventual inclusion of HR 6/PL 109-58 was merely one small part of a major energy bill. It may be that the administration and its allies in Congress were far more focused on passing certain core priorities in the bill, as well as maintaining pace on a broader legislative agenda, than the relative noise of a transaction the President would likely have the opportunity to veto if he so chose.164 Yet in a heavily polarized Congress, the Act passed through the Senate by a vote of 85 to 10.

CNOOC and Nexen

Interestingly, the CFIUS inserted itself into CNOOC’s acquisition of Nexen. While some have argued that the CNOOC-Nexen acquisition was driven far more by interest in Nexen’s interests outside of Canada than any of its Alberta holdings, Nexen’s platforms in the Gulf of Mexico were sufficient to invoke the domain of the CFIUS regardless. The CFIUS approved the investment, however there were apparently significant conditions attached, namely that CNOOC could not be the operator of the assets.165 This meant that CNOOC had to be removed as the decision maker and find a partner to assume the role as necessary.166 Consistent with the broader national security focus, the perceived issue for CFIUS was not so much economic impact of

164 With the decision makers of the CFIUS largely being administration appointees, it follows that the executive branch could sufficiently influence its appointee decision makers to place the matter on the President’s desk.


166 Penty & Forden, ibid.
resource control, but rather the actual physical control of assets geographically close to military installations.\textsuperscript{167} Also, unlike with the ICA where the proponent will make a bid and that bid will either be approved or sent back for revision, the CFIUS will impose conditions for approval specific to that transaction. This provides certainty on what needs to be done on the back-end, but not so much on the front-end, especially with the lack of an official release of previous decisions. Although obviously the ICA process provides no public reasons, a proponent subject to ICA review will have the body of trial and error of previous actions to inform their bid.

Australia

As a prosperous nation with an economy driven in large part by natural resource extraction and with its 23 million people largely concentrated in a few major cities, Australia is a natural peer for Canada. When combined with shared traditions, similar historical development, demographics, political and legal institutions, Australia is likewise the best comparator country for Canada in the context of this paper.

Australia’s foreign investment framework stems from the Foreign Acquisitions and Takeovers Act 1975 (FATA).\textsuperscript{168} As with the ICA, a cabinet minister (the Treasurer\textsuperscript{169}) is the formal decision-maker, advised by the Foreign Investment Review Board (FIRB), a “non-statutory” advisory body supported by a secretariat.\textsuperscript{170} Unlike the ICA, the FATA and FIRB also respectively govern and review foreign acquisitions of real estate interests.\textsuperscript{171}

The FATA has a 30 day review period which can be supplemented by an interim order for a further 90 days.\textsuperscript{172} Decisions come with no right of administrative or judicial review.\textsuperscript{173} Private foreign investment is subject to review based on thresholds in a similar manner (and with comparable thresholds) to the ICA,\textsuperscript{174} but Australia has higher thresholds for certain

\begin{footnotes}
\footnotetext[167]{Ibid.}
\footnotetext[168]{Foreign Acquisitions and Takeovers Act 1975 (Cth) (Australia).}
\footnotetext[169]{The Canadian equivalent is the Minister of Finance as opposed to the President of the Treasury Board.}
\footnotetext[171]{Ibid.}
\footnotetext[172]{Ibid at 12.}
\footnotetext[173]{Ibid.}
\footnotetext[174]{Ibid at 3.}
\end{footnotes}
countries with which it has a free trade agreement, including China. However, where foreign government owned firms, SOEs and SWFs are concerned, any investment is reviewable. A “foreign government investor” is defined as a “foreign government or separate government entity” holding a “substantial interest” of at least 20 percent (increased from 15 percent), or an aggregate of foreign government or entities from more than one country holding a substantial interest of at least 40 percent. Additionally, the 20 percent threshold is invoked if the combined interests of more than one foreign government investor from the same country breach the threshold. This means that if two SOEs originating from the same state run as separate and directly competitive businesses happen to have interests adding up to over 20 percent, there will be a review triggered without any opportunity for it to be proactively avoided.

Instead of the NBT, Australia’s investment regime calls for an evaluation of whether an investment is in the national interest. Framed in similar terms to the NBT, the national interest test takes into account the following factors in all industry sectors: national security, competition, impacts on other Australian government policies, impact on the economy and the community, and “character of the investor”. With the final criterion, character includes transparency of operation and regulatory oversight, corporate governance practices, and in the case of fund managers (including SWFs), the investment policy and proposed use of voting rights.

