Whole Issue (Winter 2021 / 21.2)

Follow this and additional works at: https://chicagounbound.uchicago.edu/cjil

Part of the Law Commons

Recommended Citation

This Article is brought to you for free and open access by Chicago Unbound. It has been accepted for inclusion in Chicago Journal of International Law by an authorized editor of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.
Environment, Mobility, and International Law:
A New Approach in the Americas
David James Cantor*

Abstract

The role of international law in regulating international movement in the context of global environment change and hazards remains a topic of intense debate among both legal scholars and practitioners. Yet, as this Article shows, we have largely reached the limits of what existing international law methods and approaches can tell us about the future of the law in this area. By contrast, this Article draws on a detailed regional case study to offer a distinct perspective to that ongoing debate about the role and future of international law. Against the backdrop of emerging patterns of mobility linked to devastating environmental disasters in the Americas, this Article derives new legal insights from in-depth analysis of a developing body of comparative and international legal practice by countries from across this key region.

* The author is grateful to Bruce Burson, Geoffrey Cantor, Jean-François Durieux, Walter Kälin, Ana Luquerna, and Atle Solberg for their useful comments on earlier drafts of this Article. The views expressed in the Article, and any errors that it contains, remain those of the author alone. The research on which the Article builds was generously supported by Economic and Social Research Council [grant number ES/K001051/1].
# Table of Contents

I. Introduction .................................................................................................................. 265

II. The Environment-Mobility Nexus as a Legal Problem ............................................. 268
   A. Understanding the Empirical Phenomenon ......................................................... 269
   B. Defining the Legal Problem .................................................................................. 277
   C. Framing the Case Study of the Americas ............................................................ 283

III. Empirical Dynamics in the Americas ........................................................................ 284
   A. International Movement Linked to Environmental Threats in the Americas ....... 285
   B. Other Forms of Environmental Impact on International Mobility ................. 289

IV. International Protection Law in the Americas .......................................................... 291
   A. Disasters and International Protection ................................................................. 291
   B. An Alternative Legal Approach ......................................................................... 294

V. Immigration Law in the Americas .............................................................................. 297
   A. “Ordinary” Migration Categories ....................................................................... 298
   B. “Exceptional” Migration Categories ................................................................... 302

VI. International Cooperation and Frameworks in the Americas ............................... 310
   A. Early Ad Hoc Actions ......................................................................................... 311
   B. Promoting Frameworks for Cooperation ............................................................. 312
   C. Subregional Frameworks: Approach and Scope .............................................. 315

VII. Conclusion .................................................................................................................. 320
I. INTRODUCTION

Global society appears ever more conscious of how environmental phenomena shape human mobility. The immobility enforced on populations by lockdowns in many countries as a response to the COVID-19 pandemic is only the most recent example. Yet environmental threats can also help to push the movement of persons. For example, in the context of climate change, the well-publicized risk posed to the ongoing viability of human settlement of small islands in the Pacific Ocean by rising sea levels fuels globalized concern that their populations will end up as “climate refugees.” This long-term “sinking” Pacific island scenario is but one of many scenarios where movement is shaped by environmental processes. Some reflect hazards that are more sudden-onset in character. For example, in the Americas, as recently as 2017, around 160,000 inhabitants of Puerto Rico fled to the United States mainland after the sudden devastation wrought on that island by Hurricane Maria, some temporarily and others on a more permanent basis. Indeed, a diverse range of environmental threats generates a far-reaching mobility impact on populations across the world. These are global challenges, both in the sense that few countries are immune to their effects and also in that such environmental phenomena and their consequences do not respect the territorial boundaries claimed by nation states, and they are often cross-border in nature.

The risks posed by such environmental phenomena, especially in the context of a process of global climate change, have prompted attempts by states, civil society, and other actors to coordinate international action. This includes efforts to develop appropriate structures of international law in such fields as climate change mitigation and adaptation and also disaster risk management. More recently, normative frameworks in each of these fields have begun to directly acknowledge the human mobility dimensions of these environmental phenomena. Most prominently, under the 1992 United Nations Framework Convention on Climate Change, the 2010 Cancun Agreement invites states to “enhance understanding, coordination and cooperation with regard to climate change mitigation and adaptation.”

---

1 The term “mobility” is used here to signal an emphasis on agency in movement, in other words, not only the act of movement itself but also the wider capacity to move, and to attempt to avoid importing implicit value judgments as to the voluntary or involuntary nature of such movement that are often implicit in the use of terms such as “migration” or “displacement.”

2 For a critical perspective, see Carol Farbotko & Heather Lazrus, The First Climate Refugees? Contesting Global Narratives of Climate Change in Tuvalu, 22 GLOB. ENV’T CHANGE 382 (2012).


induced displacement, migration and planned relocation.”

Like wise, in the disaster risk management field, the non-binding 2015 Sendai Framework for Disaster Risk Reduction calls on states to address “disaster-induced human mobility,” including by “transboundary cooperation.”

Up to this point, the global frameworks calling for cooperation on human mobility challenges in the context of environmental threats offer little guidance on the form that such responses should take. In this regard, a largely blank canvas appears to exist, waiting for legal development. Yet, as this Article will show, a long-standing parallel body of legal research and debate seeks to fill this apparently blank canvas. Based on a preoccupation that existing international law does not adequately protect people who leave their countries due to environmental push factors, particularly those linked to climate change, these international law studies already articulate a diverse range of innovative potential solutions to this perceived gap in the law. They are complemented by the small number of extant judicial decisions that explore how existing international law rules on refugee status and human rights protection might apply to such scenarios. Overall, this body of

---

5 A Displacement Task Force was also created under the UNFCCC Warsaw International Mechanism for Loss and Damage Associated with Climate Change Impacts in 2015. See 2010 Cancun Agreement, U.N. Doc. FCCC/CP/2015/10/Add.1. Yet, while climate-related mobility has become increasingly embedded as a topic of concern within the UNFCCC loss and damage mechanism, some suggest that its placement there may actually weaken efforts to promote climate-related mobility as a standalone issue and to develop consensus on responses, due to the particularities of that mechanism. See generally Chloé Anne Vlassopoulos, *When Climate-Induced Migration Meets Loss and Damage: A Weakening Agenda-Setting Process*, in *Research Handbook on Climate Change, Migration and the Law* 376 (Benoît Mayer & François Crépeau eds., 2017).

6 Sendai Framework for Disaster Risk Reduction 2015–2030, U.N. Doc. A/CONF/224/CRP.1, ¶ 28, ¶ 30. Yet, while subsequent policy developing this DDR framework acknowledges the number of “permanently displaced people” as a potential indicator for the “human impact” and “economic impact” of a disaster, specific guidance on measures to respond to such impact remains lacking. See, e.g., U.N. OFFICE FOR DISASTER RISK REDUCTION, WORDS INTO ACTION GUIDELINES, (2017), https://perma.cc/UTF7-JAXK.

7 See supra notes 5–6.

8 The parallel literature on the mandate and role of institutions at the international level will not be addressed here except as it bears on the question of international law development on the status of affected persons. See, e.g., Sinja Hantscher, “The UNHCR and Disaster Displacement in the 21st Century: An Organizational Analysis” (2019); Nina Hall, *Displacement, Development, and Climate Change: International Organizations Moving Beyond Their Mandates* (2016); Andrea C. Simonelli, * Governing Climate-Induced Migration and Displacement: IGO Expansion and Global Policy Implications* (2015); *Organizational Perspectives on Environmental Migration* (Kerstin Rosenow-Williams & François Gemenne eds., 2016).

9 See, for example, the national judicial decisions on how international refugee law concepts are to be interpreted in relation to claims for asylum by persons fleeing climate change or disasters, including the Supreme Court of Canada and the Supreme Court of New Zealand. See Canada (Attorney General) v. Ward, [1993] 2 S.C.R. 689; Teitiota v. Chief Executive of the Ministry of
scholarly insight and creative thinking represents a rich resource for states and other international actors as they consider how the global response to human mobility in the context of environmental threats might be further developed in the face of evidence that global warming is accelerating rapidly.

This Article contributes to this topical international law debate by offering a new perspective rooted in empirical evidence and legal practice from the region of the Americas. It starts by highlighting key features of the existing legal literature on what we might call the “environment-mobility” nexus (Section II). It shows that most legal studies adopt a particular approach, focusing on how international law, usually at the global level, could be developed to address a perceived gap in protection for people who are displaced to other countries due to environmental threats. However, it contends that we have largely reached the limits of what that methodology can tell us about the current or future role of law in this area. It suggests that studying the legal issues as they play out in practice in one specific region provides a useful complementary perspective. Moreover, as a region, the Americas offer a useful counterpoint to emerging legal scholarship with a regional focus on the sinking islands in the Pacific. Section II leads us not only to revisit certain widely held assumptions in the existing legal literature but also to reconsider the likely pathways for future development of international law in this field.

This Article continues by evaluating international mobility linked to environmental factors in the Americas to gain a sharper empirical understanding of where exactly the law might usefully act in this region (Section III). It then challenges the widely held assumption that states lack the legal tools to respond

---

10 Benoît Mayer and François Crépeau coined the idea of a “climate-mobility nexus.” That of an “environment-mobility nexus” encapsulates a similar understanding that human mobility can be shaped in many different ways, and often indirectly, by environmental factors more broadly and not just climate change. Benoît Mayer & François Crépeau, Introduction to Research Handbook on Climate Change, Migration and the Law 1 (Benoît Mayer & François Crépeau eds., 2017).

to such mobility by illustrating how pertinent provisions exist, and are used for that purpose in practice, by many states in the Americas. Crucially, such provisions are found less in refugee and human rights law on “international protection” (Section IV) than in immigration law (Section V). The analysis of how states actually approach the issue in practice is helpful in that it adds an understanding not only of where international rules may be needed but also the specific form that they might take. The Article also shows how these understandings are being actively promoted by intergovernmental bodies at the subregional level in the Americas (Section VI). On the environment-mobility nexus, the findings support the view that the international law predicament will be resolved not by producing new legal or analytical concepts but by thinking differently about existing concepts (Section VII).^{12}

II. THE ENVIRONMENT-MOBILITY NEXUS AS A LEGAL PROBLEM

Legal scholarship is increasingly preoccupied with the challenge posed to human mobility by climate change and other environmental factors. Students of international law, in particular, have led this debate, and most legal studies pursue the inquiry in terms of international law.\footnote{However, a couple of studies examine the issue in relation to the national law or policy of one country. See, e.g., Eric Omeziri & Christopher Gore, Temporary Measures: Canadian Refugee Policy and Environmental Migration, 29 Refuge 43 (2014); Chelsea Krombel, The Prospective Role of Temporary Protected Status: How Discretionary Designation Has Hinder the United States’ Ability to Protect Those Displaced by Environmental Disaster, 28 Conn. J. Int’l L. 153 (2012).} On its face, the fact that international law is at the core of this research agenda is hardly surprising. Indeed, climate change, the environment, and human mobility are all global phenomena and thus seem appropriate topics for international law. Yet many legal studies are rooted in highly particular assumptions about the nature of both the underlying empirical phenomena and the resulting legal problem. This Section illustrates these assumptions by sketching out some of the main areas of legal debate. In this regard, it does not claim to be a comprehensive survey of the burgeoning literature on this topical concern. Rather, it builds on critical review of the existing legal scholarship to elucidate where and how a case study of the region of the Americas might advance the wider legal debate in this field.

\footnote{Calum T.M. Nicholson, ‘Climate-Induced Migration’: Ways Forward in the Face of an Intrinsically Equivocal Concept, in RESEARCH HANDBOOK ON CLIMATE CHANGE, MIGRATION AND THE LAW, supra note 10, at 49, 50. In this sense, it is not a “new” challenge needing “new” responses. See Etienne Piguet, Antoine Pécout & Paul de Guichetene, Migration and Climate Change, in MIGRATION AND CLIMATE CHANGE 1, 24 (Etienne Piguet et al. eds., 2011).}
A. Understanding the Empirical Phenomenon

The underlying empirical phenomenon is described using diverse overlapping and often competing terms, each loaded with assumptions about how states should respond.\textsuperscript{14} However, “[w]e should not be distracted by semantic discussions with little practical meaning about whether to call affected persons ‘climate change refugees’, ‘environmental migrants’ or something else.”\textsuperscript{15} Rather, analyses must focus on how the broad nexus between “environment” and “mobility” is constituted empirically across a range of contexts and forms. Although a paucity of robust empirical studies on this nexus was long a cause for concern,\textsuperscript{16} the evidence base has begun to expand over the last decade or so.\textsuperscript{17} In tandem, while the superficial engagement of many legal scholars with this empirical evidence is regularly criticized,\textsuperscript{18} a growing number are now reflecting more seriously on the empirical research and its potentially far-reaching implications for understanding the role of law in this context. As a result, several important discussions about the empirical nature of the environment-mobility nexus can now be discerned as pertinent to shaping the approach and direction of legal studies.

Firstly, on the nature of the nexus between mobility and environmental factors, most legal studies frame it in terms of a causal relationship.\textsuperscript{19} Moreover, in general, these legal studies are concerned with causality in one direction only—in other words, environmental change as a cause of movement (although migration as a cause of environmental change is also considered by the social sciences).\textsuperscript{20} Likewise, the legal literature focuses squarely on adverse environmental conditions as a “push” factor prompting people to leave the country of origin. Especially in the climate change context, it frames the resulting mobility as a new challenge, although environmental adversity and change have

\textsuperscript{14} Piguet et al., supra note 12, at 17–21.
\textsuperscript{16} Dominic Kniveton, Kerstin Schmidt-Verkerk, Christopher Smith & Richard Black, Climate Change and Migration: Improving Methodologies to Estimate Flows, 32–33 INT'L ORG. FOR MIGRATION MIGRATION RESCH. SERIES I (2008).
\textsuperscript{17} Stephen Castles, Concluding Remarks on the Climate Change-Migration Nexus, in MIGRATION AND CLIMATE CHANGE, supra note 12, at 415, 419–22.
\textsuperscript{18} This critique has been advanced by many scholars within and outside the field. See, e.g., Benoît Mayer, Who Are “Climate Refugees”? Academic Engagement in the Post-truth Era, in CLIMATE REFUGEES: BEYOND THE LEGAL IMPASSE? 89, 94 (Simon Behrman & Avidan Kent eds., 2018); Richard Black, Environmental Refugees: Myth or Reality? 1 (Univ. of Sussex, Working Paper No. 34, (2001), https://perma.cc/W3J4-X9XF.
\textsuperscript{19} See generally Calum T.M. Nicholson, Climate Change and the Politics of Causal Reasoning: The Case of Climate Change and Migration, 180 GEOGRAPHICAL J. 151 (2014).
\textsuperscript{20} See, e.g., RICHARD BLACK, REFUGEES, ENVIRONMENT AND DEVELOPMENT (1998).
probably shaped human mobility throughout history. Yet this primary interest in how environmental conditions act as a “push” factor for mobility has led legal researchers to overlook other pertinent ways in which environmental change can shape the experience of human mobility, including as a “pull” factor for migrants (as for example, in places where new economic opportunities emerge as a result of certain climate change impacts).

Secondly, on the content of this causal nexus, legal scholars often adopt an “alarmist” or “maximalist” understanding of “environmental migration.” Rooted in natural sciences and security studies, this view uses deductive methods to forecast vast future waves of migration driven by environmental change. The approach posits the nexus in mono-causal terms, with environmental factors acting as the sole drivers of predicted movement. However, empirical evidence from local level studies in the social sciences instead points to the multi-causal nature of migration and shows how environmental change is often just one of many interconnected factors influencing mobility. On this basis, “skeptical” or “minimalist” scholars have argued that, empirically, environmental factors cannot be isolated as a primary driver of movement, questioning whether “environmental migration” really exists as distinct phenomenon. Others, though, attempt to transcend the divide by analyzing environmental factors as a proximate cause of movement that, even if it does interact with other factors, may produce distinct forms of mobility, as in circumstances of sudden or extreme environmental change.

21 Anthony Penna, The Human Footprint: A Global Environmental History 4–8, 56–58, 106–7 (2d ed. 2014). Indeed, environmental factors have been recognized by migration theories as early as the 1880s, although they only made a resurgence in the 1980s after references to them dwindled during much of the twentieth century. See Etienne Piguet, From “Primitive Migration” to “Climate Refugees”: The Curious Fate of the Natural Environment in Migration Studies, 103 ASS’N. AM. GEOGRAPHERS 148 (2013).


change. As the empirical evidence base grows, this approach seems to be gaining increasing acceptance.

Such debate about the multi-causal nature of migration has crucial implications for legal scholarship. Certainly, empirical evidence of the multi-causal reality of movement suggests that the legal studies that adopt a mono-causal understanding of this nexus adopt a faulty premise. This view matters because the perception of a gap in legal protection in fact emerged from the “maximalist” literature that assumes a distinct class of migrants forced to leave their homes as a result of environmental change can be identified for the purpose of intervention. However, even for those legal studies that frame environmental factors as but one proximate cause of movement in this context, the question of how to accommodate the multi-causal nature of such mobility persists. On the one hand, it poses the question of just how proximate such environmental factors need to be in order to be treated as a legally significant “cause” of movement. On the other hand, given that vulnerability to environmental threats is mediated by social, political, and economic factors, an important question also arises about the extent to which such human factors can or should be accommodated in law.

Thirdly, many legal scholars frame the “environment” side of the nexus explicitly in terms of “climate change.” For some, this is a strategic gambit to raise the profile of the issue by linking it to powerful discourses on climate change. For others, it is a matter of global justice that responsibility for resettling poor people forced out of their homes in the Global South should fall on the states in the Global North that contribute most to global warming. Yet this

27 See, e.g., Astri Suhrke, Pressure Points: Environmental Degradation, Migration and Conflict (1993); Graeme Hugo, Environmental Concerns and International Migration, 301 INT’L MIGRATION REV. 105 (1996).

28 Castles, supra note 25, at 419–24; Piguet et al., supra note 12, at 5.

29 Robert McLeman, Climate-Related Migration and its Linkages to Vulnerability, Adaptation, and Socio-Economic Inequality: Evidence from Recent Examples, in Research Handbook on Climate Change, Migration and the Law, supra note 5, at 29; Mike Hulme, Attributing Weather Extremes to “Climate Change”: A Review, 38 PROGRESS PHYSICAL GEOGRAPHY 499 (2014); Kniveton et al., supra note 16.

30 For instance, some scholars have argued that underlying processes of discrimination in the social construction of vulnerability raise the prospect that affected persons may have a claim to refugee status. See, e.g., Matthew Scott, Climate Change, Disasters, and the Refugee Convention (2020); Bruce Burson, Environmentally Induced Displacement and the 1951 Refugee Convention: Pathways to Recognition, in Environment, Forced Migration and Social Vulnerability 3 (Tamer Afifi & Jill Jäger eds., 2010).

31 Gemenne, supra note 23, at 225.

32 Mayer, supra note 22, at 43–47.

33 Maxine Burkett, Justice and Climate Migration: The Importance of Nomenclature in the Discourse on Twenty-First-Century Mobility, in Climate Refugees: Beyond the Legal Impasse?, supra note 18, at 73; Giovanni Bettini, Sarah Louise Nash & Giovanna Gioli, One Step Forward, Two Steps Back? The Fading
approach poses challenges for legal analyses. For instance, global warming seems to act on mobility by influencing more “proximate” environmental drivers, such as storms, drought, or desertification. If it is already difficult to empirically isolate the role of such proximate environmental factors in pushing migration in any specific case, then climate change adds an additional layer of complexity, as it sits one step behind those drivers (and two if the link to human activities as a cause of climate change is also to be made). Establishing the respective contribution to climate change of particular states adds a third additional layer of complexity.

Despite these challenges, even some of those scholars who recognize the “multi-causality” of migration end up proposing solutions for “climate migrants” as if they were a definite and identifiable group of persons.

Yet, even if such factual and legal determinations were possible in particular cases, this emphasis on climate change alone has other conceptual limitations. Indeed, as a “push” factor for mobility, it is not clear that the impact of climate-related phenomena, which could be influenced by global warming, differs substantially from that of other environmental phenomena, such as volcanoes or earthquakes. Moreover, even for climate related “push” factors such as storms or flooding, it is not obvious how events caused, or exacerbated, by climate change can be distinguished, in terms of their impact on human mobility, from those that are not. For these reasons, some legal studies have instead sought to frame this side of the nexus in terms of broader concepts of the “environment.” Particularly since the late 2000s, scholars and policymakers have increasingly conceptualized the environment side of the nexus more broadly in terms of disasters, an approach that obviates many of the flaws of focusing solely on “climate change.”

Contours of (In)justice in Competing Discourses on Climate Migration, 2 GEOGRAPH J. 348 (2016); François Gemenne, One Good Reason to Speak of ‘Climate Refugees’, 49 FORCED MIGRATION REV. 70 (2015); LAURA WESTRA, SATVINDER JUSS & TULLIO SCOVAZZI, TOWARDS A REFUGEE ORIENTED RIGHT OF ASYLUM (2015).

Walter Kälin, Conceptualising Climate-Induced Displacement, in CLIMATE CHANGE AND DISPLACEMENT: MULTIDISCIPLINARY PERSPECTIVES 81, 85 (Jane McAdam ed., 2010).

Moreover, even if the contribution of particular states to climate change could be characterized as internationally wrongful acts under international law, others argue that the principle of reparation in the law of state responsibility does not extend to a duty on responsible states to adopt particular policies in relation to climate migration, such as resettlement of affected individuals. See Benoit Mayer, Climate Change, Migration and the Law of State Responsibility, in RESEARCH HANDBOOK ON CLIMATE CHANGE, MIGRATION AND THE LAW, supra note 5, 238.

Mayer, supra note 18, at 97.

MAYER, supra note 22, at 12.

Id. at 26.

Id. at 12–16.

Kälin, supra note 34, at 85; MAYER, supra note 22, at 12–16 (arguing that, conceptually, “climate migration” is a component of “environmental migration” and cannot, and should not, be addressed
disasters are widely understood as encompassing both sudden- and slow-onset events and also as constituted not only by the manifestation of hazardous events but also by societal vulnerability to those hazards.\footnote{For instance, U.N. policy defines a “disaster” as “[a] serious disruption of the functioning of a community or a society at any scale due to hazardous events interacting with conditions of exposure, vulnerability, and capacity, leading to one or more of the following: human, material, economic and environmental losses and impacts.” A sudden-onset disaster is one “triggered by a hazardous event that emerges quickly or unexpectedly” while a slow-onset disaster “emerges gradually over time.” See, e.g., U.N. Gen. Assembly, Rep. of the Open-Ended Intergovernmental Expert Working Group on Indicators and Terminology Relating to Disaster Risk Reduction, U.N. Doc. A/71/644, 13 (2016) [hereinafter U.N. Report]. In short, disasters are never solely “environmental” or “natural” in character but equally reflect societal vulnerabilities to hazards that may be “natural” or “man-made.” See also IAN KELMAN, DISASTER BY CHOICE: HOW OUR ACTIONS TURN NATURAL HAZARDS INTO CATASTROPHES (2020).} On this approach, climate change remains relevant but takes a background role in causal terms, as a process that may exacerbate more immediate climate-related hazards in particular contexts.\footnote{“A changing climate leads to changes in the frequency, intensity, spatial extent, duration, and timing of extreme weather and climate events, and can result in unprecedented extreme weather and climate events.” INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE, MANAGING THE RISKS OF EXTREME EVENTS AND DISASTERS TO ADVANCE CLIMATE CHANGE ADAPTATION 7 (2012). See also Martine Rebetsz, The Main Climate Change Forecasts that Might Cause Human Displacements, in MIGRATION AND CLIMATE CHANGE, supra note 12, at 37.}

The “disaster” concept usefully foregrounds the more proximate environmental factors influencing human mobility. Yet it also raises questions. Crucially, different definitions of the disaster concept exist, despite a similar overall approach.\footnote{The different approaches reflect consensus that disasters result from the interaction between hazards and societal resilience to them but differ in other material aspects. For instance, compare the widely-endorsed U.N. definition of the “disaster” concept, supra note 41, to that developed by the Centre for Research on the Epidemiology of Disasters (CRED) for its Emergency Events Database (EM-DAT), and apparently still used by the International Federation of Red Cross and Red Crescent Societies, which defines a disaster as a “[s]ituation or event, which overwhelms local capacity, necessitating a request to national or international level for external assistance . . .; An unforeseen and often sudden event that causes great damage, destruction and human suffering.” Glossary, EM-DAT, https://perma.cc/4K6R-U8TY.} Even the widely used U.N. definition has particularities that need consideration in the mobility context. For instance, while it recognizes that a hazard need not have the potential for collective impact,\footnote{The U.N. approach defines a “hazard” as a “process, phenomenon or human activity that may cause loss of life, injury or other health impacts, property damage, social and economic disruption or environmental degradation.” U.N. Report, supra note 41, at 18.} it requires that a “hazardous event” result in a serious collective impact in order to qualify as a
“disaster.” But do people really move in response only to “disasters” or do they also move because of perceived hazards? And which concept should we favor? Moreover, each rendering of the disaster concept also differs in how it classifies different hazards in terms of both their origins and types. For our purposes, this may complicate efforts to identify which particular hazards are to be treated as “environmental” in character (and whether by reference to origins or type). Indeed, the most consistent approach may be simply to treat all of the identified hazard types as essentially “environmental.”

Lastly, some scholars have expressed concern that the “disaster”-based approach risks introduce a false binary between slow- and sudden-onset events, which might end up privileging more easily identifiable sudden-onset disasters and temporary forms of protection when more durable solutions could be required in

45 The U.N. approach defines a “hazardous event” as the “manifestation of a hazard in a particular place during a particular period of time.” U.N. Report, supra note 41, at 20.

46 The U.N. approach views the origins of hazards as, respectively, “natural, anthropogenic or socionatural” locating both environmental degradation and climate change in the last category. U.N. Report, supra note 41, at 18. By contrast, the CRED approach distinguishes between “natural” and “technological or man-made” hazards, locating environmental degradation under the latter, but treating climate change as an “aggravating factor.” See Types of Disasters: Definition of Hazard, IFRC, https://perma.cc/K5AR-VYJX.

47 Alongside “technological or man-made” hazards (that include environmental degradation and pollution), the CRED approach sub-divides the hazards of “natural” origin into geophysical, hydrological, climatological, meteorological, and biological types. As noted above, climate change is not treated as a hazard in its own right but rather an “aggravating factor.” Types of Disasters: Definition of Hazard, supra note 46. By contrast, the U.N. approach lists biological, environmental, geological, hydrometeorological, and technological types of hazard without relating them to particular origins. Environmental degradation is listed under “environmental hazards.” However, this category is qualified by the assertion that many of the processes that fall into it “may be termed drivers of hazard and risk rather than hazards in themselves, such as soil degradation, deforestation, loss of biodiversity, salinization and sea-level rise.” U.N. Report, supra note 41, at 19. A more recent U.N. document offers a still more diverse typology of hazards as geophysical, hydrological, meteorological, climatological, extra-terrestrial, environment degradation, biological, and technological. See U.N. Office for Disaster Risk Reduction, Technical Guidance for Monitoring and Reporting on Progress in Achieving the Global Targets of the Sendai Framework for Disaster Risk Reduction 172–73 (2017).

48 In other words, it is necessary to decide how a focus on environmental factors maps onto the different approaches to classifying hazards. For instance, on the U.N. approach, is it to be done by origin, in which case does the term “environmental” cover only hazards of “natural” origin or also those of “socionatural origin” (or even those of “anthropogenic” origin); or by type, in which case, does the term cover all types (“biological,” “geological,” etc.) or only some (only “environmental,” for instance, or “not technological”)?

49 The most recent U.N. approach includes geophysical, hydrological, meteorological, climatological, extra-terrestrial, environment degradation, biological, and technological hazards. U.N. Office for Disaster Risk Reduction, supra note 47. However, with reference to the origins of the hazards, the U.N. approach expressly excludes “armed conflicts and other situations of social instability or tension.” U.N. Report supra note 41, at 18.
some situations. Such criticisms foreground important questions about whether disparate types of hazardous events might impact in different ways on mobility decisions or on any resulting patterns of movement, pointing to a need for distinct kinds of legal responses.

Meanwhile, on the “human mobility” side of the nexus, legal studies tend to privilege movement with an international character. This mirrors wider public concern, which engages mainly with the cross-border aspect of climate and disaster mobility. Yet empirical evidence suggests that international movement is a less significant form of mobility in this context, in terms of numbers and vulnerability, than internal displacement or enforced immobility. Many legal studies also seem to assume that movement caused by environmental factors will be from Global South to North. Moreover, they regularly cite the predicament of “sinking” Pacific islands as the archetypal empirical problem scenario for the law and, among international environmental lawyers, resettlement of the climate-displaced to the Global North is often advanced as a solution. Much of the legal literature also seems to assume that the movement has an essentially “forced” character, reflecting its framing of environmental change as a “push” factor.

50 McCAdam, supra note 40, at 247–49; Mayer, supra note 22, at 87–89.

51 For discussion of the empirical evidence in relation to the distinct climate-related hazards of (1) storms, rains and floods, (2) droughts and desertification, and (3) sea level rise. See Piguet et al., supra note 12, at 6–12, 14–16. Indeed, the distinctions between the various forms of migration are not always neat. See generally Graeme Hugo, Lessons from Past Forced Resettlement for Climate Change Migration, in Migration and Climate Change, supra note 12, at 260.

52 Piguet et al., supra note 12, at 15. However, some of the “solutions” proposed are extended also to those internally displaced by climate change. See, e.g., David Hodgkinson, Tess Burton, Heather Anderson & Lucy Young, ‘The Hour When the Ship Comes In’: A Convention for Persons Displaced by Climate Change, 36 Monash U.L. Rev. 69 (2010); Frank Biermann & Ingrid Boas, Preparing for a Warmer World: Towards a Global Governance System to Protect Climate Refugees, 10 Glob. Envtl. Pol. 60 (2010). Most legal studies, however, tend to treat internal displacement in this context as adequately addressed by the U.N. Guiding Principles on Internal Displacement.