Specific to resource development, the FATA also applies to interests in mining or production tenements, with no caveat for production status found in the ICA. Instead foreign government investors “require approval to acquire a legal or equitable...interest in a tenement or an interest of at least 10 percent of securities in a mining, production or exploration entity” making for a stronger approval requirement than found with non-producing interests in Canada.

---

176 FIRB Policy, supra note 170 at 3.
177 Ibid.
179 FIRB Policy, supra note 170 at 9.
180 Ibid at 5.
Where proposed investments involve SOEs (and more broadly, where the defined “foreign government investors” are concerned), the most recent Foreign Investment Policy sets out that the review will consider the commercial nature of the operation versus any indicators that decisions could be motivated more by political or strategic considerations than prudent business practice.\textsuperscript{181} The Policy outlines that the following factors will be considered in whether a proposal is “contrary to the national interest”:

- the existence of external partners or shareholders in the investment;
- the level of non-associated ownership interests;
- the governance arrangements for the investment;
- ongoing arrangements to protect Australian interests from non-commercial dealings;
- whether the target will be, or remain, listed on the Australian Securities Exchange or another recognised exchange.\textsuperscript{182}

Shadowing a significant rise in certain commodity prices,\textsuperscript{183} Australia’s resource sector became a significant target for (overwhelmingly Chinese) SOE investment in the mid-2000s, providing a large sample for investment review outcomes. Wilson compiled forty-nine investments with a total value of (AUD) $40 billion by Chinese firms between 2002 and 2010, forty-three of which were made by SOEs.\textsuperscript{184} Wilson argues that this trend of investment demonstrated a strategic objective of “increasing the market power of Chinese firms vis-à-vis major mining companies and securing lower cost minerals supply independent of them.”\textsuperscript{185}

Under its previous policies, the FIRB held the position that an investment would be deemed a risk by virtue of potential strategic motivation at the threshold of 15 percent of an existing enterprise or 50 percent of a new one.\textsuperscript{186} In the most active subset of the studied period, transactions were approved at a 9 to 1 ratio, and the 15 and 50 rule was not strictly followed in cases where a denial would kill a project (in one such instance, there was an acquisition of a bankrupt firm).\textsuperscript{187} Two of four rejections arose from mine locations near remote military sites, while the other two stemmed from violations of the 50 percent rule with the proponents being invited to

\textsuperscript{181} Ibid.
\textsuperscript{182} Ibid at 10-11.
\textsuperscript{183} Wilson, supra note 112 at 291.
\textsuperscript{184} Ibid.
\textsuperscript{185} Ibid at 295.
\textsuperscript{186} Ibid at 297.
\textsuperscript{187} Ibid.
resubmit. This highlights that despite considerable political and public attention due in part to a high volume of transactions, there was a very strong inclination for approval.

A further four FIRB reviews saw conditional approval, where the Treasurer set out binding legal conditions for the transaction to move forward where there was major potential for non-market driven behaviour or where the transaction was especially large. Of the examples Wilson cites, all occurred in mining transactions and included a requirement for arm’s length market output, abstaining from director positions, support for local infrastructure, and a public stock market listing for at least some of the company. This aspect of the process more mirrors the American CFIUS approach of the reviewing body shaping the transaction to meet its review standards instead of the Canadian tack of outlining broad goalposts and certain actions that are more likely to place a deal within those posts, but leaving the specifics to the proponent. However, the conditions do share resemblance to undertakings under the ICA and more generally, Australian parameters in terms of national interest in the first instance obviously resemble Canada’s ICA criteria more than the strategic security driven American review guidelines. Comparably speaking though, it is worth noting that despite similarities between the Canadian and Australian frameworks, Canada’s is viewed as more restrictive and, considering the volume of SOE investment in Australia relative to Canada, Canada may have lost out on investment ultimately directed to Australia.

On a similar tangent, Wilson deduces that the review process is not engaging in “resource nationalism” as the Australian government is not seeking to leverage the review system for local ownership shares or local value added activities, nor is there evidence of specific targeting by national origin (despite regional fears of China’s ever expanding economic and military clout). Combined with a commitment to using the review process to guarantee the normal and free operation of the market as opposed to edging it in a particular direction, Wilson finds Australia’s system to be a positive example of resource liberalism.