53 On the former, see Khalid Koser, Climate Change and Internal Displacement: Challenges to the Normative Framework, in Migration and Climate Change, supra note 12, at 289. On the latter, those most vulnerable to the effects of climate change often do not have the resources to move internationally or at all. See Dug Cabie, In-Site Adaptation: Non-Migration as a Coping Strategy for Vulnerable Persons, in Climate Change, Migration and Human Rights: Law and Policy Perspectives 99 (Dimitra Manou et al. eds., 2017).

54 Carol Farboko, Representation and Misrepresentation of Climate Migrants, in Research Handbook on Climate Change, Migration and the Law, supra note 5, at 67, 70–77; Piguet, et al., supra note 12, at 15; Gemenne, supra note 23, at 231–35.

55 Katrina M. Wyman, Ethical Duties to Climate Migrants, in Research Handbook on Climate Change, Migration and the Law, supra note 5, at 347. It is evident too in the emphasis on legal proposals for international “resettlement” of climate migrants that seem to presuppose the unavailability of internal mobility options. See, e.g., Biermann & Boas, supra note 52. This may reflect wider stereotypes about this issue. See Simonelli, supra note 8, at 23–53.

56 Gemenne, supra note 23, at 253; Piguet et al., supra note 12, at 15.
Even where the potential for “voluntary” movement is acknowledged, the main focus of legal studies remains on responding to the forced aspects of mobility in this context.\(^{57}\) Similarly, it is well recognized that we should avoid characterizing migration merely as a failure to adapt to environmental change, since movement is not only a reactive last resort but can also be a proactive adaptive coping strategy.\(^{58}\)

Finally, returning briefly to the intersection between “mobility” and “the environment,” it is important to acknowledge the recent surge of interest among scholars in how the coronavirus pandemic will shape the movement of persons globally.\(^{59}\) In tandem, many governments around the world have imposed measures to strictly limit international movement into their territories, especially by non-nationals travelling from any territory where the virus appears to have been poorly contained.\(^{60}\) On the one hand, the situation in 2020 is a stark illustration of the fact that the “environment-mobility” nexus can manifest itself in diverse forms. On the other hand, it shows that their legal implications may differ. In this regard, epidemics and pandemics, as specific kinds of biological hazard, represent something of a special case. Given that human mobility within and between states is often one of the main vectors by which the hazard is transmitted to new communities, along with the attendant risk of disaster, they raise particular sets of questions in the mobility context around the legal framework for (exceptional) measures regulating or restricting entry and free movement to minimize the transmission of infection.\(^{61}\) Since these legal issues are

\(^{57}\) See, e.g., Kälin, supra note 34, at 96.

\(^{58}\) Piguet et al., supra note 12, at 15–16; Richard Black et al., Migration as Adaptation, 478 NATURE 447, 449 (2011).

\(^{59}\) See, for example, the contributions to the Coronavirus and Mobility Forum hosted by the Centre on Migration, Policy, and Society at the University of Oxford, available at The Coronavirus and Mobility Forum, COMPAS, https://perma.cc/6QVV-VL6G.

\(^{60}\) In the U.S., for example, the President issued a proclamation in April 2020 suspending entry into its territory for certain immigrants who may present risk to the U.S. labor market’s ability to recover from economic downturns related to the COVID-19 outbreak. Proclamation No. 10,014, 85 Fed. Reg. 23,441 (Apr. 22, 2020). Globally, it is reported that “nearly all” states have imposed entry restrictions for persons travelling from territories where the virus has become widespread, with some temporarily prohibiting the entry of all non-citizens and non-residents. Moreover, “nearly all” states have introduced additional health screening procedures at ports of entry, with “most” requiring travelers from affected territories to be quarantined for a period of time on entry. Most countries also advise their nationals against non-essential international travel or to specific jurisdictions where the outbreak is more severe. See Immigration Update: Coronavirus, FRAGOMEN NEWS (Sept. 21, 2020), https://perma.cc/2S49-L6LA.

important in and of themselves, and they are separate from those relating generally to the entry and stay of people affected by other kinds of hazards, they deserve study in their own right and will not be addressed further here.

B. Defining the Legal Problem

Legal debate on the environment-mobility nexus is underpinned by certain assumptions about not only the nature of the underlying empirical situation but also the framing of the legal problem. This debate assumes the essential legal problem to be that the law does not adequately regulate the situation of persons who leave their country due to environmental factors, especially climate change. Implicit in that statement are empirical assumptions about which parts of the environment-mobility nexus are important for legal regulation, as outlined above. But the way that legal scholarship addresses this perceived gap in the law also reflects certain legal assumptions about how that gap in the law is itself constituted and, ultimately, resolved. Those assumptions serve both to channel the resulting legal debate in particular directions and to eclipse other productive lines of inquiry. By elucidating these underpinning premises, we can better understand where and how a case study of the Americas might contribute to debate on legal responses to the environment-mobility nexus.

Firstly, legal scholarship reflects an international law standpoint. Indeed, in essence, this is a debate about international law. This focus on international law in particular is hardly surprising, since both migration and the environment are intuitively global phenomena. Naturally, it seems to follow that international mobility due to environmental drivers, as a global problem, calls for an international legal response. Yet this conception of the legal problem as inherently one of international law shapes the resulting analyses. As the following discussion will show, not only is the problem framed as a gap in international law, but also solutions to this problem are both located within international law and built from existing international law.62 Although international law at the global level is the focus of most legal studies, growing numbers of legal scholars now argue that new norms of international law are more likely to be developed at the regional or even bilateral level, at least in the first instance.63 Cooperation of this kind at the regional

---

62 See, e.g., sources cited notes 83–90.
63 Platform on Disaster Displacement, State-led, Regional Consultative Processes: Opportunities to Develop Legal Frameworks on Disaster Displacement, in CLIMATE REFUGEES: BEYOND THE LEGAL IMPASSE? 126
level is seen as attractive to states since most migration is intra-regional in nature already, regions are likely to face similar kinds of environmental processes, and regional forms of international cooperation are already the status quo in most parts of the world.\textsuperscript{64}

By contrast, law at the national level is seen as largely irrelevant by the legal literature. Even scholars who assess the few national law provisions on environmental displacement ultimately dismiss them as “ad hoc” and “inadequate,”\textsuperscript{65} “unpredictable” in terms of application and status,\textsuperscript{66} and full of “vague language.”\textsuperscript{67} Such national law is further characterized as “inconsistent” and “varying from one country to another.”\textsuperscript{68} It is also said to lack “legal certainty,” as it is “not rooted in existing legal duties” but relies on “discretion rather than legal obligation.”\textsuperscript{69} Of course, many of these complaints about vague language, inconsistency, and so on appear overstated since they could be leveled equally at international law. Likewise, the notion that national law cannot create legal rules and duties for the state concerned is simply incorrect. Moreover, it is notable that most scholars simply cite the same few protection provisions of national law from states in the Global North.\textsuperscript{70} As a result, national law from states


\textsuperscript{64}Popp, supra note 63, at 230; see also Black et al., supra note 58, at 449.

\textsuperscript{65}Christel Cournil, \textit{The Protection of “Environmental Refugees” in International Law, in MIGRATION AND CLIMATE CHANGE} 359, 369–70 (Étienne Piguet et al. eds., 2011); Thelki Anastasiou, \textit{Public International Law’s Applicability to Migration as Adaptation, in CLIMATE REFUGEES: BEYOND THE LEGAL IMPASSE} 172, 183–84 (Simon Behrman & Avidan Kent eds., 2018).

\textsuperscript{66}MCADAM, supra note 40, at 117.

\textsuperscript{67}Anastasiou, supra note 65, at 183–84.

\textsuperscript{68}MCADAM, supra note 40, at 117; Kälin, supra note 34, at 100.

\textsuperscript{69}Susan F. Martin, \textit{Towards an Extension of Complementary Protection?, in RESEARCH HANDBOOK ON CLIMATE CHANGE, MIGRATION AND THE LAW, supra note 5, at 449, 449–50; Anastasiou, supra note 65, at 183–84.}

in the Global South is largely absent from the analysis.  

Furthermore, the focus in these analyses on “international protection” provisions means that the wider provisions of national immigration law are also largely overlooked.  

This is a direct consequence of setting the legal debate so firmly within international law parameters: in contrast to the law on international protection, immigration law is not yet well established as a distinct body of international law.

Secondly, the perception of a gap in international law is the starting point for most legal studies. It is clear that, in general, persons displaced across borders by environmental factors do not benefit from international legal guarantees relating to “refugees” (or those on “migrant workers”).  

Certainly, the extant treaties dealing, respectively, with refugees, statelessness, human rights, or the environment do not specifically address this situation.  

As will be discussed below, many legal scholars seem to take this fact as sufficient evidence of a legal gap in relation to the “protection” of such persons (and thus, by extension, of a gap in relation to their envisaged need for “resettlement” to the Global North).  

However, some scholars argue for a narrower gap that exists only in respect of certain specific needs that are not covered by international human rights law, which continues to benefit such persons.  

In particular, they argue that the general gap in legal protection actually exists mainly in relation to the “legal status” of these persons, to aspects of their “admission [and] continued stay [in the reception country], and [to] protection against forcible return to their country of origin.”

---

71 This may partly reflect the perception that the Global North will be the recipient of arrivals in this context. As an exception, see the few counter examples from Africa and one from Argentina cited in passing by McAdam, supra note 40, at 105, 107.

72 For exceptions, see text at note 96 below. “International protection” refers to the assumption of an obligation to provide protection to people who are outside their own country and face certain specified kinds of persecution or harm in that country but lack the protection of their own state.


74 See supra note 73; see also McAdam, supra note 40, at 39–98. McAdam equally dismisses the much-debated prospect of the international law on statelessness resolving the situation of “sinking” small island states. See McAdam, supra note 40, at 119–60.

75 See Kälin, supra note 34, at 87–89. See also Siobhán McInerney-Lankford, Climate Change, Human Rights and Migration: A Legal Analysis of Challenges and Opportunities, in Research Handbook On Climate Change, Migration AND THE LAW, supra note 5, at 131, 131.

76 Platform on Disaster Displacement, supra note 63, at 145; Kälin, supra note 34, at 89. The recent comments of the UNHRC in Teotia Teotia v. New Zealand, U.N. Doc. CCPR/C/127/D/2728/2016, ¶¶ 9.11–14 (2020) suggest that, in principle, the effects of climate change (and possibly other forms of environmental degradation) in an applicant’s country of origin could generate a sufficient threat to the right to life to prevent refoulement on human rights grounds, although that threat would have to be highly imminent. It would also not require states to grant admission or stay.
Among legal scholars, the gap in legal international standards also tends to be perceived in terms of an absence of “international protection” for persons who flee their countries due solely to the impact of environmental factors. This analogy with the situation of refugees, and other beneficiaries of international protection, is evident from the literature’s principal concern with persons unwillingly outside their country, recalling the “‘exilic’ bias” of refugee law. Some scholars even explicitly frame the legal problem in this context as an absence of international protection for forced movements, with “voluntary migration” simply left to the discretion of states in national law. Yet those scholars not only underestimate the difficulty of distinguishing “forced” and “voluntary” movement in this context but also misrepresent the logic of international protection, which turns on prospective risk in the country of origin and a lack of national protection rather than the supposedly forced quality of movement. Even so, they show that, analogous to the situation of refugees, the legal gap in relation to mobility on environmental grounds is conceived principally as one of “international protection” under international law.

Thirdly, the legal debate is “not about the law as it exists (lex lata), but about what the law ought to be (lex ferenda).” In other words, the problem is largely accepted, and the debate is really about solutions. In most cases, legal scholars turn to international law to close this legal gap. Two main methods are evident. On the one hand, certain scholars, particularly those from the international refugee and human rights law fields, argue in favor of more expansively interpreting existing norms of international law. Many of them advocate for interpreting international refugee definitions broadly to properly take account of how “human” inputs also shape “natural” disasters in any particular society.

See generally McAdam, supra note 40, at 98; McGregor, supra note 25.

See supra note 30 and accompanying text. This approach can be discerned in Scott, supra note 30; Sanjula Weerasinghe, UNHCR Legal and Protection Policy Research Series, No. 39: In Harm’s Way: International Protection in the Context of Nexus Dynamics Between Conflict or Violence and Disaster or Climate Change 109–10 (2018); Madeline Garlick, Marine Franck & Erica Bower, Enhancing Legal Protection for People Displaced in the Context of Disasters and Climate Change, in Climate Refugees: Beyond the Legal Impasse, supra note 18, at 118, 121; Selwyn Fraser, Climate Persecutors: Climate Change Displacement and the International Community as
Some also propose the development of soft law instruments to provide temporary or similar protection to the broader class of persons fleeing environmental factors. Such proposals, often touted as merely a first step on the path to creating a new treaty dedicated to this challenge, also often creatively draw on, and develop, existing international law principles from the field of international protection law.

On the other hand, an alternative approach, more common among international environmental law scholars, proposes new treaty law to fill the legal gap. A few argue for amending the terms of existing treaties in international environmental law, for which they provide draft proposals. Many others call for a new treaty, either standalone or under the framework of international law. Some also propose the development of soft law instruments to provide temporary or similar protection to the broad


85 Several of these proposals focus on the provision of “temporary protection” for people who flee disasters or climate change but do not qualify as refugees. See, e.g., Garlick et al., supra note 83, at 121–22; Volker Türk, Temporary Protection Arrangements to Fill a Gap in the Protection Regime, 49 FORCED MIGRATION REV. 40, 40–41 (2015); Wood, supra note 84, at 23–25; MCADAM, supra note 40, at 256–66.


core, though, these proposals seek to promote status-based forms of international protection for their respective classes of refugee-like beneficiaries, whether they are defined in the draft instruments as, variably, “environmentally displaced persons,” “climate migrants,” “persons displaced by climate change,” or “climate refugees.” Likewise, many of the new obligations that these proposals envisage in areas such as resettlement and distribution serve to reproduce or develop existing legal principles drawn from the law of international protection or from international environmental law. None of the proposals have yet been taken up by states.

Finally, as no clear ethical basis exists for privileging environmental factors over drivers of migration such as poverty, there is debate over whether such special protection can be justified. Rather than arbitrarily creating new regimes for a privileged few, some suggest that we should instead focus on fully promoting the basic human rights of all migrants without distinction. Similarly, recognition of the way that environmental and human factors intertwine to shape vulnerability and mobility leads some to argue that a focus on protecting the displaced misses the bigger picture “that such migration is a consequence of the human insecurity imposed on the South in the current global order.” These approaches suggest that mobility in this context cannot be addressed in isolation from the pressing need to respond to wider migration, environmental, and development challenges and their impact on countries that are particularly exposed to the risk of disaster. For instance, certain scholars working in the Pacific have begun to analyze climate mobility in the context of wider migration patterns and processes. As a result,

89 See Christel Cournil, The Protection of “Environmental Refugees” in International Law, in Migration and Climate Change, supra note 12, at 359, 361–63 (listing scholarly proposals identified as sharing a concept of “protection” as their common basis).


91 The currently limited extent of formal state interest in creating a new treaty can be evidenced from the few examples cited in Prieur, supra note 40, at 237. See also McArdle, supra note 40, at 187–201.

92 See Mayer, supra note 22, at 31–35; Peter Penz, International Ethical Responsibilities to “Climate Change Refugees,” in Climate Change and Displacement: Multidisciplinary Perspectives, supra note 34, at 151, 152–54.

93 See Mayer, supra note 22, at 159–85.

94 Stephen Castles, Concluding Remarks on the Climate Change-Migration Nexus, in Migration and Climate Change, supra note 12, at 415, 425.

95 Id. at 424–26; see also Mayer, supra note 22, at 16–35.

96 See, e.g., Bruce Burson & Richard Bedford, Nansen Initiative, Clusters and Hubs: Toward a Regional Architecture for Voluntary Adaptive Migration in the Pacific (2013), https://perma.cc/6T1L-83ZQ; Jon Barnett & Michael Webber, Migration as Adaptation: Opportunities and Limits, in Climate Change and Displacement: Multidisciplinary
they now argue that climate migration might be addressed “within existing international migration mechanisms,”87 and ask how immigration law in Australia and New Zealand could be tweaked, or how bilateral or regional arrangements may be developed, to accommodate future mobility linked to climate change.88

C. Framing the Case Study of the Americas

Legal debate on the environment-mobility nexus revolves principally around the question of how to respond to international mobility shaped by environmental factors. As such, this Article aims to contribute to that core legal debate rather than to consider the legal implications of other aspects of this nexus, such as internal mobility linked to environmental factors.89 Nonetheless, this Article draws on insights from the preceding literature review as points of entry into the legal debate. Firstly, on the “mobility” side of the nexus, it focuses on travel, entry, and stay for nonnationals, or “aliens,” as the key challenge. Secondly, on the “environment” side of the nexus, it focuses broadly on disasters and the underlying hazards rather than limiting the analysis to climate change alone. Thirdly, on causality, it addresses not only how such environmental factors contribute to displacement but also how they might impact international mobility in other ways. These insights provide a strong foundation for renewed consideration of key areas of the legal debate, such as the scope of existing legal protection, the nature of potential legal development in this field, and how to accommodate the multicausal nature of migration.

This Article interrogates these questions through a case study of the Americas. On the one hand, this approach reflects the contention that abstract analysis of international law at the global level has largely reached the limits of

87 Burson & Bedford, supra note 96, at 10; McAdam, supra note 40, at 201–11; Graeme Hugo, Climate Change-Induced Mobility and the Existing Migration Regime in Asia and the Pacific, in Climate Change and Displacement: Multidisciplinary Perspectives, supra note 34, at 9, 33; see also Jon Barnett & Natasha Chamberlain, Migration as Climate Change Adaptation: Implications for the Pacific, in Climate Change and Migration: South Pacific Perspectives, supra note 63, at 51, 56; Richard Bedford & Charlotte Bedford, International Migration and Climate Change: A Post-Copenhagen Perspective on Options for Kiribati and Tuvalu, in Climate Change and Migration: South Pacific Perspectives, supra note 63, at 89, 125.

88 See supra note 97 and accompanying text.

89 The existing human rights-based framework codified in the U.N. Guiding Principles on Internal Displacement is usually seen as sufficient to address the situation of people forcibly displaced by climate change within their own country. See Kälin, supra note 34, at 93–94. However, elaboration of the norms may be needed in relation to durable solutions and accountability for climate change drivers. See Elizabeth Ferris, The Relevance of the Guiding Principles on Internal Displacement for the Climate Change-Migration Nexus, in Research Handbook on Climate Change, Migration and The Law, supra note 5, at 108, 119.
what it can contribute to advancing these kinds of legal debates. On the other hand, it takes seriously the observation by some scholars that international mobility linked to environmental factors—as well as the development of legal responses and cooperation by states—is most likely to play out within particular regions rather than at the global level, at least initially. As a case study, the Americas offer a contrasting example to the oft-cited Pacific case. Certainly, like those in the Pacific, countries in the Americas are exposed regularly to sundry hazards. Yet, in other ways, the Americas are more diverse, comprising two continents with extensive land borders, in contrast to the small island states that make up most of the Pacific. Moreover, the Americas are twenty times more populous than the Pacific and contain not only some of the world’s largest and richest countries but also some of its poorest, as well as many others located in between these two extremes.

III. EMPIRICAL DYNAMICS IN THE AMERICAS

The gaps in protection identified by legal scholars writing on the environment-mobility nexus exist only in relation to the presumed reality of international movement caused by environmental threats in the country of origin. In other words, the legal problem corresponds to an assumed underlying empirical phenomenon. Yet many legal scholars seem merely to rely on vague and poorly-evidenced, even rather speculative, assertions about this phenomenon’s existence, scope, and characteristics, often citing disjointed and rather particular examples as if they demonstrated some general trend. However, if we want to truly assess the adequacy of the law in relation to specific empirical phenomena, such as the international movement of persons in the context of environmental push factors, then we need to engage more robustly with the growing body of natural and social science research on this topic. Toward this end, the present Article seeks to derive a more precise understanding of the environment-mobility nexus in the region of the Americas from the somewhat fragmentary research that exists thus far. From this point of entry, this Article elucidates a few of the main ways in which disasters and hazards appear, firstly, to act as a push factor for diverse dynamics of international movement in the Americas and, secondly, to shape the

100 See supra note 97 and accompanying text.
102 See, for example, the empirical “scene-setting” in Simon Behrman & Avidan Kent, Overcoming the Legal Impasse? Setting the Scene, in CLIMATE REFUGEES: BEYOND THE LEGAL IMPASSE?, supra note 18, at 3–6.
103 Calls for more robust engagement with the empirical evidence are made by some legal scholars. See, e.g., Mayer, supra note 18, at 90–91.
experience of international mobility in other ways that are important for the law to consider.

A. International Movement Linked to Environmental Threats in the Americas

Empirical research confirms that the diverse environmental threats to which countries in the Americas are exposed can act as a push factor for movement. Evidence of this impact exists for both sudden-onset events, such as storms, hurricanes, floods, and earthquakes, and slow-onset events, such as drought, erosion, desertification, and glacier retreat that may be linked to climate change. Social scientists have suggested that these different kinds of hazards produce distinct patterns of mobility in terms of duration, distance, and character, although the evidence remains somewhat mixed. Even so, the data is clear that sudden- and slow-onset disasters now push millions of incidents of internal movement by individuals in the countries of the Americas each year. By contrast, data on international movement due to disasters are more fragmentary and not routinely collected. Yet, as the following discussion shows, international movement linked to both slow- and sudden-onset events in this region is a present reality and not just an abstract legal concern, even if its scale seems less significant than that of internal movement. Of course, given that the latter is predicted to increase with time, so may the former. These trends in the Americas accord with those in other regions of the world, suggesting that this region is not an anomaly in that respect.

104 See generally Raoul Kaenzig & Etienne Piguet, Migration and Climate Change in Latin America and the Caribbean, in PEOPLE ON THE MOVE IN A CHANGING CLIMATE supra note 63, at 155. There are also areas of Suriname, Guyana, and the Bahamas where the impact of sea-level rises on economic livelihoods could force migration in the future. Id. at 169.

105 Id. at 155; Piguet et al., supra note 12, at 6–12, 15–16.

106 Calculations by the author based on figures in INTERNAL DISPLACEMENT MONITORING CENTRE (IDMC), ANNUAL REPORTS 2019 (2019), https://perma.cc/AZ7S-3W7B, suggest that, between 2008 and 2018, over 28.5 million reported instances of internal movement due to sudden-onset disasters linked to natural hazards occurred in the Americas region. Across the same time period, the figures for internal displacements by sub-region are: Central America (798,472); Caribbean (6,705,000); North America (9,851,300); South America (11,184,180).

107 For instance, for the U.S. as a key destination country, approximate calculations by the author of instances of immigration linked to disasters from other countries in the Americas, including those granted entry or stay under normal immigration categories—drawing on Onelica C. Andrade Afonso, Natural Disasters and Migration: Storms in Central America and the Caribbean and Immigration to the U.S., 14 EXPL., 1, 10 (2011)—and under temporary protected status in disaster contexts, suggest an average annual upper ceiling in the tens of thousands.

108 See WORLD BANK GROUP [WBO], GROUNDSWELL: PREPARING FOR INTERNAL CLIMATE MIGRATION; POLICY NOTE #3: INTERNAL CLIMATE MIGRATION IN LATIN AMERICA 1 (2018), https://perma.cc/9Y5L-ASPE.

109 See Kaenzig & Piguet, supra note 104, at 171; Piguet et al., supra note 12, at 6–12, 15–16.
As in other regions, such mobility is shaped by multiple, intersecting drivers, with environmental pressures often just one more push factor in contexts sometimes riven by deep inequality.\footnote{See Kaenzig & Piguet, supra note 104, at 171.} Even so, empirical evidence from the Americas suggests that at least three different strands of international movement pushed by environmental factors can be discerned. The most visible form of movement takes place shortly before or after a sudden-onset hazardous event is perceived as approaching, as people living near land borders may temporarily cross into the neighboring country to escape the impact of the event or to access better shelter or aid on the other side of the border. The movement usually follows existing, well-established patterns of daily back-and-forth migratory crossings in border regions. Examples include north Guatemalans crossing the border into Mexico to better weather tropical storms, victims of flooding seeking respite by crossing from south Colombia to Ecuador or from Amazonian Bolivia and Peru to Brazil, and Chileans affected by earthquakes or mudslides in frontier zones that are cut off from other parts of Chile seeking aid in accessible Argentinian towns.\footnote{See Cantor, Cross-Border Displacement, supra note 11, at 17; Cantor, Law, Policy and Practice, supra note 11, at 12.} Crucially, most people move temporarily to escape not only actual disasters but also perceived oncoming disasters or hazardous events.

Another strand of international movement in the Americas consists of those persons who leave their countries in the context of slow-onset disasters. The data shows that these persons, who are often from populations or social sectors whose livelihood depends on particular forms of agriculture, also tend to follow existing migration routes out of the country. For example, severe droughts linked to climatic factors are shown to increase migration from affected parts of rural Mexico to the U.S.\footnote{See Isabelle Chott & Maëlys de la Rupelle, Determinants of Mexico-US Outward and Return Migration Flows: A State-Level Panel Data Analysis, 53 Demography 1453, 1474 (2016); Raphael J. Nawrotzki, Fernando Riosmena & Lori M. Hunter, Do Rainfall Deficits Predict U.S.-Bound Migration from Rural Mexico? Evidence from the Mexican Census, 32 Population Res. & Pol'y Rev. 129, 144–47 (2013); The National Heritage Institute, Environmental Degradation and Migration: The U.S.-Mexico Case Study 25 (1997). One study suggests that climate-related international migration from rural Mexico is predominantly undocumented. See Raphael J Nawrotzki, Fernando Riosmena, Lori M. Hunter & Daniel M. Runfola, Undocumented Migration in response to Climate Change, 1 Int’l J. Population Stud. 60, 67 (2015). Some suggest that temporary migration of up to two years is the predominant form of international migration by those affected by drought and desertification, rather than permanent migration. See Michelle Leighton, Drought, Desertification and Migration: Past Experiences, Predicted Impacts and Human Rights Issues, in Migration and Climate Change, supra note 12, at 331, 349.} Similarly, slow-onset events linked to changing weather and rainfall patterns, soil erosion, and other environmental degradation appear to have helped push migration from rural parts of the Dominican Republic and Haiti,
sometimes to other countries. Given that the impact of such slow-onset disasters is often mediated via social factors to a greater extent than for sudden-onset events, their role in driving mobility can be highly contextual. However, it is not always possible to differentiate the respective contribution of slow- and sudden-onset events to pushing movement, especially in locations where they overlap. For instance, research in some rural areas of Honduras and Haiti shows how migration out of the country from those areas is driven by spiraling livelihood pressures resulting from the combined impact of slow-onset environmental degradation with sudden-onset tropical storms.

Similar questions about how to frame the impact of disasters and hazards as a push factor for mobility emerge in evidence of a third form of international movement in the Americas that takes place up to a year or more after a sudden-onset disaster has occurred. This “delayed” movement seems to be driven not so much by the hazard’s sudden and immediate disaster impact as by its enduring implications for the viability of long-term household livelihood plans. It is documented mainly in poorer and less-resilient Central American and Caribbean countries where a tropical storm or earthquake has had a particularly devastating effect on society and infrastructure at the national level. That data dovetails with other research showing that regular migration to the U.S. increases after severe storms in these countries, and also in Mexico, albeit only after a lag period of up to a year. Like the other two strands of movement, this one also tends to follow

---


114 See Piguet et al., supra note 12, at 8–12.


116 This lag may reflect both the diminishing access to humanitarian aid in the disaster-affected country as the months pass and the time needed for households to collect the resources for travel. See Cantor, Law, Policy and Practice, supra note 11, at 12–13. Alternatively, if people are able to rebuild homes and replant crops during the initial recovery, households or household members may then migrate to seek alternative income sources. See McLeman, supra note 29, at 43–44.


traditional migration routes from the affected country. However, where the routes are blocked, then it seems that new ones are forged, as with the new patterns of Haitian mobility that reoriented toward South American countries when some traditional Haitian migration destination countries such as the U.S. tried to close their borders after the 2010 earthquake.

This growing evidence base shows that environmental phenomena in the Americas do contribute to pushing diverse forms of international movement. Indeed, the three strands of movement identified here likely provide only a few pertinent examples of how the wider mobility dynamics play out. Certainly, it seems that only in rare cases will these forms of movement be likely to satisfy the long-established legal bases for international protection by states. Equally, though, they do suggest that, while framing the empirical problem in terms of “disasters” offers a useful point of entry for understanding how environmental factors influence human mobility, a limitation of the “disaster” concept is that it describes only one way in which hazards can act as drivers of mobility. In this regard, the examples imply that people do not leave only due to the occurrence or risk of disasters at the societal level. Rather, some movement also occurs preemptively due to the perceived potential impact of a hazard at the individual or household level, regardless of whether its collective impact at the societal level will result in a “disaster.” Likewise, other patterns of movement occur after the “disaster” phase has passed due to the hazard’s perceived longer-term or ongoing impact on the viability of individual or household livelihood strategies. In other words, while a hazardous event is a prerequisite for a “disaster,” the perceived or actual impact of a hazard or hazardous event at the household level can be sufficient to drive movement by the affected people, even in the absence of disaster conditions at the societal level.

These three strands of movement also suggest that any analytical distinction between sudden- and slow-onset disasters may prove less relevant for our purposes than the recognition that hazardous events and disasters can have both

119 Social scientists have observed that international movements in this context tend to occur most frequently where pre-existing relationships of migration exist between the sending and receiving countries. See, e.g., Kaenzig & Piguet, supra note 104, at 171.


121 For instance, other scenarios might include movement away from communities that are exposed to the repeated impact of sudden-onset events.

122 See MCADAM, supra note 40, at 52–98. See Section IV below for discussion of why the legal criteria for international protection are likely to be engaged only rarely by these circumstances.

123 On the U.N. approach, even the concept of “disaster risk” is defined in terms of potential impact at the collective level on “a system, society or a community.” See U.N. Report, supra note 41, at 14.
short-term and long-term impacts on mobility at the household level, even if their relative proximity as push factors for mobility may diminish over the long-term. In general, then, these conclusions point to a need for researchers to engage with the broader ways in which hazards impact human mobility, and to recognize that the occurrence of a “disaster” is but one way in which those hazards can shape movement. But what does this mean for the law? Certainly, the “disaster” concept was not designed to be applied as a legal basis for regulating movement. At the same time, it has definite advantages over concepts such as “climate change” or “the environment” that suggest its application to this legal context merits consideration. As a potential basis for regulating entry and stay by non-nationals, though, the empirical evidence suggests that lawmakers will need to reflect carefully on whether to use the concept of “disasters” strictly as a threshold that requires that a hazard has an impact at the collective level in the affected country or, instead, to advance a more granular approach to the wider ways in which “hazardous events” or the underlying “hazards” impact mobility options at the individual or household level.

B. Other Forms of Environmental Impact on International Mobility

The legal literature is principally concerned with environmental factors as a driver of international mobility or “push factor.” It is recognized that this causality can play out in diverse scenarios—for instance, as a result of sudden-onset disasters, slow-onset disasters, the impact of climate on conflict over natural resources, etc.—but the emphasis remains on how such phenomena act as drivers of movement by affected persons. 124 More recently, though, some legal scholars have argued in favor of a broader conception of this causal nexus by pointing to the possibility that environmental factors might also shape mobility by acting as a “pull factor” due to the new opportunities created by climate change or by mitigation or adaptation activities. 125 Building on those analyses, the present Article contends that engagement with the empirical evidence from the Americas discloses still other ways in which disasters may shape the phenomenon of international mobility. Moreover, these further configurations of the environment-mobility nexus pose particularly acute questions for the law, especially in relation to travel, entry, and stay for affected persons.

Certainly, the evidence from the Americas confirms that not all international mobility in the context of environmental threats will take the form of a spontaneous movement by affected persons. In the Pacific region, scholars have long debated the prospect of inhabitants of “sinking” small island states being

---

124 See, e.g., Kälin, supra note 34, at 84–92.
125 See Mayer, supra note 22, at 22–25.
relocated to other countries and the legal implications of such measures.\textsuperscript{126} However, data from the Americas verifies that organized transfers of disaster-affected persons by states already take place in the form of evacuations carried out before or shortly after a sudden-onset disaster.\textsuperscript{127} Such evacuations are often undertaken by a foreign state for its nationals unfortunate enough to be caught up in a serious disaster overseas. In the Americas, however, empirical data shows that nationals of the disaster-affected country have sometimes also been evacuated to other countries by those states. Examples include the evacuation to other Caribbean islands as well as to the U.K. and U.S. of most of the population of Montserrat when the volcano erupted in 1995 and certain profiles of Haitian nationals evacuated on medical or similar grounds by Canada, Mexico, and the U.S. after the 2010 Haiti earthquake.\textsuperscript{128} This raises the question of how the law treats such organized transfers in terms of travel, entry, and stay.

Crucially, studies of the Americas region show that disasters can impact a range of other legal aspects of international mobility for aliens.\textsuperscript{129} Thus, where a disaster occurs in the country of origin, it may limit possibilities for return, whether voluntary or enforced, with particular legal implications in terms of removal and stay for nationals of that country. It may also reduce the flow of resources from family or businesses in the home country that are needed for the alien to support maintenance during studies or other lawful forms of stay in the host country. Likewise, where a disaster occurs in the host country, it may impede the alien’s basis for stay as a result of the death of the family member on whom legal status depends, the destruction of the business that provides the basis for a work permit, or an inability to comply with immigration reporting restrictions due to damage to transport and communication infrastructures in the host country. The disaster may also reduce or interrupt the capacity of immigration authorities in the host country to process applications from aliens for travel, entry, and stay. Overall, these scenarios suggest that the law needs to take a broader conception of the environment-mobility nexus if it is to adequately regulate international mobility in the context of environmental threats.


\textsuperscript{127} Evacuations can also involve internal movement, as with the 2017 precautionary wholesale mandatory evacuation by Antigua and Barbuda of the island of Barbuda and by the Bahamas of its southern islands. See Kate Lyons, \textit{The Night Barbuda Died: How Hurricane Irma Created a Caribbean Ghost Town}, \textit{The Guardian} (Nov. 20, 2017), https://perma.cc/A4RN-SPNF; Bahamas to Evacuate Islands in Path of “Irma”, \textit{Associated Press} (Sept. 6, 2017), https://perma.cc/PSH7-ZCTZ.

\textsuperscript{128} \textsc{Cantor, Law, Policy and Practice}, supra note 11, at 13.

\textsuperscript{129} See id. at 13–14 for these examples.
IV. INTERNATIONAL PROTECTION LAW IN THE AMERICAS

International law scholarship treats international mobility in the context of environmental pressures as a new challenge that the existing law does not yet adequately address. Moreover, most contributors to this debate frame that legal gap, and its solution, principally in terms of international protection for affected persons. At first glance, recent practice in the Americas appears to confirm this point. Certainly, states rarely extend international protection to persons fleeing environmental threats, despite claims by scholars as to the relevance of certain regional legal instruments. Yet a detailed analysis of legal practice in this region offers a more nuanced understanding. Crucially, this Section will show that certain states have actually long recognized the challenge of international mobility caused by disasters and, initially at least, some dealt with it as a matter of refugee protection. Further, although this international protection approach waned as states in the Americas increasingly became integrated into the global refugee law regime, the underlying legal challenge was not discounted. Rather, a distinct new legal approach toward the entry and stay of persons affected by a disaster can be discerned in some of the national laws on refugees and international protection adopted by states in the Americas.

A. Disasters and International Protection

At present, most states in the Americas are parties to the main binding U.N. treaties on refugee protection and have incorporated pertinent aspects of the “universal” refugee definition into national law.130 In Latin America, fifteen states have also incorporated the regional expanded refugee definition endorsed by the non-binding Cartagena Declaration into their national law.131 In a small number

---

130 These U.N. treaties define a “refugee” positively as any person who “owing to well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country.” See Convention relating to the Status of Refugees art. 1A(2), July 28, 1951, 189 U.N.T.S. 150 (entered into force Apr. 22, 1954); Protocol relating to the Status of Refugees art. 1(2), Jan. 31, 1967, 19 U.S.T. 6223, 606 U.N.T.S. 267 (entered into force Oct. 4, 1967). The Caribbean is the exception in this region: only eight of thirteen states are parties to the Protocol and only four have incorporated the refugee definition into national law and policy. See CANTOR, CROSS-BORDER DISPLACEMENT, supra note 11, at 64.

131 The Cartagena Declaration defines “refugees” also as “persons who have fled their country because their lives, safety or freedom have been threatened by generalized violence, foreign aggression, internal conflicts, massive violation of human rights or other circumstances which have seriously disturbed public order.” Cartagena Declaration on Refugees § 3(3), Nov. 1984, reprinted in 3 UNHCR, COLLECTION OF INTERNATIONAL INSTRUMENTS AND OTHER LEGAL TEXTS CONCERNING REFUGEES AND DISPLACED PERSONS: REGIONAL INSTRUMENTS 1196, 1197 (2007). Argentina, Belize, Bolivia, Brazil, Chile, Colombia, Ecuador, El Salvador, Guatemala, Honduras, Nicaragua, Paraguay, Peru, and Uruguay incorporate the expanded definition in national legislation.
of states, national law also provides for complementary forms of international protection based on non-refoulement standards in international human rights treaties. None of these international instruments specifically mention the occurrence or risk of a “disaster” in the country of origin as a basis for international protection. Yet some legal scholars argue that these universal and regional treaty norms could be interpreted to provide international protection to persons fleeing disasters. Some limited practice exists in support of this proposal. For example, after the 2010 earthquake in Haiti, several Latin American states did recognize a small number of Haitians as refugees due to the violence unleashed by the disaster. Similarly, the French territories of the Antilles and Guiana granted “subsidiary” forms of complementary international protection to some Haitian asylum seekers in light of the security and other risks generated by the earthquake.

Crucially, though, in each case where states in the Americas did grant international protection, whether under refugee law or complementary forms of international protection, this was conferred due to the breakdown in the institutions of national protection in Haiti and associated risks of violence generated by the disaster. The fact that such violence and the lack of national protection resulted from a disaster, as opposed to some other cause, was thus treated as legally irrelevant for the purposes of determining international


Non-refoulement standards forbid the sending of a person to another country where they may face specific kinds of persecution or other serious harms. Alongside the well-established rules on the non-refoulement of refugees in refugee law, international human rights law prohibitions on sending a person to a territory where the risk of torture exists form the main source of “complementary” protection (in other words complementary to refugee protection) against refoulement for persons who may not qualify as refugees stricto sensu under refugee law. In the Americas, national law in countries such as Canada, Chile, Colombia, Costa Rica, Guatemala, Mexico, and the U.S. mostly reflects the non-refoulement standards expressed by the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment art. 3, Dec. 10, 1984, 1465 U.N.T.S. 85, (entered into force June 26, 1987). See CANTOR, CROSS-BORDER DISPLACEMENT, supra note 11, at 41, 52–53.

See, e.g., KALIN & SCHREIFER, supra note 73, at 34 (arguing that the Cartagena Declaration definition may accommodate disaster-affected persons under the element relating to “other circumstances which have seriously disturbed public order”).

Mexico, Panama, Ecuador, and Peru recognized some Haitians as refugees under the definitions provided by one or other of these international instruments for refugee protection, based on the rise of insecurity in Haiti resulting from the 2010 earthquake. See CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 17–18.

protection. In fact, as a matter of national law, almost no state in the Americas treats a disaster as, in itself, a basis for international protection under “universal” or “regional” refugee definitions at the international level.\textsuperscript{136} Indeed, in some states, national law expressly rules out such an interpretation.\textsuperscript{137} In practice, certain states have even gone so far as to channel asylum claims by nationals of a disaster-affected country out of the international protection procedures so they can be resolved instead under other legal provisions unrelated to international protection.\textsuperscript{138} Thus, states in this region tend not to view persons fleeing a disaster as requiring international protection, except in certain specific cases where its impact includes clear risks of persecution or violence that fit with existing concepts of international protection. The fact of the disaster itself, though, is treated as legally irrelevant to determining international protection.

On its face, the current approach might seem to confirm the presumption that law in the Americas has yet to come to terms with the challenge of international mobility linked to environmental threats. However, a retrospective analysis of legal practice in this region shows that states have not always taken this approach to the application of international protection law. Indeed, between 1952 and 1980, national law in the U.S. expressly provided for different categories of “persons uprooted by catastrophic natural calamity” to be resettled to the U.S. as “refugees.”\textsuperscript{139} Moreover, during this early period, the U.S. was not alone in viewing the challenge of persons displaced by disasters as a matter of refugee protection. For instance, in 1978, Cuba adopted a definition of refugees including, inter alia, persons who leave their country “due to cataclysm or other phenomena of nature.”\textsuperscript{140} In 1979, the government of Trinidad and Tobago also contemplated the challenge of “refugees from natural disasters” and decided that such cases “be decided, when the need arises, on the basis of the circumstances prevailing in Trinidad and Tobago at the particular period in time.”\textsuperscript{141}

\textsuperscript{136} Cuba is the exception to this consensus. See Decreto No. 26, art. 80, July 19, 1978 (Cuba), https://perma.cc/F9MW-25VR (hereinafter Cuba Decree).

\textsuperscript{137} For example, Mexico expressly interprets the “other circumstances which have seriously disturbed public order” element of the Cartagena Declaration refugee definition as applicable only to “acts attributable to man.” See Reglamento de la Ley Sobre Refugiados y Protección Complementaria [Regulations of the Law on Refugees and Complementary Protections] [RLRPC], art. 4(XI), Diario Oficial de la Federación [DOF] 21-02-2012 (Mex.), https://perma.cc/5ZDY-A6TF (author translation).

\textsuperscript{138} See, for example, Weerasinghe, supra note 83, at 64–75 for a discussion of the Brazilian procedural response to Haitian asylum seekers after the earthquake.


\textsuperscript{140} See Cuba Decree, supra note 136 (author translation).

\textsuperscript{141} Cabinet Decision, Minute No. 4809, Nov. 16, 1979 (Trin. and Tobago).
These examples show that, contrary to the assumption by some legal scholars that international mobility linked to environmental threats represents a new legal gap, this challenge has long been recognized in the practice of certain states in the Americas. Indeed, the initial approach of those states to legally resolving the challenge by adopting unilateral and *sui generis* refugee definitions in national law waned only during the 1980s, as states across the Americas increasingly joined the U.N. refugee treaties and incorporated their “universal” refugee definition in national law. As a result, in this region today, the earlier approach persists solely in Cuba, which remains outside the U.N. refugee treaty regime. The role played by international law in this shift in approach is noteworthy. In this instance, whereas the legal scholarship usually envisages a positive role for international law in extending international protection to persons fleeing disaster contexts, here it appears to have curtailed the protection available to such persons under existing national law and thus helped create a “gap” as a result of promoting the harmonization of national law with U.N. refugee treaty law. Although this observation cannot be generalized beyond these specific examples, it calls us to think more critically about the relationship between national and international law in responding to this challenge.

B. An Alternative Legal Approach

From the 1990s, an alternative legal approach to the challenge of international mobility linked to environmental threats can be discerned in the broader “international protection” practice of certain states in the Americas. The creation of “temporary protected status” (TPS) in the national law of the U.S. offers one prominent example. Indeed, legal scholarship routinely cites the “environmental disaster” limb of TPS as one of a small number of protection provisions at the level of national law.\footnote{See note 70 and accompanying text.} This provision allows the U.S. authorities to designate a foreign state (or part of it) for TPS relief if:

(i) there has been an earthquake, flood, drought, epidemic, or other environmental disaster in the state resulting in a substantial, but temporary, disruption of living conditions in the area affected,

(ii) the foreign state is unable, temporarily, to handle adequately the return to the state of aliens who are nationals of the state, and

(iii) the foreign state officially has requested designation under this subparagraph.\footnote{Immigration and Nationality Act (INA), Pub. L. No. 101-649, § 244A(b), 104 Stat. 4978 (1990) (codified at 8 U.S.C. § 1254a).}
At the individual level, access to TPS is usually limited to nationals of the designated country who are already present in the U.S.\textsuperscript{144} It thus serves principally to temporarily regularize the immigration status of persons present irregularly. In most cases, though, TPS has turned out to be anything but temporary, with the affected countries repeatedly re-designated for TPS owing to the continuation of unstable conditions.\textsuperscript{145}

Foreign countries are designated for TPS only relatively infrequently. However, over the years, the status has benefitted a substantial number of persons.\textsuperscript{146} Thus, over 331,000 nationals of Honduras, Nicaragua, and El Salvador benefited from stay in the U.S. from TPS designations under this “environmental disaster” limb following the 1998 Hurricane Mitch in Honduras and Nicaragua and the 2000 earthquake in El Salvador. Likewise, an additional 55,000 Haitian nationals received TPS in the U.S. after the 2010 earthquake in Haiti. The Haiti designation, though, was done under a separate TPS limb that requires instead that the U.S. authorities determine the existence of “extraordinary and temporary conditions in the foreign state that prevent aliens who are nationals of the state from returning to the state in safety,” unless “permitting the aliens to remain temporarily in the United States is contrary to the national interest of the United States.”\textsuperscript{147} The 1997 volcanic eruption in Montserrat was designated simultaneously under both this “extraordinary and temporary conditions” limb and that relating to “environmental disaster.”

The TPS provision reflects the recognition that wider humanitarian circumstances beyond the rules of international protection law may legitimately be considered for stay. Indeed, it is not granted on the basis of any international obligation. Moreover, its relationship to “international protection” is somewhat

---

\textsuperscript{144} See, e.g., Designation of Haiti for Temporary Protected Status, 75 Fed. Reg. 3476 (Jan. 21, 2010) (illustrating that, in this regard, the re-designation of Haiti in 2011 to offer access to TPS for Haitians who had been continuously residing in the U.S. from a date prior to one year after the earthquake is exceptional); see also Extension and Redesignation of Haiti for Temporary Protected Status, 76 Fed. Reg. 29,000 (May 19, 2011).

\textsuperscript{145} Attempts since 2017 by the Trump administration to terminate long-standing TPS designations for nationals of certain countries continue, at the time of writing, to be litigated before the U.S. courts. See, e.g., Ramos v. Nielsen, 336 F. Supp.3d 1075 (N.D. Cal. 2018), vacated and remanded by Ramos v. Wolf, 975 F.3d 872 (9th Cir. 2020). The attempts to terminate TPS were apparently pursued by the administration in the face of advice from the U.S. State Department that this would put national security, foreign relations, and the beneficiaries’ American-born children at risk. Nicole Narea, State Department Officials Warned Trump Not to Revoke Protections for Immigrants, Vox (Nov. 7, 2019), https://perma.cc/DMX2-23C3.

\textsuperscript{146} See generally Jill H. Wilson, CONG. RISCH. SERV., RS20844, TEMPORARY PROTECTED STATUS: OVERVIEW AND CURRENT ISSUES 2, 5 (2020) (illustrating that TPS has also been used for contexts of war); see also Temporary Protected Status, U.S. CITIZENSHIP AND IMMIGR. SERVS. (Sept. 8, 2020), https://perma.cc/SH48-PMPD (providing full list of TPS countries and designation documentation).

\textsuperscript{147} § 1254a (b)(1)(C).
tenuous. For instance, while “extraordinary and temporary conditions” at least speaks to protection concerns in terms of the “safety” of returning nationals, the “environmental disasters” limb turns on relations between the U.S. and the disaster-affected state and the latter’s capacity to “adequately” receive returns. Moreover, even for disasters that meet the formal criteria for one or other limbs, no legal expectation exists that TPS will be designated. Nor can individuals apply for protection absent a determination of TPS for their country by the U.S. authorities, which remains at the complete discretion of the U.S. authorities. Although legal scholars have criticized TPS on those grounds,\(^\text{148}\) the analysis here is not intended to downplay its utility but to simply point out that it reflects a distinct legal approach not easily aligned with wider notions of international protection based on the severance of the protection relationship between individuals and their state of origin and an envisaged risk of serious harm if returned.

Most scholarship views TPS as an isolated example of states legislating for mobility in the disaster context. However, a brief review of national refugee law instruments in the Americas suggests that it actually forms part of a wider tendency to legislate for discretionary powers to allow entry and stay on broader humanitarian grounds, particularly where protection claims are not recognized. These powers have been used to benefit persons affected by disasters. In the Caribbean, for instance, the power to grant leave to remain to rejected asylum-seekers on “humanitarian grounds” was applied by Jamaica to Haitians after the 2010 earthquake.\(^\text{149}\) Similar powers exist in the refugee laws of the Cayman Islands and Montserrat in respect of rejected asylum-seekers who cannot be returned for “obvious and compelling reasons.”\(^\text{150}\) Suriname allows a residence permit to be granted to a rejected asylum-seeker if “he cannot in the light of the social and political situation in his country of origin and his personal circumstances reasonably be required to return to that country.”\(^\text{151}\) Like TPS, these provisions treat the disaster as an event that, in its own right, may engage broader humanitarian considerations for the entry and stay of persons who do not qualify for international protection. Rooted in the positive exercise of state discretion in immigration matters, these provisions in turn reflect a wider approach to addressing such situations in the broader immigration law of this region.

\(^{148}\) See CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 37–40; see also Bill Frelick, What's Wrong with Temporary Protected Status and How to Fix It: Exploring a Complementary Protection Regime, 8 IMMIGRATION & HUM. SEC. 42 (2020).

\(^{149}\) See Refugee Policy, ¶¶ 12(a)(iii), 13(f), 2009 (Jam.), https://perma.cc/U2RB-AT8Y.


V. IMMIGRATION LAW IN THE AMERICAS

The legal literature on the environment-mobility nexus largely overlooks how wider immigration law could address international mobility challenges linked to adverse environmental conditions.\textsuperscript{152} This partly reflects a perception that the situation of persons fleeing environmental threats is analogous to that of refugees, thus requiring international protection rather than immigration relief.\textsuperscript{153} Yet it also reflects a tendency to view the problem and its solution in terms of international law, thus discounting the relevance of immigration law as a field constituted principally at the national level.\textsuperscript{154} Indeed, most of the legal scholarship is quite dismissive of the role of national law in general.\textsuperscript{155} Even so, in the Pacific region, certain scholars have argued that international movement linked to climate change could be accommodated within existing migration mechanisms at the bilateral or regional level or by making tweaks to national immigration law in common destination countries in the region such as Australia or New Zealand.\textsuperscript{156}

In the Americas, the insight that immigration law might accommodate international mobility linked to environmental threats represents an important starting point for analyzing state practice. Yet, in contrast with the relative paucity of documented legal practice in other regions of the world, the use of immigration law to address this mobility challenge by states in the Americas is not merely a matter of speculation for the future.\textsuperscript{157} Indeed, the creation by the U.S. of an “environmental disaster” limb within TPS, which is essentially an immigration law provision for regularizing status in disaster contexts rather than a tool of international protection,\textsuperscript{158} in 1990 is just one early example of pertinent practice in this region. In Central America, for example, various states adopted legal

\begin{footnotesize}
\begin{enumerate}
\item The same is also true for disaster risk management law and policy frameworks. These mostly address cross-border mobility issues only in relation to the entry of personnel and assistance to a disaster-affected state, although some regional disaster risk management forums in the Americas have recently made general reference to the promotion of mechanisms to receive persons displaced across borders by a disaster. \textit{See} CANTOR, CROSS-BORDER DISPLACEMENT, \textit{supra} note 11, at 26–30. Yet such provisions seem to remain absent from disaster risk management frameworks at the national level in the Americas, except in Costa Rica. \textit{See} MICHELLE YONETANI, \textit{PLATFORM ON DISASTER DISPLACEMENT, MAPPING THE BASELINE: TO WHAT EXTENT ARE DISPLACEMENT AND OTHER FORMS OF HUMAN MOBILITY INTEGRATED IN NATIONAL AND REGIONAL DISASTER RISK REDUCTION STRATEGIES?} 29–31 (2018), https://perma.cc/8MRL-ZHBQ.
\item \textit{See} supra note 78 and accompanying text.
\item \textit{See} supra note 72 and accompanying text.
\item \textit{See} supra notes 65–71 and accompanying text.
\item \textit{See} supra note 97–98 and accompanying text.
\item As such, this existing practice also provides a counterpoint to analyses that claim the security fears of states have prevented them from using immigration law to address mobility linked to environmental factors. \textit{See}, e.g., Anastasiou, \textit{supra} note 65, at 187–89.
\item \textit{See} supra notes 143–148 and accompanying text.
\end{enumerate}
\end{footnotesize}
decrees in 1998 to regularize the immigration status of irregular migrants from other countries in the region that had been devastated by the effects of Hurricane Mitch.¹⁵⁹ In Costa Rica alone, the resulting program regularized around 150,000 disaster-affected migrants. ¹⁶⁰ It thus seems that the use of national immigration law in the Americas to resolve mobility challenges linked to the environment already represents fairly long-standing practice by some states in this region.

This Article advances legal debate on the environment-mobility nexus on several points. Most importantly, it shows just how widespread the use of immigration law instruments and concepts to resolve these challenges is among states in the Americas. It starts by illustrating how “ordinary” migration categories in national immigration law in this region have accommodated international mobility challenges linked to environmental factors. It then shows how a range of “exceptional” migration categories have also been created and deployed by states to accommodate persons whose legal situation cannot be resolved by application of “ordinary migration categories.” Overall, this analysis reinforces the impression of a shift in this region away from treating such challenges as matters of international protection to an approach based on immigration law.¹⁶¹ This means that, contrary to the assumptions of the existing legal scholarship, we cannot simply treat the regulation of mobility in this context as a blank canvas for international law. Rather, we must acknowledge that a distinctive legal approach to the problem already exists in the Americas and that it finds articulation among states not only in the Global North but also in the Global South.

A. “Ordinary” Migration Categories

A standard function of national immigration law is to codify and regulate access to what we might call “ordinary” migration categories. These ordinary migration categories usually exist to facilitate migration that is based primarily on pull factors in the country of destination or, in other words, an actual or prospective link on the part of the individual applicant with that country. Examples of short-term ordinary migration categories include such categories as


¹⁶¹ See supra Section IV.
visits or tourism, while longer-term ones include studies, employment, or joining family in the country of destination. Thus, as a basis for travel, entry, or stay by non-nationals, circumstances in the country of origin do not provide the principal rationale for these categories, which turns rather on certain forms of connection to the country of destination. Even so, and despite the scant attention paid to these migration pathways in existing legal research on the environment-mobility nexus, this Article shows that, in the Americas, they have accommodated mobility linked to environmental factors in a number of important ways.

Firstly, it is clear that these ordinary migration categories are used in practice by persons leaving a disaster-affected country as a way to enter or stay in another country. For sudden-onset disasters, the empirical data points to a spike in long-term regular migration to the U.S. from Central America and the Caribbean in the year after a sudden-onset disaster hits one of those countries. For slow-onset disasters, the documented increase in migration to the U.S. from parts of Mexico affected by such phenomena provides a similar indication. In tandem, short-term ordinary migration categories have also provided a legal basis for entry by inhabitants of border regions fleeing the impact of an oncoming sudden hazardous event on their side of the border. In the Americas, the use of such ordinary migration categories has particular salience. Not only is this a region with significant intra-regional diaspora populations, but in this region international mobility linked to environmental threats tends to follow existing migration routes and pathways where possible. As a result, in the Americas, the prospect that people from a disaster-affected country might have family or other links to a destination country in this region is not remote, especially among those sectors of society that possess the resources to migrate internationally.

The fact that ordinary migration categories are capable of accommodating a proportion of international movement linked to environmental push factors has far-reaching implications for debate about the legal gap in relation to such mobility and its resolution. On the one hand, it suggests that less of a gap in the law exists in reality than may be assumed in the abstract. This also implies that “solutions” must not focus exclusively on creating special new legal regimes for affected

---

162 This is the case even for categories, such as those relating to seasonal labor migration, that aim also to provide a secondary benefit to countries of origin, such as development gains through remittances, skills acquisition and knowledge transfer, alongside the principal benefit of temporarily linking foreign workers to gaps in the labor market of the receiving country.

163 See supra notes 116–119 and accompanying text.

164 See supra notes 112–113 and accompanying text.

165 See supra note 111 and accompanying text.


167 See examples supra Section III.
persons but more generally must also seek to ensure that states fairly apply these ordinary migration categories, especially in relation to disaster-affected countries. The imposition of undue restrictions on such migration categories might well have a greater negative impact on prospects for the entry and stay of persons affected by environmental threats than the absence of a dedicated protection regime. On the other hand, in principle, the fact of the disaster is legally irrelevant to the application of the ordinary migration categories, which turn instead on links to the destination country. Indeed, in the Americas, states clearly treat that criterion as the principal basis for determining the entry or stay of nonnationals, rather than any particular kind of push factor that may exist in the country of origin.

Secondly, for applicants affected by a disaster, some states in this region apply the formal criteria of these ordinary migration categories in a flexible manner. For instance, in Canada, immigration law allows the authorities to expedite applications under the ordinary migration categories or waive one or more formal criteria, if justified by “humanitarian and compassionate considerations.” This is applied in response to disaster situations, and, for some serious disasters, “special measures” policies are adopted by the government that instruct officials to exercise these powers in order to expedite applications or waive formal criteria where requested by applicants “seriously and directly affected” by the disaster. Meanwhile, the U.S. standing policy of “temporary relief measures” encourages immigration officials to exercise their innate discretion to expedite applications or waive the formal criteria for certain ordinary migration categories at the request of an individual applicant. Based on the periodic announcements reminding migrants of this policy, these relief measures

168 The use of visa regimes, which are often imposed on poorer countries that are more vulnerable to the impact of hazards, is a particular cause for concern. In the Americas, the countries whose nationals are most frequently required to secure a visa for lawful travel to another country within this region are Cuba, the Dominican Republic, and Haiti, although many other poor countries in this region that are regularly affected by disasters do not experience such extensive visa requirements for travel in the Americas. See CANTOR, CROSS-BORDER DISPLACEMENT, supra note 11, at 36, 47, 59.


170 These “special measures” policies have been adopted, inter alia, for the 1998 Turkey earthquake, the 2004 Asian tsunami, the 2010 Haiti earthquake and the 2013 Typhoon Haiyan in the Philippines. See CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 33–34.

171 See Immigration Help Available to Those Affected by Natural Disasters, U.S. CITIZENSHIP & IMMIGR. SERVS., https://perma.cc/8P8L-J8ZH (Oct. 6, 2017); see also CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 34–35.
seem to be applied mainly to those affected by natural disasters, not only overseas but also in the U.S. itself.¹⁷²

This flexible approach to the criteria for entry or stay under the ordinary migration categories is particularly codified in the law and policy of these Global North states. However, as a legal practice in the Americas, the approach is also evident among states in the Global South. In Central America, for example, Costa Rica has applied a broader understanding of the “family” category than normally permitted by law so that Nicaraguans present irregularly but personally affected by a sudden-onset disaster in Nicaragua could stay lawfully as family members, with all the benefits of that regular status.¹⁷³ Likewise, in South America, Colombia regularized some Haitians arriving after the 2010 earthquake by flexibly applying work and student categories.¹⁷⁴ In the Caribbean, Dominica and Antigua and Barbuda relaxed certain eligibility requirements of the ordinary migration categories for Haitians in 2010.¹⁷⁵ In the 2017 hurricane season, territories such as Montserrat and the British Virgin Islands also lifted immigration restrictions or waived visa requirements to facilitate entry by affected persons.¹⁷⁶ Overall, assimilating disaster-affected persons to ordinary migration categories has the advantage of access to ensuing regular status and rights. The states’ flexibility in this respect contrasts sharply with their rigid application of refugee law, supporting a view that they see mobility in this context principally as a matter of immigration law rather than international protection.