---

188 Ibid at 298.
189 Ibid.
190 Ibid.
191 Ibid.
192 Golding, supra note 119 at 573.
193 Wilson, supra note 112 at 299.
194 Ibid at 299-300.
Chinalco and BHP Billiton

Drysdale and Findlay likewise conclude that there is little reason to have any special hesitation about Chinese SOE investment relative to SOE investment in general or even non-state involved private foreign investment. At the core of the Drysdale and Findlay study was perhaps the most significant test of Australia’s foreign investment framework in the SOE context, the multifaceted BHP Billiton (BHP) bid to takeover Anglo-Australian Rio Tinto, which also happened to have a significant Canadian connection. Once BHP had announced its bid, the Chinese SOE Chinalco purchased 9 percent of Rio Tinto, and following the withdrawal of BHP's bid in the wake of Rio Tinto’s own acquisition of Canadian giant Alcan, Chinalco set its sights on an 18 percent stake in the new Rio Tinto in a friendly agreement.195

These maneuvers attracted review from the FIRB via its then guidelines on SOE investment and were accompanied by public statements from the ultimate formal decision-maker, the Treasurer, voicing concerns about commercial versus strategic and political interests being particularly pertinent for large, already existing resource developments where an acquisition could viably exert control over market production and pricing.196 Still, an acquisition of 11 percent of Rio Tinto was eventually approved by the Treasurer, accompanied by certain undertakings on the part of Chinalco. These undertakings were reminiscent of those suggested by the Canadian SOE guidelines and included non-appointment of directors and not making further share acquisitions beyond the approved threshold.

The transaction is also notable for the competition law aspect involved and a formal opinion from the country’s version of the Competition Bureau, the Australian Consumer and Competition Commission (ACCC). The ACCC disputed that even if the Chinese state had the intent to do so, Chinalco’s capacity for spurring a sustained drop in global iron ore prices with the interests of consuming Chinese domestic steel producers in mind was insufficient, outlining that for such tactics to be effective, Rio Tinto would have to consistently continue to generate new projects beyond any realistic scenario for doing so.197 Although neither the FIRB/Treasurer process, nor the ACCC found immediate reason to shut down Chinalco

195 Drysdale & Findlay, supra note 116 at 142.
196 Ibid.
197 Ibid.
making a substantial acquisition in Australia’s resource industry, ultimately Rio Tinto had second thoughts and scuttled the deal.198

Thoughts on Australia

As in Canada, there is reason to believe that Australian investment review policy is driven by political whims, and that that tendency is compounded by state controlled investment. With Chinese investment in Australia being a driving force behind astronomical spikes in the real estate markets of the country’s larger cities, it is no surprise that public polling does not reveal favourable opinions199 on inward flows from China, regardless of industry sector. As Chinese investment in Canadian residential real estate is having similar effects in Vancouver and Toronto as it has had in Sydney and Melbourne, future Canadian public opinion on (and in turn, political response to) natural resource sector Chinese investment might be shaped in part by resentment of occurrences external to traditional resource nationalism concerns. However, in Canada this might take on a geographical dynamic with resource-rich Alberta and Saskatchewan being largely off the current residential real estate map for Chinese investment in Canada. Still, commercial concerns might well prevail at the first instance, with a common thread between Australia and Canada being the questionable returns on already existing SOE investments.

V. REFORMING THE NET BENEFIT TEST

Any reform to the NBT will be based on a change in policy direction from the federal government of the day. While historically there were significant differences between the investment review regimes conceptualized by the governments of Pierre Trudeau and Brian Mulroney, there have not been overhauls on a comparable scale in the over twenty years since Mulroney’s retirement. This is despite both the Liberal and Conservative parties each having ample stints in power and no shortage of divergence in policy preferences or political will to undo the initiatives of the other.

Where the Mulroney-era moves from the FIRA to the ICA were viewed as liberalization, any moves to modify the ICA itself within the context of the NBT would be reform of a somewhat different scope. With the shift from

198 Ibid.
199 Golding, supra note 119 at 543.
the FIRA to the ICA focused on limiting the realm of government oversight through increased review thresholds, the policy discussion on the current ICA has revolved more on adding certainty, predictability and transparency. Although already implemented policy guidelines and statements in more recent years have endeavoured to tinker with thresholds and scope, especially as applied to certain sources of investment in certain sectors, any clarification has done little to remove the veil of darkness surrounding the assessment of net benefit in any particular transaction.

Having already touched on natural resource nationalism versus liberalization, the reality is that most national policies contain some degree of compromise between these competing sets of objectives. The current system has certain benefits in that there is a body of previous review and guidelines, as well as practical expertise from local professionals, for prospective investors to base their proposals on, while still providing the government the heavy hammer of an unrestricted blanket veto with no explanation necessary in spite of its own guidelines. The drawbacks are likewise obvious – namely uncertainty, inefficiency and arbitrariness. The policy design challenge is to address the problems that the literature consensus has identified while not eliminating the state’s ability to filter and potentially reject undesirable transactions.