Thirdly, these migration pathways are also beginning to be shaped by free movement accords. At present, citizens of certain subregional integration mechanisms in Central America, South America, and the Caribbean benefit from specific forms of treaty-based free movement across borders within the respective

¹⁷² Situational “temporary relief measures” have been announced, inter alia, for such disasters overseas as: tropical storms in the Caribbean in 2008; the 2010 Icelandic volcano eruption; the 2010 Chile earthquakes; Tropical Storm Agatha in Guatemala in 2010; the 2011 earthquakes and tsunami in Japan; extreme flooding in Central America in 2011; Hurricane Sandy in the Caribbean in 2012; Typhoon Haiyan in the Philippines in 2013; Hurricane Harvey in the U.S. in 2017; California Wildfires in 2007 and 2018; Hurricane Florence in the U.S. in 2018; and the 2018 Typhoon Mangkhut in the Philippines. See Previous Special Situations, U.S. CITIZENSHIP & IMMIGR. SERVS., https://perma.cc/VMM3-HHQ9 (Nov. 26, 2018) (listing existing announcements of “temporary relief measures”).

¹⁷³ CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 32.

¹⁷⁴ Id. at 33.

¹⁷⁵ Id. at 35.

bloc. Some scholars have already argued in favor of extending free movement arrangements as a means of facilitating migration in the context of environmental threats. Yet states in the Americas have already begun to use such free movement provisions specifically to facilitate entry and stay by nationals of a disaster-affected country in their subregional bloc. For instance, after Hurricane Maria devastated Dominica in 2017, the authorities in Trinidad and Tobago used the Caribbean Community, or CARICOM, short-term visa-free stay provision to shelter affected Dominicans. In tandem, Antigua and Barbuda, Grenada, St. Lucia, and St. Vincent also welcomed Dominicans under the OECS provision for entry and short-term stay, expediting those cases and waiving documentary requirements where documents had been lost in the disaster. Lacking disaster-specific provisions, these accords now seem to offer additional useful ordinary migration categories for states to apply in disasters.

B. “Exceptional” Migration Categories

National law also often provides for what we might call “exceptional” migration categories. These categories usually take the form of general legal provisions, or powers conferred on immigration officials, created to regulate special or exceptional situations that fall outside the ordinary migration categories. They are squarely concerned with areas of immigration law where states enjoy a wide sovereign discretion, such as where an applicant lacks a substantive connection to the destination state or where a claim falls under binding rules of international protection, but other countervailing factors still exist. In the Americas, pertinent legal practice makes it clear that many states view disasters as precisely one such factor where special consideration may be required in relation to the application not only of ordinary migration categories but also these exceptional migration categories. Given the diversity of legal systems and juridical perspectives across this region, it is appropriate to analyze exceptional migration categories, and their application to disaster situations, along a spectrum of “codification” that reflects the different degrees to which they are expressly codified by national immigration law.

177 Relevant mechanisms include, respectively, the System for Central-American Integration (SICA), the Common Market of the South (MERCOSUR) and, for the Caribbean, the Caribbean Community (CARICOM) and the Organization of Eastern Caribbean States (OECS). See CANTOR, CROSS-BORDER DISPLACEMENT, supra note 11, at 36–37, 47–48 and 59.


179 FRANCIS, supra note 178, at 18.

At one end of this spectrum, the discretionary power to resolve exceptional immigration cases takes the form of an inherent faculty not specifically codified by immigration legislation, as is apparently the case in Venezuela. A little further along are states where the existence of this power is confirmed by immigration law but its scope is left open to the discretion of the national authorities, as in Colombia and Paraguay. Similarly, in the Caribbean, the law in British overseas territories and former colonies often gives officials the discretion to postpone, or overlook, deciding whether a non-national falls into one of the categories of “prohibited” immigrants who must be denied entry and instead granted leave to stay. Crucially, across the Americas, states have exercised this kind of broad discretionary power to grant entry or stay to disaster-affected persons on a case-by-case basis. In 2010, the Dominican Republic used apparently innate discretionary powers to grant entry to certain categories of Haitians affected by the earthquake on a humanitarian basis. After Hurricane Irma in 2017, the U.S. unincorporated territory of Puerto Rico used inherent discretion to grant entry to thousands of people evacuated from the British Virgin Islands, Dutch Sint Maarten, and French Saint Martin. In Chilean law, a discretionary power to grant stay in cases outside the ordinary migration categories was applied to benefit a small number of Haitians after the earthquake.

However, these broad discretionary powers are also used to facilitate stay on a group basis. In some cases, this involves creating special regularization programs to which nationals of the disaster-affected country who are already present

---

181 See CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 37–39.
182 For example, Colombian law provides for a power to authorize entry and stay on extraordinary grounds where this is necessary, see Decreto No. 1067, mayo 26, 2015, DIARIO OFICIAL [D.O.], art. 2.2.1.11.2.5 (Colom.), as modified by Decreto No. 1325, agosto, 12, 2016, https://perma.cc/3SG6-99AP. In Paraguay, the law gives the Director General a general discretionary power to “carry out other acts” necessary with complying with the objectives of the immigration authorities. See Ley No. 978/96, art. 146(g), June 27, 1996, (Par.), https://perma.cc/9BDQ-USAJ.
183 In the Caribbean, this is the case for Antigua and Barbuda, the Bahamas, Barbados, Dominica, Grenada, Jamaica, Saint Vincent and the Grenadines, and Trinidad and Tobago, as well as the British Overseas Territories of Anguilla, British Virgin Islands, Cayman Islands and Montserrat. See CANTOR, CROSS-BORDER DISPLACEMENT, supra note 11, at 60. On the mainland, the same is true for the former British colonies of Belize and Guyana. Id. at 37, 48. In Canada, officials can grant temporary resident status to persons who do not meet the requirements of the regular migration rules where they are “of the opinion that it is justified in the circumstances.” Immigration and Refugee Protection Act, supra note 169, s 24(1) (Can.).
184 CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 61.
186 Decreto No. 597 art. 49–50, Aprueba Nuevo Reglamento de Extranjeria, junio 14, 1984, DIARIO OFICIAL [D.O.] (Chile), https://perma.cc/GL7W-3GBA; see CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 43.
irregularly in the destination country can apply. In 1998, Central American states created such regularization programs for migrants from countries affected by Hurricane Mitch.\textsuperscript{187} In 2010, similar one-off regularization programs were created for Haitians present in Ecuador and Venezuela, using broad immigration discretion based on, respectively, statutory and innate powers.\textsuperscript{188} Such powers have also been invoked to create legal measures that fall short of formal stay but which still temporarily suspend removals to a disaster-affected country on a group basis. For instance, many states in the Americas drew on broad innate discretionary powers to temporarily suspend the removal of Haitians after the 2010 earthquake.\textsuperscript{189} Overall, the breadth of such powers gives states considerable latitude in fixing the criteria for their application, as well as excluding individuals in relation to whom security or crime related concerns exist.\textsuperscript{190}

Toward the other end of the codification spectrum are those national laws that codify when this immigration discretion should be exercised positively. In the Americas, this usually turns on the existence of “humanitarian” considerations in the individual case. Although the specific wording varies among countries,\textsuperscript{191} the law of at least fifteen states in this region include an exceptional migration category based on some variation of the concept of “humanitarian considerations.”\textsuperscript{192} In

\textsuperscript{187} See supra notes 158–159 and accompanying text.

\textsuperscript{188} Decreto No. 248, Feb. 9, 2010 (Ecuador), https://perma.cc/C8VB-8DFP; see also CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 37–39.

\textsuperscript{189} They include the U.S., Mexico, Bahamas, Dominican Republic, Jamaica, as well as the British Turks and Caicos Islands and the French Antilles territories of Martinique and Guadeloupe. See Michel Forst, Report of the Independent Expert on the Situation of Human Rights in Haiti, Michel Forst, Addendum: Forced Returns of Haitians from Third States, at 6–7, U.N. Doc. A/HRC/20/35/Add.1 (June 4, 2012); CANTOR, CROSS-BORDER DISPLACEMENT, supra note 11, at 38, 61, 63.

\textsuperscript{190} Nonetheless, certain profiles of person, such as those whose cases involve a national security or serious criminal element, are often deemed ineligible to benefit from these measures. See CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 40–41 (providing a discussion of how this played out for Haitians in Canada).

\textsuperscript{191} Examples from Central America include “exceptional humanitarian reasons” (Panama – entry and stay); “humanitarian cause” (Mexico – entry and stay); “humanitarian motives” (Honduras – entry); “humanitarian reason” (Costa Rica – entry); “humanitarian reasons” (Guatemala – entry and stay; Honduras – stay; Mexico – travel and stay; Nicaragua – stay); “humanitarian visa” (Mexico – travel; Nicaragua – entry and stay); “reasons of humanity” (Costa Rica – stay). See CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 38 n.207, 49 n.306, 60, 63 (concerning South America and the Caribbean).

\textsuperscript{192} By sub-region, those countries include: Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, Panama (Central America); Argentina, Bolivia, Brazil, Ecuador, Peru, Uruguay (South America); Trinidad and Tobago, the Dutch Antilles islands of Bonaire, Sint Eustatius, and Saba (Caribbean);
some countries, this concept is not defined further by national immigration law, leaving the potential for inclusion of disaster victims open to official discretion in individual cases. This is true of the special residence permit that may be granted for “humanitarian reasons” in Honduras and the extension of stay category for “humanitarian reasons” in Nicaragua. Likewise, the law in Uruguay allows entry as a temporary resident to be granted for “exceptional reasons . . . of a humanitarian character” but does not define what that means. In the Caribbean, the law in Trinidad and Tobago allows leave to remain to be granted if “humanitarian considerations” that warrant the granting of special relief from deportation exist, similar to the law in the Dutch Antilles. In addition, this immigration law concept is articulated in several national refugee laws. None of these laws further define the “humanitarian considerations” concept but, in practice, such broad provisions have sometimes been applied by officials to persons affected by a disaster in their country of origin.

More commonly, though, and particularly in Latin American countries, national immigration law more closely defines the scope of “humanitarian considerations” concept. This is usually done by reference to three general sets of circumstances, although it is important to point out that not all three are always codified in the law of any particular state. The circumstances are that the applicant is: (1) the “victim” of serious adversity, such as grave crimes or human rights

---


194 Ley No. 18.250, Ley de Migraciones, art. 44, 34(b), Dec. 27, 2008 (Uru.), https://perma.cc/X96P-X454 (author translation).


196 For the Dutch Antilles islands of Bonaire, Sint Eustatius, and Saba, an official who has doubts about refusing entry may refer the case to the immigration authorities in the Netherlands who can decide to grant entry due to, inter alia, “compelling humanitarian reasons,” although this concept is not further defined. Circulaire toelating en uitzetting Bonaire, Sint Eustatius en Saba oktober 2010, Stert. 2010, § 2.3.6 (Neth.), https://perma.cc/4H6C-QZG5 (translation by author’s colleague).

197 See supra notes 149–151, 151 and accompanying text.

198 For instance, in Honduras, the authorities were preparing to receive Haitians in the aftermath of the 2010 earthquake using these provisions, although none actually arrived. See CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 46.
violations; or (2) “vulnerable” in the destination country, due to factors such as age, gender, disability or ill health; or (3) “facing serious danger” to life or integrity in the country of origin. Although this last scenario may resemble a rule of international protection, its application is usually discretionary. Disaster-affected persons are sometimes accommodated within such broad renderings of the “humanitarian considerations” concept. For instance, in 2017, Haitians present irregularly in Argentina were granted stay under a general provision of this kind, which was interpreted as applying to natural disasters and their effects. Likewise, the “humanitarian and compassionate” considerations provision of Canadian law is interpreted as a test of “unusual and undeserved or disproportionate hardship,” to be determined by reference to factors that include conditions in the country of origin, particularly those that have “a direct negative impact on the applicant such as . . . natural disasters.”

However, the opposite end of this codification spectrum actually consists of the growing number of immigration law provisions in the region that specify disasters as a “humanitarian consideration” or otherwise as the basis for an exceptional migration category. Ten states across the Americas take this approach, with the tendency particularly accentuated in South America and North America but also becoming increasingly common in Central America. The ubiquity of these provisions supports the analysis that states in the Americas view international mobility linked to environmental factors not through the lens of international protection, but principally through the lens of immigration law as an integral expression of their asserted sovereign right to determine who is allowed entry and stay in the territory of the state, treating the humanitarian impact of disasters as a legal basis for exercising state discretion in favor of affected persons.

Starting with South America, national immigration law in Argentina stipulates that transitory residence for “humanitarian reasons” can be granted to those who “temporarily cannot return to their countries of origin . . . due to

199 For examples of these factors the national law of countries in Central America and South America, see CANTOR, CROSS-BORDER DISPLACEMENT, supra note 11, at 38–39, 49–51. For Canada, the applicable provision is the Immigration and Refugee Protection Act, s 25(1.21). Immigration and Refugee Protection Act, supra note 169, s 25 (Can.).


202 They are Argentina, Bolivia, Brazil, Canada, Ecuador, El Salvador, Guatemala, Mexico, Peru, and the U.S.

203 Currently, in the Caribbean, states are also reported to be considering how to make legal provision. See CARIBBEAN MIGRATION CONSULTATIONS, supra note 176, at 15.
consequences generated by natural or man-made environmental disasters.”

Brazil authorizes “humanitarian reception” for a person from “any country in a situation of . . . major calamity [or] environmental disaster.”

Ecuador gives stay for “humanitarian reasons,” including being “a victim of natural or environmental disasters.”

Peru authorizes “humanitarian residence” where migration is due to “natural and environmental disasters.”

Similarly, in North America, for the purpose of granting a humanitarian visa to a non-national outside the country, Mexico defines “humanitarian reasons” as meaning that the person seeking to travel to Mexico “finds herself in a situation of danger to her life or integrity owing to . . . a duly accredited natural disaster” or that she is “victim of a natural catastrophe.”

Albeit not expressly based on the concept of “humanitarian considerations,” immigration law in the U.S., as already mentioned, gives authorities the discretion to designate TPS for, inter alia, an “environmental disaster.”

In Canada, national law likewise allows temporary

---


208 The law charges the Bolivian migration authorities to “make viable, as necessary, the admission of populations displaced by climate effects, when a risk or threat to their lives may exist, where those are due to natural causes or environmental, nuclear [or] chemical disasters or hunger.” Ley No. 370, Ley de Migración, art. 65, May 8, 2013 (Bol.), perma.cc/P2EV-ECD7 (author translation). The law in Bolivia provides a unique definition of “Climate Migrants” as “[g]roups of persons who are forced to displace from one State to another due to climate effects, when a risk or threat to their life may exist, whether due to natural causes, environmental, nuclear [or] chemical disasters or hunger.” Id. art. 4(16).


210 Lineamientos Generales para la Expedición de Visas que Emiten las Secretarías de Gobernación y de Relaciones Exteriores, Del Procedimiento para Solicitar Visa, Trámite 9, Criterios de Resolución a.i., Diario Oficial de la Federación [DOF] 10-10-2014 (Mex.) (author translation) [hereinafter General Guidelines for the Issuance of Visas].

211 See supra notes 142–146 and accompanying text.
suspension of removals, inter alia, due to an “environmental disaster resulting in a substantial temporary disruption of living conditions” in the country of origin.\textsuperscript{212}

More recently, in Central America, some states have adopted new immigration laws that also refer to disasters in exceptional migration categories based on humanitarian considerations. For example, in Guatemala, the existence of a “natural catastrophe in neighboring countries, which obliges the persons or group of persons to flee for their lives” is listed among the “humanitarian reasons” for legal entry and stay.\textsuperscript{213} In El Salvador, factors to be taken into account by immigration officials in deciding temporary resident applications based on “humanitarian reasons” specifically include, inter alia, the existence of an “internationally-recognized crisis” or that any non-national who does not meet the criteria for an ordinary migration category is in “a situation of vulnerability or danger to life owing to … natural disasters [or] environmental [disasters].”\textsuperscript{214}

Overall, legal practice in the Americas shows that surprisingly few states have not applied such exceptional migration categories as a matter of national immigration law to accommodate disaster-affected persons whose legal situation cannot be resolved via ordinary migration categories. Indeed, there is no real absence of legal tools to resolve the challenge of entry and stay in light of prevailing humanitarian considerations in this mobility context, and these tools are applied in practice. Yet, while some ordinary migration categories may provide permanent stay, exceptional migration categories mostly give temporary stay. The initial period varies between one year (for example, Costa Rica) and six years (for example, Panama), although this is normally renewable and can offer a pathway to longer forms of stay under ordinary migration categories. Such stay also usually affords the entitlements to work and access services specified by immigration law in that country. Indeed, in many countries, these immigration categories provide a defined regular status, a period of stay and a range of rights no less favorable than those conferred by refugee status.\textsuperscript{215}

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{212} Immigration and Refugee Protection Regulations, SOR/2002-227, s 230 (Can.). For details of how this was applied in the Haitian case, see CANTOR, LAW, POLICY AND PRACTICE, supra note 11, at 40–41.
\item \textsuperscript{213} Decreto No. 44, Código de Migración, art. 68, Oct. 23, 2016 (Guat.), https://perma.cc/7FUB-ZFL5 (author translation).
\item \textsuperscript{214} Decreto No. 35, Reglamento de la Ley Especial de Migración y de Extranjería, art. 181(2), 181(7), May 24, 2019, 2019 (El Sal) (author translation).
\item \textsuperscript{215} Although they do not benefit from any specific guarantee against refoulement, such issues do not easily arise in the disaster context. Indeed, most disaster migrants in the Americas do not strictly require “protection” from their own state by another state. Likewise, the mere fact of a disaster does not automatically turn removal into refoulement, only where disaster conditions are so serious that human rights standards will anyway temporarily prohibit removals. Finally, disaster migrants already established in the destination country may also be protected indirectly from any return or potential refoulement by due process guarantees in law against the arbitrary expulsion of aliens.
\end{itemize}
\end{footnotesize}
Finally, the legal practice in the Americas raises a question about the kinds of environmental threats that such measures accommodate. Here, the empirical evidence shows that people who move in the context of both sudden- and slow-onset disasters have both been accommodated under ordinary migration categories. However, where the disaster is expressly contemplated as a basis for entry or stay, whether in the flexible application of ordinary migration categories or via exceptional migration categories, the legal practice in the Americas suggests that such provisions are mostly applied by national immigration authorities to resolve the situation of persons affected by major sudden-onset disasters. On its face, this seems to confirm the preoccupation of scholars that persons who migrate as a result of the impact of slow-onset disasters will be denied access to such special measures because the link to mobility is easier to establish in the context of sudden-onset disasters. The implicit requirement of most states as a matter of law or practice is that a person must be directly and seriously affected by the disaster in order to benefit from the application of such special measures might seem to further reinforce this risk, since that link seems more straightforward to evidence in situations of sudden-onset disasters as compared to slow-onset processes.

At the same time, many of the national law provisions underpinning these special measures—under both ordinary and exceptional migration categories—do not refer expressly to a “disaster” but rather to broader “humanitarian considerations.” In principle, then, they do not rule out the application of these special measures to persons affected, on the one hand, by slow-onset disasters, or, on the other, by sudden- or slow-onset hazards or hazardous events that have not resulted in a disaster at the societal level. Moreover, not one of the provisions that refer to disaster situations as a basis for special measures gives any reason to think that slow-onset disasters fall outside their scope. It is also the case that several of these provisions refer to “calamities” or “catastrophes” alongside, or instead of, “disasters,” suggesting concepts which could equally include hazardous events more generally. Further, as to the origins of the events, these provisions

---

216 See supra note 50 and accompanying text.

217 See supra notes 190–214 and accompanying text. The same is true for similar provisions in national legislation on international protection. See supra notes 148–150 and accompanying text.

218 Indeed, the provision in Bolivian law specifically includes wider “climate effects.” See Ley No. 370, supra note 208, art. 65 (Bol).

219 Alongside the “disaster” concept, the provision in Brazilian law refers to “major calamity.” See Lei No. 13.445, supra note 205, art. 30. That in Mexican law refers to “catastrophe.” See General Guidelines for the Issuance of Visas, supra note 210. The provision in Guatemalan law refers to a “catastrofe” rather than a disaster. See Decreto No. 44, supra note 213, art. 68.
often refer to “natural” disasters or catastrophes, although many refer also, or instead, to “environmental” disasters or catastrophes. The latter concept appears to include events that are “man-made” in origin. In short, although the current practice may be to apply special measures mainly to persons who are seriously and directly affected by major sudden-onset disasters linked to natural hazards, relevant national law provisions suggest that a broader set of scenarios may ultimately be contemplated.

VI. INTERNATIONAL COOPERATION AND FRAMEWORKS IN THE AMERICAS

In the Americas, state practice in relation to the challenge posed by international mobility linked to environmental adversity also takes the form of joint action at the international level. This practice plays out principally in forums concerned with migration rather than international protection and at the level of subregional initiatives rather than regional or global forums. Crucially, the main focus of such joint action is on promoting appropriate legal responses to the challenge by participating states at the level of national law rather than creating new treaty law. However, the scope of cooperative ambition has increased over time. Thus, while early forms of collective action represented ad hoc responses to the devastation wrought by certain very serious disasters, since the mid-2010s, several subregional forums have been engaged in developing normative frameworks that promote more predictable responses at the level of national law. As yet, states appear to be content with this form of international action and no serious efforts have been made to “harden” the legal character of these frameworks through creating treaties. Even so, the existence and scope of these frameworks raise important questions about the future development of international law in this field.

220 The law in Mexico refers to natural disasters or catastrophes. See Ley de Migración, supra note 209, art. 41. While the law in Guatemala refers only to natural catastrophes. See Decreto No. 44, supra note 213, art. 68).

221 For the respective legal provisions in Ecuador, see Ley Orgánica de Movilidad Humana, supra 206, art. 58; Peru, see Decreto Legislativo No. 1350, supra 207, art. 29(2)(k); Bolivia, see Ley No. 370, supra note 208, art. 65; and El Salvador, see Decreto No. 35, supra note 214, art. 181, which refers to natural and environmental disasters.

222 For the respective legal provisions in the U.S., see supra notes 143–147 and accompanying text; and for Canada’s reference to “environmental” disasters, see Immigration and Refugee Protection Regulations, SOR/2002-227, s. 230.

223 For instance, the legal provision in Argentina refers to “natural or man-made environmental disasters.” See Law No. 25871, supra note 204, art. 24(h). The Bolivian provision adds “nuclear [or] chemical disasters or hunger” too. See Ley No. 370, supra note 208, art. 65.

224 Other legal studies appear to overlook the extent of such existing practice on precisely this issue at the international level in the Americas. See, e.g., Pires Ramos & de Salles Cavdeon-Capdeville, supra note 11.
A. Early Ad Hoc Actions

State practice at the subregional level is relatively well-established in the Americas. Certainly, state cooperation regarding the challenges of the international mobility posed by sudden-onset disasters is not new in this region. However, the early instances of state practice at the subregional level tend to involve the collective recognition of the international mobility consequences of certain extremely serious sudden-onset disasters and the promotion of suitable responses at the level of national law. This is evident in the way that states in such subregional forums responded to the devastation wrought by Hurricane Mitch. In 1998, for instance, the Meeting of Central American Presidents appealed for:

... the understanding of the International Community [sic] in order that a general amnesty be conceded to undocumented Central American immigrants who currently reside in different countries, with the objective of avoiding their deportation and, consequentially, greater aggravation of the current situation of our countries.225

At the national level, this declaration by the four most affected states facilitated the designation of TPS for Hondurans and Nicaraguans by the U.S. as well as the adoption of special measures for affected persons in national immigration law by various Central American states.226

The mobility impact of Hurricane Mitch was also addressed by other subregional forums, albeit also in an ad hoc or responsive manner. Thus, in early 1999, the matter was raised by states participating in the Regional Conference on Migration (RCM), a specialized regular subregional forum for facilitating joint discussion and action on shared migration challenges among a wider constituency of eleven member states, mostly from North and Central America. In the resulting Communication, the RCM member states “made special mention of the adoption, by Costa Rica and the United States of America, of migration measures benefiting nationals of the countries affected by Hurricane Mitch.”227 Both subregional forums, then, called not only for a common response on the part of states to the international challenges posed by this particularly serious disaster but also for the use of national immigration law by relevant states to resolve those challenges.


226 For Central American examples, see supra notes 159–160 and accompanying text. A copy of the declaration that was sent to the U.S. with a letter drawing attention to this plea appears to have influenced the granting of TPS. See Meeting of Central American Presidents, supra note 225, at 8.

Yet Hurricane Mitch is not the only disaster where such responsive collective action manifested itself. Indeed, subregional forums across the Americas also promoted these kinds of special immigration law measures as a form of collective response to the overwhelming impact of the 2010 Haiti earthquake. For instance, in 2010, the twelve states which at that time comprised the subregional Union of South American Nations (UNASUR) adopted a collective decision to promote “joint actions.” That UNASUR decision specifically exhorted “those Member States that still have not applied special processes of migratory regularisation for the benefit of Haitian citizens to do so.”

Likewise, the subregional Bolivarian Alliance for the Peoples of Our Americas (ALBA), comprised at that time of nine mostly left-leaning governments, called on member states to “decree a migratory amnesty that regularizes the migratory status of Haitian citizens resident in ALBA countries.” As a result, several states that were involved in one or both forums went on to adopt special migratory measures into national law for the benefit of Haitians present irregularly on their territories.

These examples demonstrate that, from the 1990s to the early 2010s, many states in the Americas participated in subregional forums that took joint action on international mobility linked to environmental adversity. The responsive and ad hoc nature of such action is immediately apparent. It manifested only following the occurrence of extremely serious sudden-onset disasters that posed humanitarian challenges suddenly across the pertinent subregion on a significant scale. On the mobility aspect of these challenges, they were clearly seen through the prism of immigration law rather than international protection. In particular, these subregional forums sought to promote, on a humanitarian basis, the immigration regularization of nationals of the affected state who were present irregularly elsewhere. As such, the joint action taken by states during this period was thus oriented toward encouraging a common response at the national law level, in the form of special migratory measures, rather than creating new international frameworks for the future.

B. Promoting Frameworks for Cooperation

Since the mid-2010s, states in several subregional forums in the Americas have been working to develop normative frameworks for promoting a more predictable response at the national level to future disaster displacement and cooperation between states at the subregional level. In this process, the influence

---


230 See supra note 188 and accompanying text.
of external actors has been central to encouraging states to build in this way on past practice at national and subregional levels. That work was led initially by the Nansen Initiative (2012–2015)—a global intergovernmental process focused on disaster displacement—and is being continued by its successor initiative, the Platform on Disaster Displacement (PDD) (2016). Even so, states in each subregion have clearly drawn on the expertise and other resources offered by this external actor to shape normative tools that they view as useful in responding to the challenges posed by potentially increasing levels of international mobility linked to disasters.

The development of the pertinent subregional frameworks in the Americas has taken place mainly in interstate forums concerned with cooperation on migration issues rather than international protection. Nonetheless, in 2014, engagement by the Nansen Initiative at the regional level resulted in Latin American and Caribbean states recognizing the “challenges posed by climate change and natural disasters, as well as by the displacement of persons across borders that these phenomena may cause in the region” in the Brazil Declaration on Refugees. The accompanying Brazil Plan of Action of 2014 called on the office of the United Nations High Commissioner for Refugees (UNHCR) to prepare a study on this theme in order to facilitate “regional cooperation.” That study, commissioned by UNHCR and the PDD (as successor to the Nansen Initiative), was published in 2018; it fed into the South American and Caribbean subregional processes described below. However, it seems that no further measures were taken at the regional level.

In tandem, the Nansen Initiative had been working with states at the subregional level. In 2013, a Central America consultation recommended that a set of guidelines drawing on national practice be developed through the subregional RCM forum. On the proposal of Costa Rica, this was approved by

---

231 For a description of the approach, see Platform on Disaster Displacement, supra note 63, at 126, 141. See generally Jane McAdam, From the Nansen Initiative to the Platform on Disaster Displacement: Shaping International Approaches to Climate Change, Disasters and Displacement, 39 UNIV. NEW S. WALES L.J. 1518 (2016).


233 Id. at Chapter Seven (Plan). The study was published by UNHCR and PDD in 2018. See generally CANTOR, CROSS-BORDER DISPLACEMENT, supra note 11.

234 See generally CANTOR, CROSS-BORDER DISPLACEMENT, supra note 11.

the RCM. In 2016, based on a study commissioned by the Nansen Initiative, the RCM adopted nonbinding guidance on “Protection for Persons moving across Borders in the Context of Disasters.”

A similar process in South America, initiated in 2015, led to a proposal by Chile to develop guidelines through the South American Conference on Migration (SCM), a subregional forum of twelve South American states. In 2018, with support from the PDD, the SCM in turn adopted its own nonbinding “Regional Guidelines on Protection and Assistance for Persons Displaced across Borders and Migrants in Countries affected by Disasters of Natural Origin.”

Since 2019, the PDD (as successor to the Nansen Initiative) has sought to build on this engagement elsewhere in the Americas by supporting a similar process of consultation in the Caribbean through the subregional migration forum of the Caribbean Migration Consultations (CMC). In 2019, at the first Caribbean consultation, the 2018 UNHCR study requested by Latin American and Caribbean states in the 2014 Brazil Plan of Action was presented to participants. The participating states framed the new challenges at regional and national levels as a priority “in a context of increased migration and displacement linked with climate change and disasters.” As a next step, those same states agreed on the need to “systematize approaches, harmonize them, and come up with consolidated policies through collaboration.” In the coming years, the creation of a similar subregional framework in the Caribbean, thus, looks like a possibility.


237 See generally CANTOR, LAW, POLICY AND PRACTICE, supra note 11.


239 The process within the SCM was initiated in 2016. However, this built on a regional workshop with South American states that was convened in Quito during 2015 by the government of Ecuador, the Nansen Initiative, and the Refugee Law Initiative of the School of Advanced Study, University of London.

240 Conferencia Suramericana Sobre Migraciones, Lineamientos Regionales en Materia de Protección y Asistencia a Personas Desplazadas a Través de Fronteras y Migrantes en Países Afectados Por Desastres de Origen Natural (2018), https://perma.cc/3PHH-W5EZ [hereinafter SCM GUIDELINES].

241 See generally CARIBBEAN MIGRATION CONSULTATIONS, supra note 176.

242 Id. at 10.

243 Id.
C. Subregional Frameworks: Approach and Scope

The guidelines adopted by the RCM and the SCM represent a significant addition to state practice. They reflect the views and approval of the large number of participating states in the subregions of North and Central America (through the RCM) and South America (through the SCM). Equally, as a result, their guidance extends to that same range of member states across the Americas. Moreover, as the first international instruments expressly oriented toward regulation of international mobility in the context of disasters, they provide a crucial indication of how states in this region legally frame these aspects of the environment-mobility nexus. The approaches taken in the RCM Guide and the SCM Guidelines thus offer useful insight into how, in the future, international law may come to relate to the challenges posed by international mobility in the context of environmental adversity.

There is considerable consistency in the approaches adopted by both the RCM Guide and the SCM Guidelines. Those points of convergence give an important indication of how states frame the key issues. At the outset, though, certain drafting differences between the two instruments must be acknowledged. Most crucially, they differ in how the guidance is presented. The RCM Guide describes its normative framework as “effective practices” and gives significant detail on each, while the SCM sets out broader “general guidelines” on purported “minimum standards.”244 Yet, in reality, any difference lies merely in the greater or lesser degree of detail that each instrument provides on the norms affirmed. The scope of intervention that each instrument envisages differs somewhat too. The RCM Guide mainly addresses immigration and consular authorities on measure for the benefit of affected persons post-disaster. That scenario is covered by the SCM Guidelines, but they also promote a “whole of government” approach to avoid the risk of displacement from the outset.245 In this sense, the SCM Guidelines have broader scope.

Nonetheless, on the key question of how host states should regulate international mobility in the context of disasters, the instruments are highly consistent in their approach. Firstly, they evidence a view by states that new international law norms are not required. They do not “create a new set of state obligations, extend existing state obligations, or require that new laws be passed.”246 Rather, the instruments are intended only to “support the more effective and consistent use of existing law, policy and practice” by states.247 The

---

244 See SCM GUIDELINES, supra note 240, at 15–16.
245 Id. at 28. Indeed, alongside the many examples throughout the text, this is one the orienting principles of the SCM Guidelines.
246 RCM GUIDE, supra note 238, at 8; see also SCM GUIDELINES, supra note 240, at 15, 25.
247 RCM GUIDE, supra note 238, at 8; see also SCM GUIDELINES, supra note 240, at 15.
body of existing national law in member states is thus seen as sufficient to respond to the challenges of this scenario. In tandem, such national law is also affirmed as the main basis for the resulting normative frameworks, which merely provide guidance on how to apply it. This approach of building from existing national law and practice is distinct from other soft law instruments which either express aspirational norms lacking a firm basis in existing law or interpret hard rules of international law in their application to a particular group or theme.

Secondly, both the RCM Guide and the SCM Guidelines treat this area as principally a matter of immigration law. The frameworks are primarily built upon the national practice of states in each subregion of favorably exercising their discretion in immigration matters where “humanitarian grounds,” such as a disaster exist. Accordingly, each instrument also acknowledges that states retain the inherent discretion to adopt more generous approaches than those described in the subregional norms. International obligations are acknowledged but are incorporated mainly as a set of parameters that may limit the extent to which states can decline to favorably exercise discretion in some situations, rather than the core legal basis for resulting measures. Moreover, both instruments address the legal challenges not only for people arriving due to a disaster in their own country, but also for people already outside their own country who are affected by a disaster there or by a disaster in the country in which they are living or through which they are transiting.

Thirdly, the norms in both instruments are rooted principally in the use of ordinary and exceptional migration categories. Both guides distinguish the (humanitarian) protection afforded by these categories from international protection, which each highlights as relevant to disaster displacement only in rare cases. Moreover, each instrument promotes active forms of cooperation

---

248 See RCM GUIDE, supra note 238, at 9; SCM GUIDELINES, supra note 240, at 14–15.
249 See RCM GUIDE, supra note 238, at 13; SCM GUIDELINES, supra note 240, at 20–25, 33.
250 See RCM GUIDE, supra note 238, at 9; SCM GUIDELINES, supra note 240, at 15.
251 See RCM GUIDE, supra note 238, at 13–15; SCM GUIDELINES, supra note 240, at 34.
252 See RCM GUIDE, supra note 238, at 10–11; SCM GUIDELINES, supra note 240, at 16. Indeed, each instrument sets out a range of relevant principles and norms relating to protection and assistance for migrants in a disaster-affected country. See RCM GUIDE, supra note 238, at 24–36; SCM GUIDELINES, supra note 240, at 39–42. These norms correlate with those described by another migration-oriented set of guidelines, on migrants in countries in crisis, prepared at the global level by the International Organization for Migration with input from states. See, e.g., MIGRANTS IN COUNTRIES IN CRISIS INITIATIVE, GUIDELINES TO PROTECT MIGRANTS IN COUNTRIES EXPERIENCING CONFLICT OR NATURAL DISASTER (2016), https://perma.cc/9ZG5-FTV9.
253 See RCM GUIDE, supra note 238, at 11–12; SCM GUIDELINES, supra note 240, at 34.
254 See SCM GUIDELINES, supra note 240, at 24–25 (providing a broad definition of “protection” in Part 3.2, as compared with the emphasis on lack of protection in the definition of “international protection” and indeed “complementary protection” at Part 3.2).
between the host state and state of origin, including bilateral measures to further cooperation and mutual assistance, especially where they share a border, and, in solidarity with a country of origin that lacks capacity to receive returns, granting entry or stay to persons only tangentially affected by a disaster. The SCM framework is expressly based on coordination between these two states, and it even posits a principle of “shared responsibility” between them, especially if both are SCM members.

Thus, in each subregional forum, it seems that states do not generally view the situation of disaster-affected persons through an “international protection” lens.

Finally, neither the RCM Guide nor the SCM Guidelines attempt to create a new legal status. Rather, they simply distil from the legal practice of states in each subregion a consensus approach to when the discretion to allow entry or stay on humanitarian grounds should usually be positively exercised in disaster contexts. This is when the person is “directly and seriously affected by the disaster.” Some interpretation of this concept is provided by the RCM Guide. The emphasis is squarely on the proximity and severity of the disaster’s impact on the individual, in light of any pre-existing vulnerabilities. In other words, unlike the established concepts of international protection, it is not based on a rupture in the political link between an individual and their state of origin or a risk of their fundamental human rights standards being violated there.

The interpretation in the RCM Guide of a disaster’s direct impact as a “sudden and severe change” suggests slow-onset disasters are not included. This is confirmed by the guide’s affirmation that it applies only to disasters “caused in part or in whole by a sudden and serious natural hazard.” By contrast, the SCM Guidelines expressly include other slow-onset disasters and events “that may be associated with adverse effects of climate change” when they contribute in...
fundamental ways to the affected person’s decision to cross an international border.\textsuperscript{264} Otherwise, though, the terms of each instrument strongly suggest that the principal concern is with “disasters,” a concept defined by reference to the extant U.N. policy.\textsuperscript{265} As such, it seems that hazards or hazardous events that do not reach the implicit threshold for societal impact will fall generally outside the scope of the guidance and be left purely to the discretion of national state authorities.

Despite their recent adoption, these two subregional instruments have already begun to shape state practice in the Americas. For example, in the RCM Guide, the principles on bilateral cooperation have been acted upon by some states. Costa Rica and Panama have broken new ground in the subregion of Central America by developing from earlier drafts of the RCM Guide shared by PDD a set of bilateral mechanisms and policies to manage displacement and disaster risks.\textsuperscript{266} These include a set of draft Standard Operating Procedures (SOPs) for their respective disaster response systems to address cross-border displacement in disaster contexts.\textsuperscript{267} The structure, principles, and rules in the SOPs are based directly on the RCM Guide. Simulation exercises to put the SOPs into practice have been carried out jointly by the two countries, again with the involvement of PDD.\textsuperscript{268} Similarly, the RCM Guide facilitated coordination between Costa Rica and Nicaragua to allow 150 Nicaraguans to cross the border and shelter in Costa Rica from the effects of Hurricane Otto in 2016.\textsuperscript{269}

The two subregional instruments, and the processes that led to their adoption, also seem to have encouraged states in Central and South America to incorporate national law provisions to specifically regulate the entry and stay of persons affected by a disaster when revising their immigration laws. Since the Nansen Initiative started work on its consultations in those regions in 2013, at least five states in these subregions have adopted significant new provisions of national law specific to the situation of disaster-affected non-nationals when

\textsuperscript{264} See SCM GUIDELINES, supra note 240, at 16.

\textsuperscript{265} They thus reflect later and earlier versions of this concept in UN policy post- and pre-2016. See SCM GUIDELINES, supra note 240, at 24; RCM GUIDE, supra note 238, at 9–10. For a discussion of the U.N. approach to defining the concept, see supra notes 41, 44–49 and accompanying text.

\textsuperscript{266} See International Organization for Migration, Costa Rica and Panama Prepare for Cross-Border Disaster Displacement, PLATFORM ON DISASTER DISPLACEMENT (Aug. 23, 2017), https://perma.cc/F53M-JC3M.


\textsuperscript{268} Int’l Org. for Migration, Costa Rica and Panama Effectuate First Border Crossing Simulation, PLATFORM ON DISASTER DISPLACEMENT (Aug. 21, 2017), https://perma.cc/T8Q8-8FCA.

\textsuperscript{269} See Platform on Disaster Displacement, supra note 63, at 126, 141.
overhauling their immigration legislation. They are Guatemala (2016),\textsuperscript{270} Brazil (2017),\textsuperscript{271} Ecuador (2017),\textsuperscript{272} Peru (2017),\textsuperscript{273} and El Salvador (2019).\textsuperscript{274} Paraguay is also reported to be considering such provisions as it debates adoption of a new immigration law.\textsuperscript{275}

Finally, the legal practice of states in the Americas has also had an impact on policy development at the global level. In 2015, the Nansen Initiative presented an Agenda for the Protection of Cross-border Displaced Persons in the context of Disasters and Climate Change (Protection Agenda), which—based on seven regional consultations with states and other actors—sets out norms for responding to cross-border disaster displacement.\textsuperscript{276} At that conference, over 100 states from different regions endorsed these global guidelines.\textsuperscript{277} However, on closer study, it is evident that the approach and many of the more novel norms described by the Protection Agenda for “protecting cross-border disaster-displaced persons” are derived principally from state practice in the Americas.\textsuperscript{278} Even so, in 2018, the approach in the Protection Agenda was endorsed by the Global Compact for Safe, Orderly and Regular Migration as a basis for developing “coherent approaches to address the challenges of migration movements in the context of sudden-onset and slow-onset natural disasters.”\textsuperscript{279}

\textsuperscript{270} See Decreto No. 44, \textit{supra} note 213, art. 68 (Guat.).

\textsuperscript{271} See Lei No. 13.445, \textit{supra} 205, art. 30 (Braz.).

\textsuperscript{272} See Ley Orgánica de Movilidad Humana, \textit{supra} 206, art. 58, 66(5) (Ecuador).

\textsuperscript{273} See Decreto Legislativo No. 1350, \textit{supra} 207, art. 29 (Peru).

\textsuperscript{274} See Decreto No. 35, \textit{supra} note 214, art. 181 (El Sal.).

\textsuperscript{275} Ministerio del Interior, Dirección General de Migraciones, \textit{Presentan Propuesta de Reforma de la Ley Migratoria Nacional al Poder Legislative} (Aug. 9, 2016), https://perma.cc/D4K4-5X3R.


\textsuperscript{278} See Nansen Initiative, \textit{supra} note 276, at 21–22, 24, 26–27, \textit{¶¶} 33–34, 38–40, 47–53, 64–65. See also Nansen Initiative, \textit{Agenda for the Protection of Cross-border Displaced Persons in the Context of Disasters and Climate Change}, Vol. 2 42–48 (2015), https://perma.cc/WZY4-U6UJ (describing more fully the basis in existing practice). The majority of these practices come from the Americas and were described in the background studies for consultations and workshops convened by the Nansen Initiative in that region.

\textsuperscript{279} G.A. Res. 73/195, \textit{¶¶} 18(l), 21(g) (Dec. 19, 2018). By contrast, the norms relating to disaster displacement are addressed only obliquely by the Global Compact on Refugees. See Rep. of the U.N. High Comm’r for Refugees, \textit{Global Compact on Refugees}, \textit{¶} 63, U.N. Doc. A/73/12 (Sep. 13, 2018).
This Article contributes to the debate on how international law should address international mobility in the context of environmental threats. As a whole, this Article illustrates how greater engagement with different kinds of evidence on the empirical dynamics of movement and existing legal practice by states in one region may shed light on broader questions about the current and future role of international law in shaping this response. The contention that legal scholars working on this aspect of the law should reflect more carefully on the empirical framing of the assumed underlying problem is not new. In this regard, the present Article merely adds new elements to existing scholarship on the implications of the empirical data on environmental threats and human mobility. However, it also makes a more ambitious claim, namely that discussion of the role of international law in this field cannot be divorced entirely from proper consideration of existing legal practice at the national level. In other words, legal scholars interested in how, in cross-border contexts, international law could or should address environmentally displaced persons, climate refugees, or the disaster-displaced—to use only a few contemporary terms—cannot continue to discount national legal practice as if it were irrelevant. What, then, are the principal implications of this Article for our understanding of international law development in response to this perceived legal problem?

On the underlying empirical dynamics, beyond the much-cited risk posed by rising sea levels in the Pacific, the growing evidence base from the Americas confirms that international mobility in this region is already being shaped by sudden-onset events, such as hurricanes, storms, earthquakes, volcanic eruptions and flooding, as well as slow-onset processes, such as desertification and droughts. It has been argued here that, although the concept of disasters is a useful starting point for understanding how these events contribute to the movement of persons, a broader focus on the underlying hazards and hazardous events may better capture the wider ways in which these phenomena can shape mobility, even where a disaster does not result at societal level. In tandem, contrary to assumptions in the legal literature that the resulting movement is an issue only for the Global North, data from this region show that some of this movement also extends to countries in the Global South. As such, this Article contends that the legal practice and perspectives of Global South states need to be incorporated alongside those of states in the Global North in considering the development of law in this field. Moreover, beyond a narrow focus on movement “pushed” by hazards, this Article demonstrates how such hazards, in countries of destination and transit as well as countries of origin, impinge on international mobility issues in other legally-relevant ways. Indeed, the risk is that if we continue to frame the issue as one of extending international protection to persons fleeing disaster-affected countries, we lose sight of the fact that similar legal gaps in relation to travel, entry, and stay
also exist for non-nationals whose mobility is affected by hazards in countries of destination and transit. We also risk stretching refugee law, and wider concepts of international protection, to its breaking point.

On the framing of the legal problem, it is true that international law only tangentially regulates the international mobility related challenges of travel, entry, and stay of non-nationals in the context of environmental threats. However, the existence of a gap in international law does not mean that no law exists, nor does it imply that proposed solutions to the problem can start from a blank canvas. Instead, this Article shows that many states in this region have long recognized the challenges involved and also developed legal responses at the national level to accommodate affected persons. At least for the Americas, this body of legal practice robustly challenges the contention in much of the legal literature that national law can simply be dismissed as irrelevant or as comprised merely of isolated protection provisions in the national law of states in the Global North. Rather, a broadly similar legal approach to these mobility challenges is evident in national laws across this region, including among states in the Global South. On the one hand, this existing practice raises a question about the role of international law, and its added value to the existing response in this region, especially in light of international efforts in some subregions to develop harmonized guidance on such legal practice. On the other hand, given that it is states that are the creators of international law, a better understanding of the ways in which they already see the pertinent challenges, and respond to them in law, offers an insight into how those existing views and practice might influence the development of international law in the future.

In this respect, this Article establishes that most states in the Americas treat the challenges principally as a matter of immigration law rather than international protection. This is not to say that the latter body of law is not applied where environmental events unleash persecution or violence, but simply that it is done by reference to the latter phenomena, rather than the disasters themselves. In general, though, the travel, entry, and stay of persons affected by environmental threats is resolved not by application of the law on international protection—not that on the environment—but by immigration law and cooperation with the state of origin. For persons with a link to the destination country, ordinary migration categories offer a pathway for mobility and a criterion for differentiation among the wider universe of migrants, the importance of which is often overlooked in the legal literature. The fact that states have been prepared to flexibly apply these categories to disaster-affected persons, in contrast to the rigid application of refugee law, indicates how strongly states see immigration law as the appropriate medium of response. Even for those who lack such a link with the destination country, a surprisingly wide range of states have used exceptional migration categories in immigration law, and similar provisions for humanitarian discretion in national refugee law, to facilitate entry and stay for affected persons. In both
cases, the pertinent criterion is usually that the individual is “seriously and directly affected” by the disaster. Immigration law may be largely overlooked by the legal literature, since it is not well-established in international law, but, in practice, it represents the main framework for response by states in the Americas.

What, then, of international law? As this Article demonstrates, at least in the Americas, the problem is not an absence of legal tools. Indeed, in this region, the basic elements of the approach in national law have quite a high degree of consistency. This raises the important question of whether similar legal practice can be discerned in other regions of the world. Yet, for the development of international law, even this discrete body of national legal practice in the Americas raises the prospect that these provisions provide evidence of emerging norms of “regional” custom, and they are already influencing policy at the global level. That consistency is seen also in how the existing national practice has been distilled and elaborated in soft law frameworks at the subregional level as a means to harmonize the approach in participating states. Looking to the future, this suggests that grand proposals for new global treaties on international protection or environmental law to address the legal implications of such mobility are less likely to gain traction with states in this region than efforts to develop the existing approach in international immigration law at the regional or subregional levels. In some forums, incorporating elements of the current approach into free movement arrangements looks like a distinct possibility. Overall, then, it seems that international law may still have a role to play in shaping the response to international mobility linked to environmental factors. However, for better or worse, in the Americas and elsewhere, its future development seems more likely to build on existing state practice than on the flights of fancy of us international law scholars. The law offers few truly blank canvases.

280 The fact that these provisions are rooted in the discretion of states to regulate their immigration affairs may raise a question about whether such legal practice truly reflects opinio juris, as an element of international custom. However, the codification in national law of a power to favorably resolve these types of cases, and its exercise in practice by the state concerned according to the terms of its law, may suggest that states perceive the creation of such powers as a matter of legal obligation. For discussion of the concept of “regional” or “particular,” as opposed to “general” international custom, see Rep. of the Int’l Law Comm’n, at 154–56, U.N. Doc. A/73/10 (2018).

281 This analysis is bolstered by a recent example of practice from outside the Americas region. In February 2020, the Intergovernmental Authority on Development (IGAD), an African subregional interstate forum took precisely this step, endorsing a free movement agreement that integrates specific immigration-based provisions that require states parties to allow citizens of fellow IGAD member states to enter their territory “in anticipation of, during or in the aftermath of disaster,” and to facilitate the extension of stay for such disaster-affected persons while return to their country of origin “is not possible or reasonable.” See Protocol on Free Movement of Persons in the IGAD Region, Feb. 25, 2020, art. 16 (awaiting entry into force).
Human Rights Disclosure and Due Diligence Laws: The Role of Regulatory Oversight in Ensuring Corporate Accountability*

Rachel Chambers† and Anil Yilmaz Vastardis‡

Abstract

The proliferation of human rights disclosure and due diligence laws around the globe is a welcome development in the area of business and human rights. Corresponding improvement in conditions for workers and communities in global supply chains whose human rights are impacted by businesses has not materialized, however. In this Article, we focus on the oversight and enforcement features of human rights disclosure and due diligence laws as one of the missing links to achieving the accountability objectives envisaged by such legislation. Drawing on our analysis of key legislative developments, we observe and critique that the state has almost completely withdrawn itself from the oversight and enforcement roles and assigned these crucial accountability functions solely to consumers, civil society, and investors. Without a regulatory mechanism to ensure quality of human rights disclosures and due diligence processes and to impose sanctions for failing to comply with the laws, not only may the disclosures and processes be inadequate, but there is a danger that misleading disclosures and flawed processes may mask harmful impacts and be detrimental to any hopes of vindicating the rights of workers and communities in global supply chains. We offer a new perspective on a more effective approach to oversight and enforcement in which the state should function as a key actor through which consumers, civil society, and investors can hold businesses accountable.

* An earlier version of this Article was presented at the 4th Annual Conference of the Global Business and Human Rights Scholars Association at New York University in September 2018; at a faculty seminar at the University of Connecticut in February 2019; and at the Innovate Rights Conference at the University of New South Wales in May 2019. The current draft was presented at the Academy of Legal Studies in Business Conference in August 2020. The authors are grateful to Justine Nolan, Shareen Hertel, Stephen Park, and Onyeka Osuji for their helpful comments, and to Gianni Garyfallos for his research assistance.

† Dr. Chambers is the Postdoctoral Research Associate in Business and Human Rights, Human Rights Institute, University of Connecticut, MA (Oxon) LLM (Kent) PhD (Essex). Barrister-at-law (England and Wales).

‡ Dr. Yilmaz Vastardis is a Senior Lecturer at the School of Law and Human Rights Centre, University of Essex, BA (Marmara University, Turkey) LLM (Essex) PhD (Essex).
# Table of Contents

I. Introduction ........................................................................................................................................... 325

II. What Is Accountability in This Context and What Are the Various Tools for Accountability? ........................................................................................................................................ 328

III. Key Features of Human Rights Disclosure and Due Diligence Laws ................................................. 333
    A. General Human Rights Due Diligence and Disclosure ............................................................. 334
    B. Targeted Human Right Due Diligence and Disclosure ............................................................. 337

IV. Benefits of Transparency and HRDD Requirements to Corporate Accountability for Human Rights Impacts ........................................................................................................................................ 344

V. Weaknesses and Pitfalls of this Mode of Market-led Accountability .................................................... 346
    A. Overview ...................................................................................................................................... 346
    B. Market Oversight .......................................................................................................................... 349

VI. The Importance of Regulatory Oversight and the Institutional Options ................................................ 355
    A. Key Functions of Meaningful Regulatory Oversight ................................................................. 356
        1. List and Repository .................................................................................................................... 356
        2. Monitoring Function and the Content of Reports ..................................................................... 357
        3. Enforcement Function .............................................................................................................. 362
    B. Institutional Options and Subject Matter Expertise ...................................................................... 363

VII. Conclusion ......................................................................................................................................... 366
I. INTRODUCTION

In response to intense civil society pressure and increasing public awareness of appalling human rights impacts of businesses, including working conditions amounting to slavery and forced labor,1 several governments have committed themselves to take action to prevent human rights abuses by businesses and to eradicate modern slavery in global supply chains.2 State efforts to date have primarily concentrated on increasing transparency in parent or lead companies.3 The United States and some countries in Europe have adopted legislation that requires companies to make annual public disclosures containing information about their human rights impacts.4 While transparency can fulfill complementary prevention and accountability functions alongside other measures, such as strengthening local trade union involvement and setting clear liability standards to eradicate adverse human rights impacts by business, it can only do so if designed diligently and implemented robustly. The design of existing human rights transparency rules has allowed highly ineffective reporting practices to emerge.5 For instance, a large number of businesses covered under the California Transparency in Supply Chains Act fail to disclose all the required information, and many do not have a disclosure statement at all.6 Human rights disclosures under existing legislation are at best minimal in their reporting of risks and at worst

4 See, e.g., TVPA, supra note 3; MSA, supra note 3.
misleading about human rights impacts across global supply chains and subsidiaries.\(^7\)

There have been various efforts to address the shortcomings of transparency-focused laws, including the adoption of the French Law on Corporate Duty of Vigilance,\(^8\) German businesses’ push for new human rights due diligence laws,\(^9\) and the European Union’s announcement that it will likely enact a due diligence law in 2021.\(^10\) The process of human rights due diligence (HRDD) requires companies to identify, prevent, mitigate, and communicate risks to human rights.\(^11\) If done properly, it can transform corporate behavior and prevent harms.\(^12\) Company reports of human rights impacts and due diligence provide information to stakeholders to enable them to make informed choices about their interactions with the corporation. As such, they act as an important component of an overall corporate accountability framework for human rights impacts. Push for transformation can come from external or internal sources—information found in reports might trigger external stakeholder pressure to transform corporate behavior, or the information gathering and disclosing process might influence internal decision making.\(^13\) If such a transformation materializes, disclosures and HRDD can achieve prevention and mitigation of adverse human rights impacts. But to be able to fulfill these transformation and prevention functions, HRDD must be done properly, and reporting must contain information that is accurate and provides a complete and meaningful picture of the disclosed issues. The current designs of transparency and HRDD laws do not contain adequate safeguards to ensure duties are carried out properly and in accordance with the relevant legislation.

\(^7\) See, e.g., CORE COAL., supra note 5.


\(^10\) See CORE COAL., 25 NGOS AND TRADE UNIONS CALL FOR A UK LAW ON MANDATORY HUMAN RIGHTS DUE DILIGENCE (2019); EU to Legislate for Human Rights and Environment Due Diligence, HAUSFELD LLP (2020), https://perma.cc/5TB9-V7DE.


\(^13\) See Buhmann, supra note 12.
Most worryingly, the transparency and due diligence rules discussed in this Article typically require minimal substantive disclosure and largely adopt only “non-coercive enforcement,” which leaves the watchdog role almost exclusively to consumers, investors, and NGOs. Without a regulatory oversight mechanism to ensure the quality of disclosures and to impose sanctions for misleading information, not only may the reporting be inadequate, but also the disclosures may present false realities and be detrimental to any hopes of improving rights of workers and communities in the global supply chain. The question of oversight and enforcement, however, remains underexplored in the literature.

This Article makes a novel contribution to the ongoing discussions about how to improve human rights disclosure and due diligence laws to achieve the stated legislative aims. It has been observed and critiqued that the oversight and enforcement features of these laws remain weak or nonexistent. Taking this as our starting point, we first offer a new framing of the problem by analyzing how these laws assign oversight and enforcement roles between the state and the market. We critique that the state has almost completely withdrawn itself from the oversight and enforcement roles and assigned these crucial accountability functions primarily to consumers, civil society, and investors. As discussed below, even in the case of the French Law, civil society acts as the main driver of oversight. We argue in favor of greater state involvement in transnational business regulation as it concerns the human rights impacts of businesses. We argue that state-based enforcement and oversight constitute necessary ingredients for public disclosures and HRDD to contribute to the improvement of human rights conditions in supply chains and to achieve accountability for adverse business impacts on human rights. We offer a new perspective on a more effective approach to oversight and enforcement that distinguishes human rights disclosures and due diligence from traditional corporate reporting and due diligence. Our approach assigns these functions primarily to a public authority that should have expertise in both corporate governance and human rights and would also function as an enabler for consumers, civil society, and investors to hold businesses accountable.

The Article begins in Section II by unpacking the concept of accountability in the context of business impacts on human rights. Next, Section III outlines and compares the HRDD and reporting requirements under key examples of the relevant legislation: the E.U. Non-Financial Reporting Directive (enacted 2014);

---


15 See Tim Bartley, *Rules Without Rights: Lands, Labor, and Private Authority in the Global Economy* 7, 12, 40 (Oxford, 2018) (noting that corporate codes of conduct and private social or ethical audits are the prime examples of this mode of governance).

16 See, e.g., Nolan, supra note 5, at 68–71; LeBaron & Rühmkorf, supra note 5.
the French Law on the Corporate Duty of Vigilance of 2017; the California Transparency in Supply Chains Act of 2010 (CTSCA); the U.K. Modern Slavery Act 2015 (MSA); the Australian Modern Slavery Act 2018 (AMSA); the U.S. Dodd-Frank Act of 2010 § 1502; the E.U. Conflict Minerals Regulation (enacted 2017); and the Dutch Child Labor Law (enacted 2019). In doing so, the Article briefly addresses the reasons behind the enactment of the laws, the transparency and HRDD requirements they contain, and the institutional arrangements supporting these requirements, in order to give the reader context for the discussion that follows.

Building on the existing scholarship, the analysis in Sections IV and V frames the approach of the existing legislation to oversight and enforcement as market-led and critiques this model of oversight and enforcement. In Section VI, we argue that enforcement of these laws via regulatory oversight, alongside market oversight, is essential for achieving their accountability objectives. We then evaluate the options for regulatory oversight of HRDD and reporting. Here, we analyze the distinguishing features of human rights reporting from financial and other types of non-financial reporting.

We urge policymakers to move away from placing HRDD and reporting within the realm of traditional corporate reporting and instead to adopt a sui generis model of oversight that marries corporate reporting expertise with human rights expertise.

II. WHAT IS ACCOUNTABILITY IN THIS CONTEXT AND WHAT ARE THE VARIOUS TOOLS FOR ACCOUNTABILITY?

Corporate accountability represents a movement away from the voluntarism and self-regulation that characterize contemporary corporate social responsibility (CSR). For Peter Newell, “the term [accountability] implies both a measure of answerability (providing an account for actions undertaken) and enforceability (punishment or sanctions for poor performance or illegal conduct).” In the context of business impacts on human rights, we conceive of accountability as a wider concept than liability, encompassing the idea that companies should be held responsible for the consequences of their actions via non-legal accountability (risk of loss of reputation, denial of access to foreign markets, fall in share price, and

---


18 Other types of non-financial reporting include diversity, governance, and environmental performance.


shareholder dissent) as well as legal accountability through regulators or courts of law. We concur with Simon Zadek’s common-sense opinion that for accountability to work, there must be “capability to do something about [a given action and] . . . some level of consensus about the action that needs to be accounted for and the penalties for failure . . . [Additionally, there needs to be] a reasonably well-defined ‘community’ that can reach such decisions.”

In following Zadek’s criteria, one must first determine what needs to be accounted for in the context of transparency and due diligence for human rights abuses in a company’s supply chains and subsidiaries. Second, one must determine what consequences and penalties may attach to performance that falls below the expected or required standard and what processes are needed to assess compliance and impose sanctions.