To various extents of depth and quality, proposals have been made on how to address this design challenge. As previously mentioned, in June 2008, a few months prior to the financial crisis and recession that became the Harper government’s primary focus, the Competition Policy Review Panel report suggested a reframing of the NBT to whether the acquisition is “contrary to Canada’s national interest”.\(^{200}\) This effectively served to reverse the onus of the test. The Panel report described this shift to mean that investment previously held up by the mandatory review process would instead “proceed without intervention from the minister, unless it was a case where the minister’s concern with regard to the factors required to be considered under the ICA rose to the level of the national interest.”\(^{201}\) The Panel also identified transparency and predictability as major challenges posed by the review process and recommended additional guidelines and advisory materials, suggestions that were seemingly acted upon by the Harper government.\(^{202}\)

\(^{200}\) Compete to Win, supra note 49 at 32.

\(^{201}\) Ibid.

\(^{202}\) Ibid at 36-37.
Borrowing and adding depth to the Panel report, one of the better and more comprehensive recommendations, authored by Bergevin and Schwanen for the CD Howe Institute, sets forth a “national interest” test whereby the onus would be reversed and “require the federal government to show that a foreign investment was contrary to Canadian interests in order to block a particular transaction” and further require the government to explicitly set out why a transaction failed. The primary benefit would be, in the opinion of the authors, to show that foreign investment would be no more likely to be subject to government intervention than a squarely domestic transaction. To get to a conclusion on this basic threshold question, Bergevin and Schwanen offer a set of four questions:

1. Does the acquisition threaten to prevent the effective application of competition, commercial, employment, and other laws, regulations in Canada?
2. Is the nature of the investor such that it is likely to make economically unjustified or politically motivated decisions that would disadvantage Canada at the expense of other jurisdictions where the investor operates?
3. Does the proposed investment otherwise threaten the security of Canada or its allies?
4. Does the acquisition demonstrably threaten other significant public policy goals?

All four questions have application to the SOE context, with the second question having the most direct relation in terms of the core concern surrounding SOEs - that they may make non-commercially based decisions detrimental to the local market. While these questions are certainly pertinent, they cover much of the same ground as the current test and guidelines. The difference is the implicit onus - instead of being framed in a neutral style with the need to affirmatively show that the transaction is of net benefit, the questions are phrased to outline a threshold at which a transaction should be blocked and, short of that threshold, the assumption is that a transaction should be allowed. The problem is that while the onus of the question leads to a more explicit dividing line, there is still significant room for subjectivity and arbitrariness in terms of what serves as a sufficient trigger to invoke the veto. The authors appear to temper this through the requirement of clear reasons for rejection and in turn, the creation of a body of reasons (if non-

\[203\] Bergevin & Schwanen, supra note 52.
\[204\] Ibid.
\[205\] Ibid at 15-17.
\[206\] Ibid.
binding) that will create a reference point for future decisions that will be questioned if there is a serious deviation.

A further issue is that by reversing the onus, implicitly the government may end up with a starting point for reasons why a transaction should not be accepted. However, it is possible for the government’s default position to be neutral with the reverse onus only coming into play where the initial review of the evidence indicates substantial reasons for a potential rejection. Once the preliminary neutral review shows reasons for rejection, then the reverse onus standard of being “contrary to the national interest” (or whatever it may formally be) becomes the standard by which a transaction should be conclusively reviewed. If there is no sufficient preliminary indication of deleterious impacts with the transaction, then the deal simply proceeds with no need to positively show net benefit as with the current NBT.

VI. RECOMMENDATIONS FOR THE FUTURE

Much of what Bergevin and Schwanen put forward forms a sound basis for this paper’s proposal. Indeed, this paper’s suggested test borrows an aspect of the onus shift and the requirement for clear reasons to be provided for any rejection of a transaction. However, this paper goes further by addressing remaining concerns and narrowing exposures in terms of certainty, clarity and avoiding arbitrariness. As a result, this paper proposes the implementation of a quantitative-based modification to the assessment of net benefit under the NBT. The parameters of evaluation can be shifted based on the government’s normative preferences as well as through experience and market impact. The key is that empirical criteria are the basis of a scoring system and that these criteria can be measured through quantitative means. Different thresholds could still exist for different transaction amounts, sectors and sources of prospective investment. The questions forming the NBT may be altogether changed to align with the Bergevin and Schwanen test, or simply re-phrased to suggest a preliminary neutral onus and subsequent reverse onus should preliminary review indicate sufficient reasons for concern. The questions may even be left as they are currently phrased with a threshold in the scoring system serving as the onus.