Determining the scope of accountability can be a complex task where the disclosure obligations are placed on parent or lead companies even though the presence of human rights abuses in those companies’ subsidiaries or supply chains may not be contrary to any legal requirement or obligation placed on the parent or lead company. The corporate law concepts of corporate personality, limited liability, and the contractual nature of relationships with suppliers insulate parent or lead companies from liability for harm caused by subsidiaries and suppliers. Therefore, at least in theory, legal liability falls on the subsidiary, supplier, or sub-supplier that is directly linked to the human rights harm rather than the lead company that sells the end product. For instance, the MSA imposes criminal liability only for slavery and human trafficking that take place within the U.K. Accountability arises from the presence of slavery within a business organization. By contrast, supply chain accountability for slavery and human trafficking overseas through the Act is achieved through disclosure of the steps taken to identify and eliminate these practices. Unlike the former type of criminal accountability that applies in the domestic context, the latter obligation is neither

24 However, it is still possible for a parent or lead company to be directly liable for the harms suffered as a result of a subsidiary’s acts or omissions. See Vendanta Resources v. Lungowe [2019] UKSC 20 (appeal taken from Eng.); Chandler v. Cape [2012] EWCA ( Civ) 525 (appeal taken from Eng.). The French Law places a legal requirement on the parent company to conduct due diligence in its supply chain. See Loi 2017-399, supra note 8.
25 See MSA, supra note 3, §§ 1–2.
26 Id. § 54.
an obligation of result nor an obligation of due diligence to eliminate slavery within supply chains. Instead, it is an obligation to report the steps, if any, taken to eliminate slavery. The presence of human rights abuses and modern slavery practices legally distant from the parent or lead company renders it difficult to define what action needs to be accounted for in modern slavery or human rights disclosures, particularly if a legally-mandated due diligence obligation is lacking.

At the outset, the conduct that needs to be accounted for is the failure on a parent or lead company’s part to make diligent efforts to identify, prevent, mitigate, and eliminate human rights abuses in its subsidiaries and supply chains and to disclose adequately the steps taken.27 This is just one piece of the wider puzzle of accountability. We are not looking at other pieces of the puzzle such as the role of state oversight and enforcement through civil or criminal liability for human rights harms. Relatedly, we are not looking at this from the perspective of access to remedy. Our concern is whether the current disclosure and HRDD frameworks improve accountability: (1) to victims of business human rights impacts and (2) to those affected indirectly as consumers, customers, or investors who buy from and/or invest in the company on the understanding that it is doing all that can be reasonably expected to prevent and remediate human rights abuses and modern slavery in its global supply chains.

There is a spectrum of regulatory approaches to bringing about corporate accountability in terms of the consequences that will attach to poor performance, ranging from light-touch (private-led regulation) to stringent regulation with binding standards enforced by public authorities.28 Genevieve LeBaron and Andreas Rühmkorf observe that home state regulation on business and human rights has been “enacted through a range of different institutional designs that combine elements and instruments of public and private governance.”29 That said, as will be seen in Section III, a heavy emphasis has thus far been placed on various degrees of transparency accompanied by market-centered accountability mechanisms at the light-touch end of the spectrum. This mode of regulation is a move away from traditional “command and control” regulation, in which governments adopt “legal rules backed by [civil or criminal] sanctions,”30 toward

27 This accords with the due diligence requirements of the U.N. See UNGPs, supra note 11, at 21–22.

28 See LeBaron & Rühmkorf, supra note 5, at 17–19, 26. Compare Bribery Act 2010, c. 23, §§ 1–20, sched. 1–2 (U.K.), with MSA, supra note 3 (highlighting that, under these categories, the former produced significant changes in corporate practice while the latter has not).

29 See LeBaron & Rühmkorf, supra note 5, at 17.

what is termed “reflexive” regulation, or “New Governance.” Through this mechanism, the government acts as “the orchestrator of private actors to encourage compliance” and attempts to “influence normative practices indirectly by shaping the context in which society’s various actors and subsystems interact and bargain with one another.” This model has been widely supported by the CSR literature due to its promise of effecting organizational and lasting change, whereas command-and-control type regulation has been viewed with skepticism due to its potential to produce a tick-box approach to human rights issues. Human rights disclosure laws discussed in this Article largely adopt this light-touch regulation model based on a market-led model of accountability. Recent developments suggest a slow, gradual movement toward more stringent regulation, with a new legislative approach featuring a legal duty to conduct HRDD and to publish HRDD information backed by certain penalties and civil liability for failure to comply. Karin Buhmann has argued that for the light-touch approach to be successful, it needs to properly encourage organizational learning and not merely focus on penalties for non-disclosure. While the organizational learning focus is crucial, decades of voluntarism and soft regulation in this field have not produced successful outcomes when the bottom line of business remains profit oriented. One reason for this lack of meaningful progress is the lack of stringent legal accountability mechanisms to push businesses to take disclosure


33 See Nolan, supra note 5, at 70.

34 See Doorey, supra note 30, at 357.

35 See Hess, supra note 32.


38 See Buhmann, supra note 12, at 39.

and HRDD obligations seriously. It has been argued by critics of the current transparency rules that without, as a minimum, an accompanying HRDD obligation and civil, administrative, or criminal liability for failure to comply, these rules cannot effectively contribute to corporate accountability.\(^{40}\) The analysis in the latter parts of this Article discusses the necessity for state-based oversight and enforcement supported by stakeholders as an essential ingredient for the efficacy of any laws, whether they comprise transparency obligations only or they include the additional requirement of HRDD. This is crucial to ensure the avoidance of a disconnect between what is reported in corporate disclosures and the actual human rights situation on the ground.\(^{41}\)

Both transparency and HRDD are requirements of the U.N. Guiding Principles on Business and Human Rights (UNGPs).\(^{42}\) The UNGPs, which were adopted unanimously by the U.N. Human Rights Council in 2011, represent a consensus of opinion among a number of states, companies, and non-governmental organizations, about the human rights responsibilities of corporations. The UNGPs expect business actors to “operationalize” their responsibility to respect human rights through HRDD and reporting processes.\(^{43}\) Businesses are expected to communicate the steps they take to address human rights impacts by publishing sufficiently-detailed information on the impacts and steps taken to prevent, mitigate, and remediate these in appropriate form and frequency.\(^{44}\)

The commentary on reporting explains “showing” that businesses respect human rights involves companies communicating and “providing a measure of transparency and accountability to individuals or groups who may be impacted and to other relevant stakeholders, including investors.”\(^{45}\) Thus, reporting under the UNGPs is not an end in itself but, theoretically at least, is an exercise that will provide a measure of accountability.\(^{46}\) Reporting alone is not sufficient, however, and the reporting provisions in the UNGPs are supplemented by a requirement that companies conduct HRDD. This process entails identifying whether they

\(^{40}\) See LeBaron and Rühmkorf, supra note 5; Nolan, supra note 5; U.K. HOUSES OF PARLIAMENT JOINT COMM. ON HUM. RTS., supra note 5.

\(^{41}\) For an analysis of such a disconnect, see Madhura Rao & Nadia Bernaz, Corporate Responsibility for Human Rights in Assam Tea Plantations - A Business and Human Rights Approach, SUSTAINABILITY 2020, 12, 7409.

\(^{42}\) See UNGPs, supra note 11.

\(^{43}\) See id. (noting that UNGPs 16–24 are the “operational principles” in relation to the corporate responsibility to respect pillar).

\(^{44}\) Id. For an exploration of reporting under the UNGPs, see UN Guiding Principles Reporting Framework, SHIFT & MAZARS (2015), https://perma.cc/XZV8-R7MB.

\(^{45}\) UNGPs, supra note 11, at 20.

have caused or contributed to adverse human rights impacts, integrating and acting upon the findings, tracking responses, and remediating the harm if an adverse impact has occurred. The enactment of transparency and HRDD laws are illustrative of the polycentric governance described in the UNGPs.

The knowing-and-showing approach of the UNGPs has been a welcome shift from the naming-and-shaming approach. We argue here that for the “showing” to be reliable and not misleading, there needs to be a reliable system of oversight. While the UNGPs acknowledge the role of transparency in achieving corporate accountability, Principle 3 and its commentary leave it up to each state to determine the type of transparency measures to be introduced. According to Principle 3, the states should encourage or require businesses to be transparent on how they address their human rights impacts. As far as transparency measures go, the expectations of the UNGPs could be fulfilled by states introducing or maintaining the “light-touch” regulations and mandating reporting without any follow-up measures. The UNGPs place no clear expectations on states to introduce robust measures of oversight and enforcement for transparency requirements. We argue in this Article that such measures are a crucial element of transparency and HRDD laws to ensure corporate accountability.

III. Key Features of Human Rights Disclosure and Due Diligence Laws

We divide the legislation mandating disclosure and/or HRDD into two main categories: (1) general HRDD and disclosure laws and (2) laws that target a specific human rights issue. The key features of the different laws are outlined below, with particular focus on two distinct elements: the type of disclosure required and the processes in place (or the lack thereof) to ensure accessibility and accuracy of the disclosures.


48 Buhmann, supra note 12, at 39.

49 See UNGPs, supra note 11, at 8–9.

A. General Human Rights Due Diligence and Disclosure

In this section, we focus on the E.U. Non-Financial Reporting Directive (E.U. NFRD) and the French Law on the Corporate Duty of Vigilance. The E.U. NFRD applies to all companies of a certain size governed by laws of individual E.U. Member States. These companies are required to report on human rights and related matters “to the extent necessary for an understanding of the undertaking’s development, performance and position and of the impact of its activity.” Companies must additionally disclose HRDD processes implemented by the company in pursuing policies related to non-financial matters, the corresponding outcome, and the principal risks arising in connection with the company’s operations, including how the company manages these risks. The provision adopts a “comply or explain” approach, meaning that companies can elect to comply with it either by making the required disclosures or by providing an explanation for why they have elected not to do so. The E.U. NFRD leaves it up to each E.U. member state to determine whether to require verification of reports by an independent assurance service provider and whether to have a sanctions regime for companies that fail to report adequately. At a minimum, each member state is to require checks by an auditor for the existence of a report. Some member states have implemented legislation requiring these disclosures to be made in the management report, and some have imposed additional checks. Thus, member states vary in the checks they have in place, but it has largely fallen to the “market” to oversee reporting.

52 Directive No. 2014/95, supra note 51. These are certain large companies and qualifying partnerships with more than 500 employees.
53 Id.
54 Id.
56 According to a study reviewing how member states have transposed the directive into their national law, twenty member states only require the existence of the reports to be verified and not the content. The states are Austria, Croatia, Cyprus, Czech Republic, Estonia, Finland, Germany, Greece, Hungary, Ireland, Lithuania, Luxembourg, Malta, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, and Norway. See GLOBAL REPORTING INITIATIVE, MEMBER STATE IMPLEMENTATION OF DIRECTIVE 2014/95/EU: A COMPREHENSIVE OVERVIEW OF HOW MEMBER STATES ARE IMPLEMENTING THE EU DIRECTIVE ON NON-FINANCIAL AND DIVERSITY INFORMATION 1, 16–31 (2017).
57 See id. Eight states require consistency checks with the management report: Belgium, Bulgaria, Denmark, Latvia, Netherlands, Romania, U.K., and Iceland.
58 For instance, Denmark. This subject is taken up in Section VI below.
In France, the French Law on the Corporate Duty of Vigilance requires companies meeting the threshold requirements for size to create and implement an annual “vigilance plan” aimed at identifying and preventing human rights violations in both their domestic and their international operations, including those associated with their subsidiaries and supply chain. The first plans were published in 2018. The development and the publication of the plan and a report on its implementation are among the substantive obligations prescribed by the “duty of vigilance.” The plan must set out the steps that the company will take to detect risks and prevent serious violations with respect to human rights and fundamental freedoms, health and safety, and the environment. This includes mapping out and analyzing the risks and putting measures in place to mitigate risks and address negative impacts, including an alert mechanism and a monitoring scheme to follow up on the plan’s implementation. Unlike most of the other laws discussed in this Article, the French law’s transparency requirement can only be fulfilled by complying. There is no room for compliance by explaining why no steps have been taken because taking the prescribed steps is a fundamental obligation placed on companies covered under the law. Since the development, implementation, and communication of the plan together constitute the vigilance duty, the French law brings together the HRDD and transparency elements of the UNGPs’ second pillar previously discussed in Section II.

Compliance with the law is established through a court process whereby companies can be legally compelled—at the request of a party with standing, including an NGO or a trade union—to create and implement an adequate vigilance plan. Prior to the initiation of a court process, companies will be given a three-month period to comply with the requirements of the law. Periodic penalties may be imposed by the court if companies are found to be failing their vigilance obligations. To date, a small number of notices have been served to companies, at the initiative of civil society organizations (CSOs), on the basis of

59 The law applies to any company registered in France that has (a) 5,000 or more employees, including employees of its direct or indirect French-registered subsidiaries; or (b) 10,000 or more employees, including employees of its direct or indirect French-registered or foreign subsidiaries. See Loi 2017-399, supra note 8, art. 1. It is estimated that the law applies to about 150 companies. See Anna Triponel & John Sherman, Legislating Human Rights Due Diligence: Opportunities and Potential Pitfalls to the French Duty of Vigilance Law, INT’L. BAR ASS’N (2017), https://perma.cc/SX59-5N5K.

60 See Loi 2017-399, supra note 8, art. 1, ¶ 3 (covering the companies that the company controls directly or indirectly and, moving down the supply chain, the activities of its subcontractors and suppliers “with which [it] maintains an established commercial relationship”).


62 Loi 2017-399, supra note 8.

63 Id. ¶¶ 4–9.

64 Id. ¶¶ 7–9; Brabant & Savourey, supra note 61, at 4.
inadequate vigilance plans, and two of these incidents have proceeded to the courts at the end of the three-month notice period. Accountability also takes place through a process by which victims who have been harmed by a company covered by the legislation can claim damages for negligence through an ordinary civil lawsuit, using the company’s noncompliance with the vigilance obligation as evidence of its wrongdoing.

Combining HRDD and transparency backed up with sanctions, the French law is the most promising piece of legislation presently in force to advance corporate accountability. It moves away from exclusively relying on market-led oversight and enforcement of the law. It is not without shortcomings, however. It has been highlighted that the law’s threshold for coverage is very high, the sanctions available are weak in terms of remediating harms, and, most importantly for the purposes of this Article, there is a lack of governmental monitoring and oversight for compliance by covered companies.

While the French law takes a crucial step by attaching sanctions to the vigilance obligations, in the absence of a state-initiated oversight mechanism it is left to the “market,” typically CSOs, to monitor companies’ compliance and initiate the complaints procedures available under the law. The lack of an official list and repository for vigilance plans render it challenging for CSOs, trade unions, and other stakeholders to identify shortcomings and take part in the enforcement of the obligations. A preliminary proposal has been put forward to address this by designating certain individuals within the French administration to look into which companies are within the scope of the law, how the law is implemented, and whether some provisions of the law need to be clarified. A need for formal oversight and more robust enforcement constitutes the main focus of the analysis presented in the later sections of this Article.


67 FORUM CITOYEN POUR LA RSE, LAW ON DUTY OF VIGILANCE OF PARENT AND OUTSOURCING COMPANIES YEAR ONE: COMPANIES MUST DO BETTER 1, 8 (Juliette Renaud et al. eds., 2019).

68 Brabant & Savourey, supra note 61, at 2–4.


70 Brabant & Savourey, supra note 61.

B. Targeted Human Right Due Diligence and Disclosure

Targeted HRDD and transparency legislation focuses on specific issues such as modern slavery, child labor, or conflict minerals. In this Section, we will first examine laws mandating disclosure on modern slavery. The design of the three statutes dealing directly with modern slavery is similar, although each legislative scheme has slight variation. We will then examine a law mandating transparency on conflict mineral due diligence, and lastly a law mandating due diligence on child labor. A common theme relevant for the purposes of this Article is that these laws rely on market-led oversight to ensure compliance with the law, to a greater or a lesser extent.

The CTSCA, which came into force in January 2012, requires certain large retail sellers and manufacturers doing business in the state of California to disclose their efforts to eradicate slavery and human trafficking from their direct supply chain for tangible goods offered for sale.\(^72\) The disclosed information should be posted on the retail seller or manufacturer’s website. Specifically, a company to which the legislation applies must disclose to what extent, if any, it: verifies product supply chains to evaluate and address risks of human trafficking and slavery; conducts audits of suppliers to evaluate supplier compliance with company standards for trafficking and slavery in supply chains; requires direct suppliers to certify that materials incorporated into the product comply with the laws regarding slavery and human trafficking of the country or countries in which they are doing business; maintains internal accountability standards and procedures for employees or contractors failing to meet company standards regarding slavery and trafficking; and provides relevant training.\(^73\) There is no requirement to update the report on a periodic basis. The California Franchise Tax Board produces an annual list of companies covered by its provisions based on information from tax returns.\(^74\) There is, however, no official repository where these reports must be deposited for public access. While the presence of an official list is crucially important for stakeholders to identify the companies covered by the law, the lack of a central repository for accessing the reports renders it challenging for stakeholders to identify companies failing to comply with the law. A large number of covered businesses fail to disclose information on all the required areas of activity, and many do not have a disclosure statement.\(^75\) The only

\(^72\) CTSCA, supra note 6.

\(^73\) Id. § 3(b).

\(^74\) KAMALA HARRIS, THE CALIFORNIA TRANSPARENCY IN SUPPLY CHAINS ACT: A RESOURCE GUIDE 3 (2015). Note, however, that the CTSCA does not require companies to report on an annual basis, meaning that companies can comply with the law by reporting just once.

\(^75\) See Barna, supra note 6, at 1463 (noting that only sixty-two percent of covered companies disclosed).
state-based relief under the CTSCA for failure to report is injunctive. Following a compliance review in 2015, the California Department of Justice took steps to improve compliance with the Act by writing to companies and asking them to provide either an explanation of why the legislation does not apply to them or a link to a compliant disclosure. To date, the Attorney General of California has not yet brought an action against a corporation for nondisclosure under the Act. From the foregoing, it is not possible to conclude that the CTSCA provides robust oversight and enforcement of the transparency obligations, leaving the “market” to exercise checks and hold businesses accountable.

In the U.K., Section 54 of the MSA introduced a modern slavery and human trafficking transparency requirement for certain “commercial organizations” with a turnover of at least £36 million that “carry on” business in the U.K. The law adopts a comply-or-explain approach by requiring companies to publish either a statement of the steps the organization has taken to ensure slavery and human trafficking is not taking place in any of its supply chains or in any part of its business, or a statement that the organization has taken no such steps. The MSA does not prescribe specific content for the disclosures, but provides a nonexhaustive list of items that may be included in the “slavery and human trafficking statement.” These items include information about a commercial organization’s policies and due diligence processes in relation to slavery and human trafficking in its business and supply chain; the parts of its business most at risk of slavery; and human trafficking and steps put in place to assess and manage that risk, including performance indicators for the success of these steps. The statement must be approved by the board of directors of a limited company or all members of a limited liability partnership. The MSA covers steps taken in “any of [a corporation's] supply chains,” a broader requirement than that in the

76 CTSCA, supra note 6, § 3(d).
78 Memorandum from Jena Martin on Policy Options for Addressing and Preventing Forced Labor, Modern Slavery, and Human Trafficking in Supply Chains (May 2020).
80 CORE COAL., supra note 5, at 3 (estimating that Section 54 covers between 12,000 and 17,000 companies).
81 See MSA, supra note 3, § 54(12) (defining a commercial organization as it as a body corporate or partnership “which carries on a business, or part of a business, in the United Kingdom”).
82 Id. § 54(5).
83 Id. § 54(6).
CTSCA, which covers the direct supply chain only. The MSA also requires yearly updates, in contrast to the CTSCA’s one-off approach. However, of companies that are required to report, meaning that stakeholders are not able to verify if a company is covered by the law or not, although the U.K. Home Office has written to 10,000 companies to which it believes the law may be applicable. Nor is there an official database where such reports are deposited. If a business falling under section 54 fails to report, the Secretary of State may bring court proceedings for injunctive relief. As of yet, there has been no instance of this happening in practice. There is no other means of enforcement or oversight in the Act. Furthermore, there are no mechanisms to ensure the accuracy of the report contents (though the report contents can be so vague under this law that there seems little need for assurance). Thus, oversight and enforcement are left to the “market.”

There have already been four reporting cycles under the MSA. Generally, however, disclosure has been of a low standard, not always meeting even the minimum requirements of the Act including approval of the statement by senior management and visibility on the company website. Many companies have not reported at all. Acknowledging deficiencies of the MSA, an independent review of the Act published in 2019 recommended to the U.K. Government to: abandon the comply-or-explain approach; adopt a comply approach with prescribed minimum content for the report; create a repository for statements; establish a monitoring and enforcement mechanism; and strengthen sanctions for failure to comply.

84 *Id.* § 54(4)(a)(ii). Commercial organizations caught within the definition are not, however, required to report on all the supply chains in their groups overseas, such as those of wholly owned foreign subsidiaries. See Parasha Chandran, *A Loophole in the Slavery Bill Could Allow Companies to Hide Supply Chain Abuses*, *The Guardian* (Mar. 24, 2015), https://perma.cc/R2WL-J8EY.
85 MSA, supra note 3, § 54(4)(a).
86 See Section VI.A.1 for a discussion on the importance of a formal list.
87 COMM. OF PUB. ACCT., supra note 79, at 10.
88 *See About Us, Modern Slavery Registry*, https://perma.cc/YA5W-56KU.
89 MSA, supra note 3, § 54(11).
90 MSA, § 54(6). Depending on the type of entity, senior management could consist of the board of directors, members, partners, or a general partner.
91 *See* BUS. AND HUM. RTS. RES. CTR., FTSE 100 AT THE STARTING LINE: AN ANALYSIS OF COMPANY STATEMENTS UNDER THE UK MODERN SLAVERY ACT (2016); CORE COAL., supra note 5; BUS. & HUM. RTS. RES. CTR., FIRST YEAR OF FTSE 100 REPORTS UNDER THE MODERN SLAVERY ACT: TOWARDS ELIMINATION? (2017).
92 CORE COAL., WRITTEN EVIDENCE SUBMITTED FROM CORE ¶ 8 (2018) (noting that sixty percent of companies that are covered by section 54 have failed to produce a report).
Like the CTSCA and the MSA, the Australian Modern Slavery Act 2018 (AMSA) requires companies that meet a prescribed size threshold to report on the risks of modern slavery in their operations and supply chains and actions taken to assess and address those risks. This law abandons the comply-or-explain approach adopted in the U.K. MSA. It is mandatory for companies to provide the particular information on policies and processes to detect and address modern slavery listed in the Act, meaning that companies cannot be selective about what to report. The AMSA also makes provision for a government-funded central repository for slavery and human trafficking statements, but it does not penalize companies for noncompliance, though the Minister for Home Affairs can make an inquiry if a company has not complied. If a company fails to respond, the minister may publicly disclose information about the company’s failure to comply. The law was passed in November 2018, and the first disclosure under the Act was made in 2019. In terms of “enforcement,” the AMSA creates a mechanism through which noncompliant entities can be asked to explain and remedy their failure to report, or they risk being named on the government-maintained register. The relevant minister reports to Parliament annually on compliance trends, enabling oversight of overall compliance patterns by Parliamentarians. Otherwise, as with the other modern slavery reporting laws, oversight and enforcement is left to the “market.”

Outside the arena of modern slavery, in the U.S., the Conflict Minerals Rule, adopted in 2012—with the first reports filed in 2014—requires companies reporting to the U.S. Securities and Exchange Commission (SEC) to conduct due diligence and to report on the sourcing of certain minerals (tin, tungsten, tantalum, and gold). These companies must make reasonable and good faith efforts to determine whether the specified minerals used in the manufacture of their products originated in the Democratic Republic of Congo or its neighboring countries. The companies must also disclose their determinations and describe their country of origin inquiries to the SEC and on their company websites. Where (recommendating that the government should consider publishing a list of companies that have complied and not complied with the legislation, rather than falling back on civil society to undertake this work). The government has responded to the report, agreeing to this recommendation. See HM TREASURY, TREASURY MINUTES: GOVERNMENT RESPONSE TO THE COMMITTEE OF PUBLIC ACCOUNTS ON THE THIRTY FIRST TO THE THIRTY SEVENTH REPORTS FROM SESSION 2017–19, 2018, Cm. 9634, 23 (U.K.).

94 See Modern Slavery Act 2018 (Cth) s 5 (Austl.) (requiring companies to report if they carry on business in Australia with a minimum annual consolidated revenue of AU$100 million).

95 Id. ss 18–20.

96 Id. s 16A.

this inquiry reveals that the minerals did originate in these countries, the company must exercise due diligence on the source and chain of custody of the mineral, in accordance with a nationally or internationally recognized due diligence framework. Where the due diligence confirms the company’s determination, it must file a Conflict Minerals Report with the SEC and post the same on its website. There is no list of companies that must comply with the Conflict Minerals Rule, but an annual Government Audit Office report to congressional committees examines how companies responded to the Conflict Minerals Rule in the previous calendar year and analyzes a generalized random sample of company reports. In terms of enforcement and sanctions, the Conflict Minerals Rule imposes penalties on companies for not reporting or complying in good faith. Form SD (the form used for submitting the disclosure) is deemed filed under the Securities Exchange Act of 1934 and subject to § 18 of the Exchange Act, which attaches liability for any false or misleading statements. There has, however, been no SEC enforcement action against companies for failure to comply with the Conflict Mineral Rule, despite the mixed record of engagement and compliance among companies. In the absence of enforcement by the regulator, the oversight and enforcement functions are left to the “market.”

The E.U. has also passed a disclosure law aimed at supply chain due diligence for the use of conflict minerals. This is company law, not securities law, “laying down supply chain due diligence obligations for [E.U.] importers of tin, tantalum, and tungsten, their ores, and gold originating from conflict-affected and high-risk areas.” The Regulation will enter into force in 2021. The geographical scope of the E.U. Regulation is broader than § 1502 of the Dodd-Frank Act, targeting imports not only from conflict zones and areas where a risk of armed confrontation exists but also from failed states and areas where widespread and

100 See Public Statement, U.S. Securities and Exchange Commission, Updated Statement on the Effect of the Court of Appeals Decision on the Conflict Minerals Rule (Apr. 17, 2017), https://perma.cc/FZ67-VK3B (indicating that the staff of the SEC would not recommend enforcement action if companies only file a Form SD, and not a Conflict Minerals Report). Companies are being encouraged to continue to file Conflict Minerals Reports, and many continue to do so. See id.
103 Id.
systematic violations of international law, including human rights abuses, occur.\textsuperscript{104} Nationally, implementing the Regulation depends on the responsible authorities designated by E.U. member states.\textsuperscript{105} These authorities should conduct ex-post checks on how E.U. importers comply with the Regulation.\textsuperscript{106} This includes audits of records as well as on-the-spot inspections.\textsuperscript{107} The Regulation has been criticized for its lack of sanctions.\textsuperscript{108} Member states set the rules that apply to infringements of the Regulation. When an infringement occurs, the competent authorities issue a notice of remedial action to be undertaken by the company.\textsuperscript{109} Whether compliance will be achieved without penalties for failure to take remedial action remains to be seen.

The only example of a targeted HRDD law is the Dutch Child Labor Due Diligence Act, which was approved by the Dutch Senate in 2019, and is yet to go into effect.\textsuperscript{110} Like the French law, the Act brings together the HRDD and transparency elements of the UNGP’s second pillar by pushing companies to examine their supply chains for child labor, act upon their findings, and report that they have done so. Specifically, the Act requires all companies that supply goods or services to Dutch end-users to issue a declaration that HRDD is conducted to prevent child labor from being used in the production of goods and services.\textsuperscript{111} In order to make the requisite declaration, it is implicit that the company must conduct the necessary HRDD. Should the HRDD give the company a reasonable suspicion of child labor in the production of the company’s goods or services, it must adopt and implement a plan of action to address this.

Once the obligation is in place, a new regulator \([\text{toezichtboeder}]\) will be created that will publish the corporate human rights due diligence statements in an online

\begin{footnotesize}
\begin{enumerate}
\item No definitive list of “conflict-affected” or “high-risk” countries has been published yet, and E.U. importers are encouraged to make this assessment themselves based on non-binding guidelines issued by the European Commission. \textit{See} Commission Recommendation (EU) No. 2018/1149 of 10 Aug. 2018, 2018 O.J. (L 208) 94.
\item \textit{EUR. COMM’N, Conflict Minerals: List of Member State Competent Authorities Designated Under Article 10(1) of Regulation (EU) 2017/821 (Jan. 9, 2020).}
\item \textit{See} Regulation 2017/821, supra note 102, art. 10–11.
\item Sascha Arnold, \textit{The EU Conflict Minerals Regulation – New Due Diligence Requirements for Importers}, FRESHFIELDS BRUCKHAUS DERINGER LLP (June 5, 2020), https://perma.cc/44FH-VTPM.
\item \textit{See} Regulation 2017/821, supra note 102, art. 16.
\item The expectation is that the Act will become effective sometime in 2022. The three-year period between the Act’s approval and it going into effect would give the government time to prepare a General Administrative Order that appoints the regulator and fleshes out the obligations of companies under the Act in more detail, see \textit{Dutch Child Labor Due Diligence Act Approved by Senate – Implications for Global Companies}, ROPES & GRAY (June 5, 2019), https://perma.cc/NS4E-BZZB.
\item Id.
\end{enumerate}
\end{footnotesize}
There will not be a formal list of companies that must comply with the law, however, meaning that third parties cannot check if a company is covered or not. Affected third parties such as victims cannot sue companies under the Act, but they can submit complaints that may trigger enforcement by the regulator. Any individual or entity wishing to submit a complaint must first submit the complaint to the company itself. If the company’s reaction is “inadequate” according to the complainant, and on the basis of concrete evidence of non-compliance with the Act, a complaint can be filed with the regulator. A company can be fined up to €8,200 for failing to submit a statement declaring that it exercises due diligence. If a company fails to carry out due diligence in accordance with the Act or to draw up a plan of action, or to comply with any further requirements that are established pertaining to due diligence and the plan of action, a fine of up to €870,000 or 10% of the worldwide annual turnover of the company can be imposed. Thus, in terms of regulatory oversight, the Act provides the most comprehensive oversight among the laws discussed. However, the scheme still has gaps: in particular, the Dutch authorities will not actively enforce the law except in response to a third-party complaint, meaning the law relies on the watchdog role of civil society to ensure its effectiveness.

---


113 See Anneloes Hoff, Dutch Child Labor Due Diligence Law: A Step Towards Mandatory Human Rights Due Diligence, OXFORD HUMAN RIGHTS HUB (June 10, 2019), https://perma.cc/TLY5-34GW. Plaintiffs under general Dutch tort law, nevertheless, would still be able to rely indirectly on the Act if the violation of the Act by the company could be construed as an indication of an act contrary to a duty of care to society. Where the compliance officer breaches their obligations, such as by a violation of the implementation of a due diligence process that causes serious bodily harm, the compliance officer themselves incur personal criminal liability. This can be punishment of a maximum of two years’ imprisonment and a €20,500 fine. Lise Smit et al., Study on Due Diligence Requirements Through the Supply Chain 211 (2020).


115 Ropes & Gray, supra note 110.


IV. BENEFITS OF TRANSPARENCY AND HRDD REQUIREMENTS TO CORPORATE ACCOUNTABILITY FOR HUMAN RIGHTS IMPACTS

Despite their shortcomings, which are discussed in later sections, the current transparency and HRDD requirements do move the legal framework closer toward bringing human rights standards to bear on corporate activities. The most obvious positive impact is that these laws place human rights on the corporate agenda at the highest levels of management for the covered businesses. Placing an expectation on companies to consider at the board level and, ideally, engage in the issues that external reporting and HRDD raise may influence internal business decisions which produce adverse human rights impacts.

Corporations may be prompted to monitor and change their own behavior, as well as to push for change in supplier practices. The different degrees of expectations placed on companies by different types of legislation will influence the extent of the positive changes (if any) that may occur. Whereas a mandatory HRDD law such as the French Law on the Corporate Duty of Vigilance may achieve a greater commitment from the businesses covered and more substantial change on the ground for those adversely impacted, a light-touch disclosure law, such as the U.K. MSA, is less likely to bring a substantial change for the workers and communities affected.

Human rights disclosures may reveal information that stakeholders could not previously access, if at a minimum the disclosures include a description of policies and processes. In other words, reporting on human rights impacts can contribute to legal and non-legal accountability by providing shareholders and other stakeholders with formal acknowledgement by the company of its human rights risks, policies, and processes. In terms of enhancing legal accountability, HRDD laws play a crucial role by expressly placing a legal duty on businesses to prevent, mitigate, and remediate human rights impacts. Various enforcement and liability measures reinforce these duties. As for disclosures, information disclosed

---


120 Eccles & Serafeim, supra note 12, at 3. For a discussion of how social disclosure can catalyze internally driven changes in corporate behavior, see Park, supra note 50. For an analysis of the shortcomings of the EU NFRD on affecting organizational change, see Baumann, supra note 12, at 36–39.

121 CORE COAL, supra note 5, at 6–7.

122 See generally Written Evidence Submitted From CORE, supra note 92, ¶ 6.
in human rights reports can be relied on as evidence in litigation brought against parent or lead companies by individuals harmed at subsidiary and supplier sites.\textsuperscript{123} Most recently, corporate-sustainability reports have been relied on by the plaintiffs in \textit{Lungowe v. Vedanta} in England and \textit{Jabir v. KIK} in Germany to demonstrate the existence of a prima facie duty of care assumed by the parent or the lead company toward the communities or workers harmed by the subsidiary or supplier’s activities.\textsuperscript{125}

On the non-legal accountability side, the approach of a company to its HRDD and reporting can inform stakeholder decisions in relation to the company, including investment decisions, purchasing decisions of consumers and customers, and employment decisions.\textsuperscript{126} It is expected that the stakeholders informed by the human rights disclosures will put pressure on businesses to improve their policies and practices. External agencies, such as international finance institutions or government agencies, may scrutinize these reports if they require evidence that companies have identified and managed human rights risks as a condition of providing support to them. Such support could be the provision of export credit, the granting of a procurement contract, or the loan of finance.\textsuperscript{127} CSOs have made use of the information in various ways that help the public and policymakers see the shortcomings of the legislation but also highlight the contrasts between a company’s statements on human rights and its actual performance. The value of these developments should not be overstated, however. The next Section discusses the inadequacies of the existing legal frameworks for improving corporate accountability for adverse human rights impacts.


\textsuperscript{124} Id.


\textsuperscript{126} Choudhury, supra note 31, at 207; see \textit{SHIFT & MAZARS LLP}, supra note 44.

\textsuperscript{127} See \textit{SHIFT & MAZARS LLP}, supra note 44. In the U.K., the MSA and other relevant legislation do not require public procurement processes to consider company reports under the MSA, § 54. It has been recommended by an independent review that the U.K. Government introduce standards to exclude non-compliant companies from eligibility for public contracts. \textit{INDEPENDENT REVIEW OF THE MODERN SLAVERY ACT 2015}, supra note 93, ¶ 2.6.4.
V. WEAKNESSES AND PITFALLS OF THIS MODE OF MARKET-LED ACCOUNTABILITY

A. Overview

Our analysis of the existing disclosure and HRDD laws in Section III demonstrates that, despite the variation in the obligations imposed, they rely predominantly or exclusively on the “market” to exercise checks and hold businesses accountable for human rights impacts.128 This is not unusual as, at least for corporate disclosures, the main objective is to empower market actors with information. But when the disclosure rules themselves are not designed effectively, the empowerment and accountability functions are hindered by businesses’ lack of disclosure, inadequate disclosures, and misleading disclosures. While HRDD laws represent a crucial step for improving corporate accountability by imposing substantive obligations on businesses to identify, prevent, mitigate, and remediate human rights impacts and to communicate these steps, they still heavily rely on initiatives from stakeholders for oversight and enforcement. Communication of the HRDD processes and outcomes to stakeholders via corporate disclosures is a key tool for those stakeholders to understand and react to the human rights performance of businesses. To effectively and meaningfully exercise this role, stakeholders need the support of regulatory tools to ensure completeness and accuracy of HRDD disclosures.

Studies have shown that disclosure laws have had very limited success in improving human rights conditions for affected groups and improving accountability for impacts.129 This is unsurprising if these laws fail to elevate human rights and environmental impact considerations on the priorities list of the corporate world driven primarily by increasing profits.130 We argue here that to improve the accountability function of HRDD and reporting obligations, two main weaknesses in the current rules need to be overcome. The first one relates to the content of the reports and the information that should be or is disclosed under the relevant legislation. According to benchmarking reports analyzing these disclosures, the content of the disclosures remains largely limited to disclosure of

128 See, e.g., Home Office, Modern Slavery and Supply Chains Consultation 1, 8, 13 (2015) (“We believe that once it is made clear what activity major businesses are undertaking to ensure slavery and human trafficking is not taking place in their supply chains or own business, pressure from consumers, shareholders and campaigners and competition between businesses will encourage those who have not taken effective steps to do so.”). The document continues, “[i]nstead of relying on heavy-handed regulation, this measure will encourage businesses to do the right thing, by harnessing consumer and other stakeholder pressure, which will encourage and influence businesses to do more.”).


130 Bartley, supra note 15, at 49, n.56.
information on commitments and policies rather than disclosing concrete risks to workers and communities and also the substantive steps taken to address them.\textsuperscript{131} This is particularly problematic when the reporting requirements are not accompanied by a due diligence obligation. We join scholars who have argued that the lack of a due diligence obligation preceding the disclosure places serious limitations on the law’s promise to increase corporate accountability and contribute to eradicating human rights abuses.\textsuperscript{132} Most disclosures have been largely limited to descriptions of the company’s commitments and processes in addressing human rights and modern slavery issues in their supply chains.\textsuperscript{133} Little space, if any, is dedicated to issues of substance, such as the specific risks to employees and communities identified within the company’s own business and its supply chain, as well as references to the concrete steps they have taken to eliminate those risks and remediate the grievances.\textsuperscript{134} More advanced reports typically present case studies, the company’s declared approach and commitment to tackling modern slavery or human rights issues, expectations from its suppliers, links to a list of first tier suppliers, identification of the most salient risks, and the plans, policies, programs, and procedures it has established to assess and address the risks.\textsuperscript{135} The pattern of focus on procedures and policies resembles the audit and certification processes that are widely employed by lead firms to regulate labor, human rights, and environmental performance in their supply chains and which typically focus on process rather than substance.\textsuperscript{136} Since audits and certification processes have so far been the central tool used to deal with human rights impacts in supply chains, it is unsurprising that most disclosures are limited to process as well.

\textsuperscript{131} See, e.g., Michelle Langlois, \textit{Human Rights Reporting: Are Companies Telling Investors What They Need to Know?}, SHIFT (May 2017), https://perma.cc/DEY5-GC4R.


\textsuperscript{133} An Ergon Associates study found that “most statements do not go further than general commitments and broad indications of processes.” See ERGON, REPORTING ON MODERN SLAVERY: THE CURRENT STATE OF DISCLOSURE—MAY 2016 1 (2016); ERGON, MODERN SLAVERY REPORTING: IS THERE EVIDENCE OF PROGRESS? (2018). LeBaron and Rühmkorf’s study shows that companies reporting on bribery under the U.K. Bribery Act are using much more firm and clear language, while modern slavery reporting uses a weaker and aspirational language. LeBaron & Rühmkorf, supra note 5, at 25; see also Langlois, supra note 131; CORE COAL., supra note 5, at 6.

\textsuperscript{134} ALLIANCE FOR CORPORATE TRANSPARENCY, 2018 RESEARCH REPORT 24 (2018).

\textsuperscript{135} See, e.g., MARKS & SPENCER GROUP LLP, MODERN SLAVERY STATEMENT 2017/18 (May 2018).

The content of human rights disclosures can be strengthened by including a HRDD obligation. As the HRDD framework proposed by the UNGPs clarifies, the process should focus on the risks to the rights holders rather than focusing on the risks to the business itself. The focus of the reports attached to the HRDD processes should contain rights-holder oriented communication. But as we discuss in the following sections of this Article, an HRDD obligation alone may not enhance the usefulness of disclosures in terms of accountability. The analyses of initial disclosures under the French Law on the Corporate Duty of Vigilance demonstrate that, despite the improvements in disclosures, reporting remains relatively immature. For those companies that have complied with the reporting requirement, disclosures of policies and processes remains the key message. The implementation of § 1502 of the Dodd-Frank Act also suggests that even with a due diligence requirement, more is needed to make these laws effective in achieving their aims. Many HRDD disclosures focus on commitments and processes in the abstract, which fails to provide the kind of meaningful information that stakeholders may rely on in making investment, purchasing, and campaigning decisions about a particular company. This type of disclosure can also easily transform into a publicity tool, painting a misleading picture of a company's human rights performance. More dangerously, it can mask and legitimize serious abuses, especially when they report successes based on audits and certification. This takes us to the second weakness in these laws, which is our focus in the rest of this Article. The lack or inadequacy of mechanisms for formal oversight and enforcement renders the role of the stakeholders, as guardians of accountability, extremely challenging and thus undermines the accountability objectives of these laws. It is hoped that the market forces alone will assume the oversight function and produce the desired accountability outcomes without having the support of appropriate regulatory tools in exercising this function. We


138 Galit A. Sarfaty, Shining a Light on Global Supply Chains, 56 Harv. Int'l L.J. 419, 431 (2015). In this study Sarfaty finds that only about seven percent of companies report strong due diligence measures in their 2014 reports prepared in order to comply with the regulation.

139 See ALLIANCE FOR CORPORATE TRANSPARENCY, 2019 RESEARCH REPORT 90 (2019). An analysis of the sustainability reports of 1,000 companies pursuant to the E.U. NFRD Study found that “general human rights reporting requirements are not an effective tool to ensure the disclosure of information that can help to assess a company’s management of individual risks of human rights impacts, and by extension of whether its business conduct is responsible.” Id.

argue here that even if the first weakness is overcome—as it has been in the French law—without a formal verification, oversight, and enforcement process, the utility of the disclosures to empower stakeholders to pressure for change will be undermined. This results from: (1) the reliability of the disclosed information remaining questionable; (2) stakeholders being left to search for a needle in a haystack in the absence of formal lists of covered businesses and central public repositories for reports; and (3) even where the inadequacy and the accuracy of disclosures are well established, the lack of enforcement measures and sanctions that can be triggered by stakeholders will weaken their leverage. The flaws of content and oversight are closely linked, and for mandated disclosure to contribute meaningfully to corporate accountability for adverse human rights impacts, both must be addressed by the policy and law makers. In the remaining sections of this Article, we engage in an in-depth discussion of oversight and enforcement issues.

B. Market Oversight

Even when companies do fulfill their obligation to report on issues of human rights and modern slavery, doubts remain as to the effectiveness of these reports from an accountability point of view. Early empirical research indicates that the legally mandated human rights and modern slavery disclosures are “quite limited” and “more symbolic than substantive.” Some companies even appear to copy each other’s explanations of their due diligence processes. These practices of failing to comply or selective disclosures have reportedly been unwelcome by certain businesses that want to see serious monitoring and enforcement so as to level the playing field. The Ethical Trading Initiative, in a submission to the U.K. Houses of Parliament Public Accounts Committee, reported that a large majority of companies they have engaged with stated that it is important for the Government to monitor compliance with section 54 of the

---

142 Id.
143 Steve Gibbons, What are Construction and Building Companies Reporting Under the Modern Slavery Act?, LINKEDIN (Feb. 22 2017), https://perma.cc/33W6-WSY6; see also Hess, supra note 50, at 53 (examining human rights reporting and concluding that reports are “unbalanced, incomprehensive, and inconsistent” and that reporting under CTSA and MSA “face[s] similar challenges”).
144 FLEX, supra note 129, at 20–22 (“One FLEX interview participant who is in favor of stronger enforcement of the Act, said that he was concerned that penalties in isolation could result in transparency reporting being a tick box exercise for many companies. He suggested that a penalty for non-compliance should only be introduced in combination with expectations or requirements on the content of statements as many statements otherwise are unlikely to provide meaningful information.”).
Modern Slavery Act and that the Act could not be effective without such monitoring.\footnote{Cindy Berman, Written Evidence From the Ethical Trading Initiative (2018), https://perma.cc/75GR-GQKJ.}

We saw in Section III above that the reporting and HRDD laws analyzed in this Article take varying approaches to the regulation of oversight and enforcement. The common rationale behind each law is to empower key stakeholders, such as investors, consumers, and civil society, with information that will enable them to bring human rights standards to bear on corporate misconduct. In this respect, these stakeholders can play a crucial role of oversight and enforcement by making effective use of the information disclosed through these reports. In other words, the stakeholders will take notice of businesses that fail to report, or report inadequately, and will penalize them by not purchasing products, divesting or not investing, or by running campaigns to raise awareness about the businesses’ failure.\footnote{Responding to Modern Slavery – New UK Benchmarking Report, supra note 36.} These market interventions can, in their most legalized form, include consumer suits for misleading disclosures or advertising. The thinking is that these market pressures will result in companies improving their practices and processes on human rights and modern slavery risks in their supply chains.

There are a few overly optimistic assumptions here and this approach has been challenged already from several angles, particularly with respect to the scale of the desired transformative impact of human rights reports on consumer behavior.\footnote{See Marcia Narine, Disclosing Disclosure’s Defects: Addressing Corporate Irresponsibility for Human Rights Impacts, 47 COLUM. HUM. RTS. L. REV. 84 (2015); Chilton & Sarfaty, supra note 14.} Marcia Narine argues that company human rights disclosures, including on modern slavery, are not always widely disseminated or known and that stakeholders who do know about them do not use the information they contain adequately to press for corporate reform.\footnote{See Narine, supra note 147.} She concludes that evidence of consumer behavior changing as a result of such disclosure is “inconsistent, at best.”\footnote{Id. at 137.} Narine’s reasoning would also apply to HRDD related disclosures. Her viewpoint aligns with a study of human rights disclosure conducted by Adam Chilton and Galit Sarfaty, which found that consumers perceived non-compliant or inadequate supply chain disclosures in the same way as they did detailed disclosures showing a high level of due diligence.\footnote{Chilton & Sarfaty, supra note 14, at 6.} Their study suggests that
“supply chain disclosures are unlikely to be understood and used by consumers making purchasing decisions.”

It is important to acknowledge the limitations of the impact HRDD and disclosure will have on the transformation of corporate behavior if they heavily or solely rely on a market-based enforcement model. Consumer and investor perceptions might change the longer these rules are in force, or as the ability of CSOs to raise public awareness of HRDD and reports increases, thus elevating their impact. But we advance two arguments as to why this will not overcome the limitations of the impact HRDD and disclosure laws can have on improving business behavior. First, the passage of time alone will not overcome the weaknesses relating to the content of the reports discussed in the previous Section. Second, as we argue in this Article, improving content requirements alone will also not suffice to achieve the optimal accountability and transformation objectives envisaged by these disclosure laws. Chilton and Sarfaty identify the limitations of reliance on consumers as influencers in this area. They note that corporate disclosures are generally not sufficiently effective, but that they are less likely to produce meaningful outcomes in this particular area. This is because the information communicated to the consumer relates to processes used in the making of the product and not to its characteristics. They argue that consumers might not be willing to change their purchasing decisions based merely on process if all other qualities of the product are the same. Their reasoning here would also apply to HRDD disclosures. The other obstacle they observe is the difficulty in interpreting the contents of the disclosure. For instance, MSA reports merely present the processes a company is using to try to tackle modern slavery in their supply chain, but do not report the incidences of modern slavery and how the company has responded to them. Also, companies will face different risks depending on variables like sector, business model, or location of sourcing. It is very unlikely that consumers will be able to interpret the contribution of these factors to eradicate modern slavery.

Another important reality to note is that even where awareness is high and there is a sustained reaction against a business because of its performance in this area, this can only cover businesses and brands that are consumer facing. This leaves many large businesses operating in industries, such as mining, construction, shipping, or defense, outside one of the main radars of the transparency

---

151 Id. Chilton and Sarfaty do note that the study had several limitations, like the fact that these reporting requirements are new.

152 Narine, supra note 147, at 40.

153 Chilton & Sarfaty, supra note 14, at 23; FLEX, supra note 129, at 20–21 (“[O]ne company representative [interviewed] suggested the idea that consumers are going to challenge companies for failure to comply with the Act or for publishing statements of poor quality is flawed, saying ‘we’re so far away from this being the reality.’”).

154 Berman, supra note 145.
legislation. These businesses may still be on the CSO or investor radar, but it may be harder for CSOs to garner public interest to a campaign against a non-consumer facing company, and some investors might place less importance on reputational risk posed to a business that is non-consumer facing.

In view of these hurdles for consumers, it is more likely that the greatest pressure on businesses to improve their human rights performance will come from CSOs and investors rather than from consumers. Interested CSOs and investors scrutinize human rights and modern slavery disclosures actively. CSOs publish their analyses of these reports, highlighting the levels of compliance as well as the weak and notable practices. They also invite companies to respond to allegations of human rights abuses in their supply chains, informally or as part of a process established in law—such as the process available under the French Law on the Corporate Duty of Vigilance. The information contained, or the lack of information, in a human rights or modern slavery statement can shine the spotlight on the policies and performance of the reporting companies. Similarly, investors can raise concerns and questions with businesses in which they are investing during annual meetings or directly with management. Investors might also take into consideration the human rights record of a business or its efforts to eliminate modern slavery in its supply chain when making their investment or divestment decisions. Both groups of stakeholders can also engage with policy makers to increase efforts to eliminate modern slavery, if they find the legislative framework inadequate. Recently, the CEO and two senior executives resigned from Rio Tinto after reported investor pressure prompted by the company’s

---

155 In fact, it has been observed that “due to insufficient regulation, a large burden lies with the international civil society which, in various ways, monitors the functionality of transnational corporations.” Jernej Letnar Cerči, Moving Towards Protecting Human Rights in Global Business Supply Chains, 36 B.U. INT’L L.J. 101, 109 (2018); see also Join the Alliance, INVESTOR ALLIANCE FOR HUMAN RIGHTS: A N I T I V E OF ICCR, https://perma.cc/BMD5-6DJR.

156 See, e.g., CORE ET AL., MODERN SLAVERY REPORTING: WEAK AND NOTABLE PRACTICE (2017); CORE COAL., supra note 5.

157 See Brabant & Savourey, supra note 61; Savourey, supra note 65; Cossart et al., supra note 66; CHLOE STEVENSON ET AL., THE LAW ON DUTY OF VIGILANCE OF PARENTS AND OUTSOURCING COMPANIES (Juliette Renaud et al. eds., 2019). NGOs make allegations to the company in the first instance and may elevate their concerns to the court if the company does not respond adequately.

158 See CORE ET AL., ENGAGING WITH COMPANIES ON MODERN SLAVERY – A BRIEFING FOR INVESTORS (2017).

159 For instance, investment fund BlackRock divested from Nevsun Resources Ltd. because of the allegations that the latter uses forced labor in its Eritrean mining operations, after receiving pressure from NGOs. See Campaigners Welcome Blackrock’s Divestment from Nevsun Following Campaign Over Alleged Use of Forced Labor in Eritrea, BUSINESS & HUMAN RIGHTS RESOURCE CENTRE, https://perma.cc/4PPX-APFQ.
decision to blow up a 46,000-year-old aboriginal site in Australia.\(^{160}\) It is yet to be seen how far the company will go to affect structural change in preventing adverse human rights impacts caused by its business. While public relations scandals, such as this recent one involving Rio Tinto, can lead to short-term positive changes within certain parts of a company’s business, it is important to continue keeping companies accountable beyond large scandals to trigger structural improvements.\(^{161}\) Regulatory oversight and enforcement of disclosures and HRDD are among the essential ingredients of such a structural change.

The amount, accuracy, and type of information presented in the reports or, in many cases, the lack of reporting, place a substantial limit on the contribution of investors and CSOs to improving corporate accountability in this area.\(^{162}\) Businesses are required to disclose very little to comply with the reporting requirements, and under some of the laws, this includes the option of disclosing lack of action in this area. The disclosures, even the most detailed ones available, mainly focus on process and contain little information on concrete problems. It is difficult for an investor or a CSO to extract actionable information from these disclosures on a company’s actual human rights performance. These factors significantly limit these stakeholders’ ability to use the disclosures to hold businesses accountable for human rights violations in their supply chains. With only a very small number of CSOs monitoring the legislation, it is unreasonable for policy makers to expect civil society actors with limited resources to drive the push toward business compliance with the transparency laws without any serious regulatory support.\(^{163}\)

One of the only means of private legal action challenging the accuracy of human rights disclosure is consumer litigation. When activist consumers sue companies, this has the potential to send a powerful message.\(^{164}\) The cause of action could be a suit under consumer protection law or some other statutory prohibition on misrepresentation, unfair competition, or false advertising, thereby challenging the accuracy or adequacy of the disclosure and arguing that consumers

---

\(^{160}\) See Ben Butler, Lorena Allam & Calla Wahlquist, Rio Tinto CEO and Senior Executives Reign From Company After Juukan Gorge Debacle, \textit{The Guardian} (Sept. 11, 2020), \url{https://perma.cc/X8UD-EQRG}.

\(^{161}\) For an analysis of how corporations can act as crucial actors for advancement and enforcement of international human rights standards even in the absence of home or host state willingness to do so and the limitations of this approach, see Jay Butler, \textit{The Corporate Keepers of International Law}, 114 Am. J. Int’l L. 189–220 (2020).

\(^{162}\) See \textit{Alliance for Corporate Transparency}, supra note 134.

\(^{163}\) Berman, supra note 145.

were ill informed or misled when purchasing products, causing them harm. The disclosure that is challenged through this litigation has in some instances been made pursuant to one of the laws discussed in this Article, but many of these cases concern product information from labels or other product literature. On the whole, the impact of these lawsuits has been quite limited so far. A series of cases brought against companies in California alleging inadequate and/or misleading disclosure of documented modern slavery in their supply chains were rejected by the courts. There is a concern that, even if these cases were successful, companies would make changes to their labels and product literature, rather than seek to improve conditions for workers in their supply chains. Damages are complicated to calculate in this type of case because it is hard to value the loss to the plaintiffs, when the harm they have suffered is that they would not have bought the products if they had known about the use of child and forced labor in the supply chain. The lack of success in claims to date has not prevented new

---

165 For a full list of cases, see note 167. There have been similar cases outside the U.S. In Germany, a successful complaint was filed against German retailer Lidl in 2010 for false advertising and unfair competition arising from the retailer's claims of fair working conditions in its supply chain. Following the complaint, Lidl agreed to retract the claims in its marketing material on working conditions in its supply chain. Although this was a successful outcome of the litigation, and it had an impact on the company's public statements on these issues, it is not possible to determine whether it had any substantive impact on the company’s sourcing policies or practices. See Complaint re Fair Working Conditions in Bangladesh: Lidl Forced to Back Down, ECCHR, https://perma.cc/RX3G-VPUZ. This complaint did not relate to disclosure made under modern slavery legislation but rather to statements on its supply chain that the company made voluntarily.

166 One of the early cases in this line, *Kasky v. Nike, Inc.*, 27 Cal. 4th 939 (2002), concerned representations that Nike had made about working conditions in its supply chain. A settlement was agreed for $1.5 million and involved investments by Nike to strengthen workplace monitoring and factory worker programs.

167 See *Sud v. Costco Wholesale Corp.*, No. 15-CV-03783-JSW, 2016 U.S. Dist. LEXIS 5524 (N.D. Cal. Jan. 15, 2016) (dismissing case for lack of standing); *Barber v. Nestle*, 154 F. Supp. 3d 954 (9th Cir. 2018) (dismissing plaintiffs’ false advertising and unfair competition claims on the grounds that the CTSCA creates a safe harbor from liability by defining what a company is required to disclose regarding the use of forced labor in its supply chain, and dismissing plaintiffs’ misrepresentation claim, finding that the statements about supplier adherence to law and industry standards were “aspirational”); *Hodsdon v. Mars, Inc.*, 891 F.3d 857 (9th Cir. 2018) (dismissing the case and finding that Mars does not have a duty to disclose forced labor in its supply chain because it is not a physical defect that affects the central function of the chocolate products); see also *Tomasella v. Nestlé USA*, Inc., 962 F.3d 60 (1st Cir. 2020); *Tomasella v. Nestlé USA, Inc.*, 364 F. Supp. 3d 26 (D. Mass. 2019), aff’d, 962 F.3d 60 (1st Cir. 2020); *Tomasella v. Hershey Co.*, No. 18-CV-10360-ADB, 2019 U.S. Dist. LEXIS 14488 (D. Mass. Jan. 30, 2019) (dismissing the case because the consumers’ claims were not actionable under Massachusetts law and they failed to show that the companies deceived them).

168 See Complaint re Fair Working Conditions in Bangladesh: Lidl Forced to Back Down, supra note 165. The consumer complaint against the retailer Lidl regarding its advertising campaign claiming fair working conditions in its supply chain led to the company withdrawing the advertisements, rather than the company being compelled to take steps to ensure fair working conditions.
cases being brought, however. Thus, consumer litigation has the potential to provide enforcement of reporting accuracy and adequacy but in an ancillary role compared to other types of enforcement.

VI. THE Importance of Regulatory Oversight and the Institutional Options

We argue in this Article that to achieve their stated accountability goals, human rights due diligence and disclosure requirements should be accompanied by rules establishing: (1) a formal list of businesses covered by the requirements and a publicly accessible repository for storing annual disclosures; (2) an institutional structure to exercise oversight; and (3) enforcement functions. The institutional structure should have subject matter expertise, in order to provide training and guidelines to ensure accuracy and completeness of disclosures. Without these features accompanying the disclosure requirements, stakeholders’ ability to make effective use of the information disclosed becomes significantly diminished. The presence of an oversight body with the powers to check accuracy and completeness and impose sanctions for misleading and incomplete disclosures will allow the stakeholders targeted by the transparency rules to exercise their leverage more systematically and effectively.

Admittedly, it may be a huge task for a regulatory body to scrutinize all submitted reports, especially as the number of covered companies grows, but the body could routinely review a random sample each financial year and be prompted by investors, consumers, and civil society to carry out additional reviews or investigations. Stakeholders should have standing to initiate complaints regarding suspected discrepancies and inaccuracies in reports to an expert body, equipped with legal authority and sufficient resources, that can investigate the accuracy or adequacy of the information; if needed, compel the business to correct and complement the disclosure; and impose penalties for failure to comply. In this approach, stakeholders—equipped with and empowered by the regulatory tools and the institutional infrastructure to exercise their watchdog role more effectively—continue to play a key role in holding businesses accountable. This

---

169 Walker v. Nestle USA, Inc., No. 3:19-CV-723-L-BGS, 2020 U.S. Dist. LEXIS 106858 (S.D. Cal. June 17, 2020); Myers v. Starbucks Corp., 5:20-CV-00335-CBM-SHK (C.D. Cal. Feb. 19, 2020). Both cases allege affirmative misrepresentation by the defendant companies in relation to statements made on product labels and literature that cocoa is “sustainably sourced,” “certified,” and “supports” or “helps” farmers. They rely on allegations of child and forced labor in the cocoa farms and of environmental destruction, as part of clearing the land for farms, to evidence that these statements are misrepresentations.

170 We are not alone in concluding that verification and oversight are needed. Doorey, for instance, in an article on using domestic disclosure to influence foreign labor practices, argues that the information needs to be verified by the state (and/or a credible outside auditor). See Doorey, supra note 30.
way, the role of oversight is not entirely or largely left to the voluntary efforts of investors and CSOs whose abilities to push for compliance can be limited by several factors, such as scarce financial and human resources, lack of authority to compel further disclosures, and inability to impose financial penalties. At the same time, the oversight role would not be left exclusively to the regulatory authorities. Stakeholders would continue to play a crucial role in the accountability framework from a strengthened position both by having access to a centralized list and a repository and, more importantly, by having standing to bring complaints before a body with powers to investigate and impose penalties.

A. Key Functions of Meaningful Regulatory Oversight

Among the HRDD and transparency laws discussed in this Article, the regulatory oversight feature remains either inadequate or non-existent. Even the French Law on the Corporate Duty of Vigilance—with its advanced accountability features providing legal standing for civil society before courts to bring actions to enforce the duty to publish an adequate vigilance plan—lacks a regulatory body to exercise oversight over the law’s implementation and to hold a central list and repository, thus leaving the challenging and resource-intensive monitoring function almost exclusively to stakeholders. Taking stock of the initial experiences with the transparency laws discussed in the earlier sections of this Article and most recently with the French Law on the Corporate Duty of Vigilance, we argue that all of the elements elaborated below are needed to achieve greater level of accountability via transparency and HRDD laws.