Regardless of policy preferences on the parameters of question semantics, the key is a scoring system resembling that found in government procurement. Certainty will be derived not only from further detail in criteria, but through point value attached to particular criteria. Additionally, greater certainty can be gained through a requirement of written explanations for each scoring component. Moving beyond Bergevin and Schwanen’s
proposal for reasons in the event of a rejection, the provision of reasons in any reviewable transaction, including the vast bulk of transactions that will be approved, will quickly create a body of reference, if non-binding, for future prospective investors to have to base their proposals upon.

As pertaining specifically to SOEs, given that the current test has different thresholds for review and applies certain unique criteria, the revised test could simply address additional concerns through either an increased score requirement for all or certain criteria, or additionally applicable scored questions. The scoring system could also incorporate a midrange where a transaction will be approved with conditions. As far as SOE conditions go, this paper advocates the consideration of so-called “golden share” structures, whereby many of the risks associated with SOE control can be hedged against through allowing capital influx without voting rights attaching to the acquired shares. While such conditions may be a disincentive to investment, a scoring system framed with achievable thresholds where conditions did not need to attach could also provide incentive for prospective investors to raise the bar of their initial proposals to avoid conditions.

Although the issue of subjective or arbitrary application of a scoring system is still present, beyond the check on subjectivity provided by an emerging body of reviews, temperance of politically driven subjectivity can be derived through bureaucratic administration of the test. Further, an Australian-style review board, albeit with a slightly different form and function, has also been demonstrated not to inhibit inward resource sector investment flows from SOEs. In Canada, the Competition Bureau would be a logical venue for its administration, simply because of their expertise in assessing transaction impacts on the economy and administrating competition statute and merger review in a domestic context. The Competition Bureau is also attractive for its ability to take on this new mandate with minimal incremental funding needs relative to alternatives (such as setting up an Australian-style FIRB).

Depending on government policy priorities, the scoring system can also incorporate a component for the host government human rights record. China often complains about Canada’s “lectures” on human rights and such lectures are often seen as futile and detrimental to commercial relations from a Canadian perspective. Here, Canada would not need to lecture - a suboptimal score based on the Chinese government’s policies would simply be the procedural consequence of the Chinese government’s choices. The mandatory release of the scoring, if undertaken through the auspices of the civil service, would be a bureaucratic means for the government to make a political point with fewer diplomatic consequences. Of course, this can be
viewed as the inverse of state-controlled investment being driven by considerations external to typical private business – Canada would be potentially basing its investment review on criteria with a tenuous connection to market impacts to drive its own political preferences.

It is also worth again noting that unfavourable market outcomes can serve as their own filter – the considerable struggles experienced by CNOOC with its Nexen takeover may prove to be more of a disincentive to further SOE investment than any hurdle placed by government. The largely Canadian former owners of Nexen sold at a relative value peak, making it on its face a strong commercial outcome from a liberal perspective. Again, considering the current government is seemingly China-friendly bent on trade and investment issues, market outcomes might in practice end up being the strongest available barrier.207

VII. CONCLUSION

In its assessment of the ICA NBT as applied to SOE-proposed transactions with a focus on the resource sector, this paper has evaluated the history and literature on Canadian foreign investment review as well as the present framework and its application to certain transactions. These domestic aspects have been complemented with the broader context of literature on international investment flows and SOEs, in addition to an examination of the American and Australian investment review frameworks and select transactions from those countries. This process has highlighted the many issues associated with the NBT as well as how those issues are amplified and complicated when SOE investment is involved. Namely, how the lack of (or absence of) clarity and certainty, when combined with an often arbitrary process, can serve to dissuade or delay prospective acquirers and in turn preclude Canadian assets from achieving their maximum valuation by making these assets less attractive than they would otherwise be relative to those in other jurisdictions. Where an SOE is involved, the additional procedural and practical hurdles compound the challenges of approval and increase the burdens for achieving successful review. The potential for SOEs to alter or adversely impact commercial outcomes does provide justification for enhanced scrutiny. However, the above listed issues still remain, to varying

degrees, as impediments to optimal outcomes and create a difficult review landscape.

To address these challenges, this paper has proposed replacing the current NBT with a modified test that would impose a reverse onus on the government to justify its decision to reject approval of proposed foreign investment. More crucially, this test would mandate that government engage in a quantitative scoring system with written reasons made public for all its decisions. In so doing, investment review in Canada would become more transparent, certain, and predictable, while removing the scope for arbitrariness. The unique risks of SOE investments could be controlled through more stringent scoring thresholds and through this new framework, the long troubled Canadian process of foreign investment review could instead become of net benefit to Canada.