1. List and Repository

CSOs have expressed concern about the lack of information regarding companies covered by various HRDD and reporting laws currently in force. It is often left to their investigative skills to identify which companies may be covered by disclosure requirements and confirm whether covered companies have published disclosures. Having an annually-updated, formal list of companies covered by human rights disclosure requirements provides the stakeholders and the regulator with an essential tool for identifying which companies have complied with the most basic obligation under these laws. A central repository accessible by the public to store annual disclosures will allow stakeholders more efficient access both to the most recent reports and all the other years since the introduction of the relevant laws.

As discussed in Section III, some disclosure laws already provide for a list and/or a repository, such as the Australian MSA. For the disclosure laws that fail to provide for the establishment of a formal list and a central repository for

---

171 See, e.g., Savourey, supra note 65.
Human Rights Disclosure and Due Diligence Laws

There have been calls for governments to introduce these safeguards (for example, the U.K. MSA) to improve the effectiveness of these laws. In agreement with these calls, we argue that the establishment of a formal list and a repository to be an essential ingredient for improving the accountability mission of HRDD and disclosure laws.

2. Monitoring Function and the Content of Reports

The added value brought by having a regulator with monitoring responsibility is in ensuring submissions are made in a timely fashion and in exercising checks on the content of an appropriate size sample each year. These sample checks can ensure coverage of all required elements and accuracy disclosures. As we discussed above in Section V, the existing HRDD and disclosure laws rely primarily on stakeholders to monitor business performance and compliance with the applicable law's requirements. Even the most evolved statutory regime, the French Law on Corporate Duty of Vigilance, does not establish regulatory oversight and relies on stakeholders to monitor whether covered businesses have developed and published an adequate vigilance plan. The enforcement and sanctions mechanisms of the French law (discussed in Section III) depend solely on the stakeholders identifying the lack of compliance with the law or the misleading statements within published plans and triggering the relevant court processes stipulated in the law. We will return to the subject of stakeholders raising complaints as a trigger for regulator action below, but first we address the primary source of monitoring currently in place in this legislative field, namely the monitoring of non-financial disclosure. Auditors usually monitor non-financial disclosure, as the complement to financial disclosure, in the first instance, with financial regulators holding a further oversight function. The level of monitoring provided by the financial regulator varies from state to state.

In some countries, limited oversight of human rights disclosures is exercised by accounting or securities regulators such as the U.K. Financial Reporting Council (FRC) and the U.S. SEC. For instance, the SEC enforces liability for any false or misleading statements under Dodd Frank §1502. The Trump Administration opposed §1502 and made a proposal to repeal it. Verification and enforcement were stepped down. Whether coincidentally or as a result of these developments, there is very little verification and enforcement of this law.

---

173 The other potential source of monitoring under existing legal regimes is the US securities regulator, the SEC. In theory at least, the SEC enforces liability for any false or misleading statements under Dodd-Frank Wall Street Reform and Consumer Protection Act. See Dodd-Frank Wall Street Reform and Consumer Protection Act, § 1502, 111 P.L. 203, 124 Stat. 1376 (2010).
175 RESPONSIBLE SOURCING NETWORK, supra note 101, at 4, 9.
that has occurred in practice. The appropriateness of the SEC as a regulator on corporate social impacts is debated.\textsuperscript{176} The current civil society effort in the U.S. to push for a new corporate transparency law on social and environmental impacts is focused on publicly listed companies—its press statement arguing that “the SEC is the right agency, given its expertise in corporate disclosures, and broad mandate to protect investors and the public interest.”\textsuperscript{177} On the other hand, the ability of the SEC to be a “humanitarian watchdog” has been questioned, due to the organization’s lack of specialist knowledge.\textsuperscript{178}

The E.U. NFRD, which has been transposed into the law of E.U. Member States, provides a sample of instances on the monitoring of human rights disclosures. Eight states require that disclosure under the NFRD forms part of the company’s management report.\textsuperscript{179} The allocation of the human rights report within the management report allows for a basic level of auditor scrutiny over the content of the report, as the E.U. Accounting Directive requires that an auditor check the entire management report to verify its consistency with the financial statements and its compliance with legal requirements and also to check for the presence of material misstatements.\textsuperscript{180}

For a verification of a human rights report, there is not much to be gained by consistency checks with financial statements. Compliance with legal requirements is also easily verifiable as the legal requirements for human rights reporting are minimal and relatively vague. Only the checks for material misstatements could prove useful in the human rights reporting context, but such checks will require expertise and access to information that may exist either within the company as well as beyond the company, and often beyond the country. Four states require verification of information beyond checks for consistency with the management report.\textsuperscript{181} In Denmark, for instance, the implementing legislation envisages a regulatory review of ten to twenty percent of listed companies that are selected for full scope enforcement each year, checking presence and content of


\textsuperscript{178} See Woody supra note 176.

\textsuperscript{179} GLOBAL REPORTING INITIATIVE, supra note 56, at 16, 17, 19, 24, 26, 27, 30, 31. (Belgium, Bulgaria, Denmark, Latvia, Netherlands, Romania, U.K., and Iceland).


\textsuperscript{181} GLOBAL REPORTING INITIATIVE, supra note 56, at 19, 20, 23, 26 (Denmark, France, Italy, and the Netherlands).
The enforcement approach is based on the materiality of the disclosed information. Material misstatements may result in the imposition of fines in accordance with the Danish Financial Statements Act. The experience with the human rights transparency laws to date shows that the unique features of human rights reporting call for a sui generis approach to oversight and enforcement. As with financial reporting, the rules on the required information and the oversight and enforcement related to the completeness and accuracy of the content disclosed may be designed around the concept of “materiality,” but the meaning of “materiality” in a human rights context is unclear in the existing reporting regulations. Materiality has been viewed as a misleading concept in the human rights context, and instead the use of “salient risks” has been proposed by the UNGP Reporting Framework. Companies are typically required to disclose material information in non-financial reports including human rights disclosures. The E.U. Accounting Directive, for instance, describes material information as “the status of information where its omission or misstatement could reasonably be expected to influence decisions that users make on the basis of the financial statements of the undertaking. The materiality of individual items shall be assessed in the context of other similar items.” For non-financial statements, the E.U. NFRD adds that the required disclosures shall contain information “to the extent necessary for an understanding of the . . . impact of [the company’s] activity.” The NFRD recognizes that in determining materiality, the context in which the business is operating needs to be taken into account. Recital 8 of the Directive states that the information disclosed should cover “principal risks of severe impacts,” which will be assessed by the scale and gravity of impact. One study points out that this standard introduces a different approach to materiality by focusing on the “scale and gravity of the materialization of the risk, rather than whether knowledge of a principal risk would influence readers’

See generally Forslag Til Lov om Ændring af Årsregnskabsloven og Forskellige Andre Love [Act Amending the Danish Financial Statements Act], LOV nr 738 af 01/06/2015 (2015); GLOBAL REPORTING INITIATIVE, supra note 56, at 19.


Note the difference here between due diligence and reporting obligations. The due diligence laws are not framed in terms of materiality, but for instance, the French Law requires companies to detect risks and prevent serious violations with respect to human rights (the threshold is “serious”).

See SHIFT & MAZARS LLP, supra note 44.


economic decisions.” The approach fleshed out in the Recital 8 does align with the conception of “salient risk” embedded in the Pillar Two of the UNGPs.

In a shift away from the usual investor-risk rationale for non-financial reporting, UNGP Pillar Two focuses on risks to rights-holders, and the importance of taking into account the perspectives of those who may be directly affected by companies’ actions. One study notes that this aspect of the “materiality” concept has not been reflected in the Member State implementing legislation covered in the study, however. A guidance published by the U.K.’s FRC on the non-financial reporting, for instance, emphasizes materiality for investors, thus following the classic shareholder-centric understanding of materiality in the reporting context. But such an understanding of materiality for human rights reports does not align with the understanding of risk under the UNGPs, which are instead centered around the affected individuals and communities. We argue here that human rights disclosure laws should impose mandatory minimum content, covering salient human rights risks posed to individuals and communities affected by the activity of the business and the steps taken to prevent, mitigate, and remediate impacts.

In assessing the completeness of human rights disclosures, auditors and regulatory bodies will have to determine which human rights issues relating to a company’s business can be categorized as salient. The size and geographical spread of a covered company’s business is likely to render checking the completeness of the disclosure challenging due to external information covering each overseas or domestic subsidiary or supplier’s human rights impact not being readily available to the external auditors and the regulatory bodies. The distinction between this type of reporting and reporting on diversity and governance is apparent here: diversity and governance reports are more amenable to verification by domestic regulatory bodies because they are driven by quantitative data, meaning they are more easily fact checked. Local and international CSOs and inter-governmental bodies, such as the International Labor Organization, can play a supportive role here. They are aware of business-related human rights impacts on the ground, and their documentation of impacts

---

188 Claire Jeffrey et al., Comparing the Implementation of the EU Non-Financial Reporting Directive in the UK, Germany, France, and Italy 5 (2017).
189 Jeffrey et al., supra note 188, at 4.
191 See Jeffrey et al., supra note 44, at 48–51 (explaining salience in this context).
192 There are also restrictions under international law on extraterritorial verification and enforcement. As Doorey reminds us, “Canada obviously cannot send inspectors to examine workplaces in Bangladesh to verify the accuracy of the information provided by MNCs.” Doorey, supra note 30, at 385.
can at least help regulators raise red flags for problematic areas that may prompt a more detailed investigation. The dichotomy in approaches to human rights reporting identified in this Section highlights the need for a state-based approach to monitoring and enforcement for human rights disclosures that marries human rights with business and accounting expertise.

Civil society and investor groups can alert the regulators and other relevant authorities of suspected false or misleading statements or to omissions of salient risks from reports. The effectiveness of the existing procedures is variable. As noted above, the French Law on the Corporate Duty of Vigilance allows any person with legitimate interest to give official notice to the company to comply with the law. If the company does not comply within three months of the notice, a judge could oblige the company to publish a plan, under financial penalty if necessary. The judge would also rule on whether a vigilance plan is complete and appropriately fulfills the obligations described in the law. But much of the heavy work of identifying and locating vigilance plans and identifying and investigating their inadequacies falls to the civil society actors. Under the Dutch Child Labor Law, any stakeholder with concrete evidence that a company’s goods or services were produced with child labor will be able to submit a complaint to that company. If the issue is not resolved, the stakeholder will be able to submit the complaint to a regulator. Once a complaint is filed, the regulator may issue a legally binding instruction ordering the company to conduct the required due diligence and make the appropriate declaration. Again, the process of monitoring is very much stakeholder led. Even without specific power contained within the disclosure law, a regulator can invite and welcome complaints from external parties such as CSOs, as the Attorney General of California has with respect to CTSCA. As discussed above in Section III, this has not proven an effective means of oversight and enforcement. Reports from external parties are also used as part of the operating procedure for regulators reviewing companies’ non-financial reports, such as the FRC in the U.K., which accepts complaints and reviews reports on the basis of these. The FRC’s procedures are fairly limited in their effectiveness, however, as the discussion that follows in next Section illustrates.

194 FINANCIAL REPORTING COUNCIL, THE CONDUCT COMMITTEE: OPERATING PROCEDURES FOR REVIEWING CORPORATE REPORTING (2017). For information about making a complaint to the FRC or raising a whistleblowing concern, see Whistleblowing, FINANCIAL REPORTING COUNCIL https://perma.cc/Y2GF-WK9Q. The U.K. implemented the EU NFRD in the Companies, Partnerships and Groups (Accounts and Non-Financial Reporting) Regulations 2016, which added § 414CB to the Companies Act 2006. On March 11, 2019, the Business Secretary announced that the FRC will be abolished and replaced by a new regulator, the Audit, Reporting and Governance Authority. Audit Regime in the UK to Be Transformed with New Regulator, GOV.UK (Mar. 11, 2019), https://perma.cc/R8DG-8UBX.
3. Enforcement Function

The added value brought by having a regulator with enforcement responsibility is the investigation of instances of alleged noncompliance and the imposition of sanctions and penalties when noncompliance is found. On the other hand, sanctions for noncompliance are not the hallmark of transparency provisions but do feature in the two due diligence laws. Although the French Law on the Corporate Duty of Vigilance is relatively new, the enforcement mechanism was triggered for the first time in 2019 and has now been used in five instances, with two cases so far having reached a court. There is yet to be a substantive judgment on whether a company has breached its duty of vigilance, however, because the question of which court is competent is still being litigated; this nonetheless represents a significant departure from the status quo of minimal, if any, enforcement under the other laws discussed.

Most transparency laws discussed in this Article lack an effective enforcement mechanism for noncompliance. Efforts have been made to seek sanctions for noncompliance with human rights reporting requirements placed on certain large or listed companies under U.K. law. In one instance, CSO ClientEarth referred mining company, Rio Tinto, to the relevant regulator, the FRC, for failing to report the reality of the company’s environmental and social impacts. The regulator found that Rio Tinto had failed to make material disclosures about serious environmental, employee, social, and community issues at a mine site in Indonesia. Following this finding, Rio Tinto’s directors included more information, as advised by the FRC, in their report and accounts for the following year, and the regulator closed its inquiry. No other sanction was

---


196 The two cases are against oil company Total. See Total Lawsuit (Failure to Respect French Duty of Vigilance Law in Operations in Uganda), BUSINESS & HUMAN RIGHTS RESOURCE CENTRE, https://perma.cc/2XFJ-L2ZG.

197 Id.

198 The complaint concerned, in particular, the group’s non-managed Grasberg mine in Indonesia. The Norwegian Sovereign Wealth Fund had divested from Rio Tinto on account of “severe environmental damage” at Grasberg. The mine had been subject to bombings and other attacks from local resistance groups, and the mine operators continued to pay for mine security provided by the Indonesian military, despite the military’s history of human rights violations in Papua New Guinea. See CLIENTEARTH, REFERRAL TO THE FINANCIAL REPORTING REVIEW PANEL, RE: THE RIO TINTO GROUP ANNUAL REPORT 2008 (2010); see also Charlotte Villiers, Narrative Reporting and Shareholder Value, in DIRECTORS’ DUTIES AND SHAREHOLDER LITIGATION IN THE WAKE OF THE FINANCIAL CRISIS 120–126 (Joan Loughrey ed., 2012).

199 CLIENTEARTH, supra note 198.
applied. There is power under the Companies Act 2006 for the regulator to apply to court for a declaration that the annual reports of a company do not comply with the relevant requirements and for an order requiring directors of the company to prepare revised accounts. However, this power has never been used. This case is illustrative of the role adopted by regulators thus far with respect to human rights reporting: accepting complaints, reviewing reports, but using sanctions as very much a last resort. To avoid situations of ineffective enforcement, we envisage the introduction of sanctions and penalties prescribed by the transparency or HRDD regulations that have to be imposed by the regulator in accordance with established law and not on a discretionary basis. A further added value of having a regulator with the power to impose sanctions and penalties for noncompliance is the possibility of channeling financial penalties applied to a fund that can be used as a contribution to reparations to individuals or communities affected adversely by the acts and omissions of the penalized corporation.

B. Institutional Options and Subject Matter Expertise

With political will and support, a regulator can be empowered and resourced to acquire subject matter expertise on human rights and also business and accounting. A regulator staffed with appropriate experts and supported by sufficient resources would develop greater expertise over time to establish indicators on human rights risks on a sectoral and geographical basis. This knowledge would enable the regulator to evaluate corporate disclosure and due diligence to determine whether it reflects the salient risks to human rights from the company’s operations. In this respect, the regulatory body is not expected to penalize companies for human rights violations in their supply chains directly. Rather, it would focus on the company’s compliance with the HRDD and disclosure standards. The regulator would not evaluate the substance of a human rights claim against the company, nor apply international human rights standards to determine a violation of such standards. For instance, the regulator would have

---


202 The fund envisaged here is different than the reparation orders made by U.K. courts under the MSA §§ 8–9. In a similar vein, in June 2018 the relevant U.K. authorities established the “General Principles to compensate overseas victims (including affected States) in bribery, corruption and economic crime cases,” a common framework set up “to identify cases where compensation is appropriate and act swiftly in those cases to return funds to the affected countries, companies or people.” See New Joint Principles Published to Compensate Victims of Economic Crime Overseas, SERIOUS FRAUD OFFICE (June 1, 2018), https://perma.cc/BQ4X-Y5JD.

203 JOINT COMM. ON HUM. RTS., supra note 5.
the power to inspect whether a lead company has adequately disclosed the human rights risks that are present in its supplier factories and the HRDD steps it has taken to prevent, mitigate, and remediate any impacts. The regulator can sanction the lead company, if it is satisfied that either the disclosed information does not adequately capture the risks present and the steps taken or that the information is misleading or inaccurate.

In terms of understanding risk, the U.S. Department of Labor commissions an annual child labor report known as “Sweat and Toil.”204 These are detailed reports of instances of child labor around the world and the gravity of each case. This subject matter expertise, if held by regulators of corporate HRDD and disclosure, would serve to enhance the ability of stakeholders to verify the content of company reports and would also provide information for companies to consider when they assess human rights and modern slavery risks. An additional function of the regulator could be disseminating this information and developing guidelines for businesses and other stakeholders.

We see slow movement in this direction. For instance, proposals made so far to improve the MSA disclosures from an oversight perspective include provision for a government funded central repository for published statements,205 for the government to publish a list of companies that must report under the Transparency in Supply Chains clause of the Act,206 and for the establishment of an independent review of modern slavery statements made by companies.207 The last of these is the crucial piece, according to our argument for regulatory oversight. Whether through an enhanced role for the Anti-Slavery Commissioner, or through the creation of the sui generis body we recommend, independent review is a necessary step toward the accountability goal of the legislation. As noted above, there is a proposal currently under consideration in France for additional state oversight for the Law on the Corporate Duty of Vigilance. Rather than taking the form of a regulator, this would entail designated individuals within the relevant ministry providing guidance to companies on implementation and checking on compliance.208

Commentary to the UNGPs Principle 3 acknowledges the role which can be played by national human rights institutions (NHRIs) “in helping states identify


205 As noted above, a government funded central repository was set up by the Australian Act; it was also recommended by the U.K. Houses of Parliament Joint Committee on Human Rights in its 2017 report on the MSA. JOINT COMM. ON HUM. RTS., supra note 5.

206 FLEX, supra note 129, at 20.

207 JOINT COMM. ON HUM. RTS., supra note 5.

208 Duthilleul & de Jouvenel, supra note 71.
whether relevant laws are aligned with their human rights obligations and are being effectively enforced, and in providing guidance on human rights also to business enterprises and other non-state actors.\textsuperscript{209} To move the business and human rights agenda forward meaningfully at the domestic level, there are multiple benefits to be gained from a specialized regulator in this area, as acknowledged in the UNGPs regarding the role that can be played by NHRIs for guidance and enforcement. While we do not envisage the role of the regulator proposed here to be carried out by a NHRI,\textsuperscript{210} close cooperation between NHRIs and the regulator overseeing human rights reporting would be beneficial for the latter to establish and develop human rights expertise.

We recommend the establishment of a \textit{sui generis} body, or a specialized department within an existing body, to tackle both corporate and human rights aspects of the reporting. This independent oversight mechanism should have responsibility for reviewing reports and providing feedback to a sample of companies on an annual basis, similar to the process established through the Danish implementation of the E.U. NFRD. We take the view that there should be provision for external parties to alert the oversight body, which can then investigate the accuracy or adequacy of the information, and, if needed, compel the company to correct and complement the disclosure. The oversight body should be able to impose meaningful penalties for failure to comply akin to those in the Dutch Child Labor Due Diligence Law. We stress the need for the oversight body to have specialist subject matter knowledge that goes beyond that of a corporate regulator to include the complexities of the human rights and modern slavery issues which are the subject of the reports. Such specialist knowledge could, for instance, come from the commission of “Sweat and Toil” type reports or from close cooperation with NHRIs. The oversight body should analyze trends in reporting and company practice, and also develop training and guidance in relation to human rights, modern slavery, cases of forced labor, and human trafficking, including their drivers and outcomes.\textsuperscript{211}

\textsuperscript{209} UNGPs, supra note 11, at 6.

\textsuperscript{210} Though this does not mean that NHRIs cannot be tasked with such a role. In the U.K., the Equality and Human Rights Commission oversees and enforces gender pay gap reporting regulations that place obligations on the public and private sector. \textit{See} \textit{EQUALITY & HUM. RTS. COMM’N}, \textit{CLOSING THE GENDER PAY GAP (2018)}.

\textsuperscript{211} The SEC Foreign Corrupt Practices Act Resource Guide is an example of the kind of guidance that can be offered to companies by a regulator. \textit{Crim. Div. of U.S. DOJ & Enf’t Div. of U.S. SEC, FCPA: A RESOURCE GUIDE TO THE U.S. FOREIGN CORRUPT PRACTICES ACT (2020)}. The lack of equivalent state-sponsored guidance for the French Law on the Corporate Duty of Vigilance is conspicuous.
VII. CONCLUSION

We began this Article by acknowledging that HRDD and transparency can contribute to improving human rights and labor conditions in global production networks. But for due diligence and transparency to make a genuine contribution to improving corporate accountability and to avoid potentially masking and legitimating abuses, legislation should move away from heavily relying on a market-based model of accountability. Numerous studies and reports show that the existing frameworks have been inadequate. One obvious area of improvement concerns the content of the disclosures.

There have also been calls for more regulatory involvement and a move away from the market-led model of oversight, such as the establishment of a registry of MSA reports in the U.K., or the introduction of effective sanctions for noncompliance with the reporting standards. In this Article, we argued that state-based oversight and enforcement is an essential element for human rights reporting to be effective. Without this element, even where mandatory HRDD is introduced, there remain serious limitations on ensuring accuracy and completeness of reports.

Our contribution to this reform agenda is twofold. First, we argue that there is a need to support HRDD and transparency frameworks with a state-based oversight mechanism that can also be supported by stakeholders. Second, we emphasize that oversight for human rights reporting requires a fundamentally different approach to institutional expertise and to risks and materiality than financial or governance reporting. So far, oversight of a limited number of reporting frameworks were entrusted to bodies specializing in traditional corporate reporting without staffing these bodies adequately with human rights expertise. We urged policymakers to move away from this one-size-fits-all model and adopt a sui generis model of oversight marrying knowledge of corporate reporting with human rights expertise to verify and enforce human rights and modern slavery transparency regulations.
Remediation in Foreign Bribery Settlements: The Foundations of a New Approach
Samuel J. Hickey*

Abstract

A handful of nations spearhead the global anti-corruption regime through the transnational enforcement of foreign bribery laws. These laws prohibit corporations with a connection to the enforcing nation from paying or offering bribes to the officials of a foreign nation. Enforcement agencies construe the extraterritorial application of these laws broadly, establishing their global prominence. The most notable example is the United States Department of Justice’s enforcement of the Foreign Corrupt Practices Act of 1977 (FCPA). Enforcement agencies typically resolve investigations against corporations through deferred prosecution agreements and other consensual settlement mechanisms known generally as non-trial resolutions. Fines and penalties paid pursuant to these agreements can extend beyond the billion-dollar mark. In most cases, money paid in fines and penalties goes to the treasury of the enforcing nation. However, a movement has emerged that advocates for the sharing of proceeds of non-trial resolutions with the victims of foreign bribery, namely, citizens and governments in the developing world. This movement is complemented by a small number of instances in which non-trial resolutions have been used to provide remediation in this manner. However, these cases do not reveal a coherent approach to remediation, and enforcement agencies do not have the benefit of any kind of conceptual or practical framework to guide the provision of remediation. The extant literature also fails to consider the many political and practical difficulties of coupling transnational foreign bribery enforcement with a remedial agenda. The purpose of this Article is to address the practicalities of when and how remediation might be written into the terms of non-trial resolutions. To achieve

* LLB (Hons I) (Queensland), LLM (Harvard). I am grateful to Professor Matthew Stephenson, Phil Mason OBE, Branimir Hock, William Thomas, Jordan English, Ashrutha Rai, Sagnik Das, and the participants of Harvard Law School’s 2019 Global Anti-Corruption Lab for their thoughts and feedback on this Article. I am indebted to Zach Heater, Jared Mayer, Neema Hakim, Daniel Sung, Keila Mayberry, Maria O’Keefe, and the editorial team of the Chicago Journal of International Law for their assistance in the preparation of this Article. An earlier version of this Article was presented at the OECD 2019 Anti-Corruption & Integrity Forum, and I am grateful to those attendees who provided comments and suggestions, including the prosecutors from anti-corruption enforcement agencies in the U.S. and U.K. who provided insight, counterargument, and criticism. The views and opinions expressed in this Article are mine alone, as are any errors.
this goal, this Article assumes two functions. First, it offers a conceptual framework to underpin remediation by defining elusive notions such as harm, victimhood, and remediation itself. Second, it presents a list of factors to guide the provision of remediation in foreign bribery cases. The shared benefit of these conceptual and practical frameworks is that they allow remediation in foreign bribery settlements to be approached with newfound precision. These frameworks are ultimately geared toward moving practice forward in this fledgling field by improving the consistency of outcomes and developing a body of precedent and best practices.

Table of Contents

I. Introduction........................................................................................................................................369
II. Background, Challenges, and Current Approaches .................................................................372
   A. Fundamental Aspects of Foreign Bribery Enforcement .................................................................372
      1. The FCPA and the Foundations of the Global Foreign Bribery Regime ..................................372
      2. The Extraterritorial Application of Foreign Bribery Laws .........................................................374
      3. The Reliance on Non-Trial Resolutions and Prosecutorial Common Law ..................................375
      4. The United Nations Convention Against Corruption.................................................................377
   B. Practical Barriers to Remedial Settlement Distribution .................................................................378
      1. The Demands and Implications of Prosecutorial Discretion .......................................................379
      2. Justifications for Improving Extant Practice ..............................................................................380
   C. Current Approaches to Remedial Settlement Distribution .............................................................383
      1. United Kingdom ..........................................................................................................................384
      2. United States ..............................................................................................................................388
      3. Canada .......................................................................................................................................390
III. Terminology and Concepts ..............................................................................................................392
   A. Direct and Indirect Harm ..............................................................................................................392
   B. Categories of Remediation ............................................................................................................399
IV. A Framework for Compensation ....................................................................................................402
   A. The Identifiability of the Victim ......................................................................................................402
   B. Whether Harm is Ascertainable ......................................................................................................405
   C. The Risk of Repeat Corruption ....................................................................................................406
   D. The Appropriateness of Compensation .........................................................................................408
V. A Framework for Reparations ...........................................................................................................409
   A. Nexus Connecting Foreign Bribery and Harm .............................................................................409
   B. Risk of Repeat Corruption ............................................................................................................414
   C. Whether Reparations are Appropriate in the Circumstances .........................................................417
VI. Conclusion .......................................................................................................................................417
I. INTRODUCTION

Corruption can have corrosive economic consequences within the developing world. A veritable mass of literature demonstrates how both the governments and populaces of developing nations are harmed when foreign corporate actors bribe government officials.¹ The anti-corruption efforts of Western nations focus on policing and punishing those under their jurisdiction who commit the criminal offense of foreign bribery by offering or paying either bribes or things of value to influence the acts or decisions of public officials in foreign nations.² This regulatory activity, known as supply-side foreign bribery enforcement,³ often involves schemes in which corporate actors from developed nations offer or pay bribes to the public officials of developing nations in order to procure lucrative public infrastructure contracts.⁴ Controversially, anti-corruption enforcement agencies like the U.S. Department of Justice (DOJ) and the United Kingdom Serious Fraud Office (SFO) broadly construe the extraterritorial application of their national foreign bribery laws, such that enforcement actions are often initiated against foreign corporations with only tenuous connections to the enforcing nation.⁵ This regulatory market is remarkably robust; enforcement agencies levy substantial fines, while entities subject to regulation dedicate significant resources to foreign bribery compliance.⁶ Since 1999, prosecutors around the globe have collected in excess of US$15 billion


³ OECD, FOREIGN BRIBERY ENFORCEMENT: WHAT HAPPENS TO THE PUBLIC OFFICIALS ON THE RECEIVING END? 3 (2018) (defining supply-side foreign bribery as “relating to what bribers do—it involves offering, promising or giving a bribe to a foreign public official to obtain an improper advantage in international business. In contrast, the demand side of foreign bribery refers to the offence committed by public officials who are bribed by foreign persons.”).

⁴ KEVIN E. DAVIS, BETWEEN IMPUNITY AND IMPERIALISM 41–54 (2019) (setting out the hallmarks of foreign bribery practice, labelling the modern enforcement landscape as “the OECD paradigm”).

⁵ Id.; see also Section II.A.2.

through foreign bribery enforcement. Unsurprisingly, enforcement is sometimes justified with reference to the harmfulness of bribery in the developing world. However, the money extracted through foreign bribery enforcement is typically retained by the treasury of the enforcing nation. These proceeds, extracted almost exclusively through criminal settlement agreements made between prosecutors and corporations, are rarely used to remedy, or to attempt to remedy, the harm caused by corruption. Several scholars and commentators have proposed that the money extracted through foreign bribery settlement agreements should go toward more constructive ends that would assist the victims of corruption or the global anti-corruption efforts generally. However, the academic literature to date has yet to reveal satisfactorily how this proposal might be operationalized on a practical level. These proposals are nonetheless buttressed by a small handful of instances in which either the DOJ or the SFO has negotiated settlement agreements that include terms providing some type of remediation to the victims of corruption. However, neither the DOJ nor the SFO has adopted any sort of comprehensive criteria to determine whether remediation should be pursued in a given case, which victims should receive it, or how they should receive it. As such, seeking remediation through foreign bribery settlement agreements has proven unprincipled and inconsistent in practice. The purpose of this Article is to provide a coherent framework to support the use of foreign bribery settlement agreements to assist the victims of corruption. The term “remedial settlement distribution” has been coined to describe this practice.

This Article offers three contributions to the extant literature. First, it explains how remedial settlement distribution has thus far functioned in practice and situates such distribution within the modern foreign bribery enforcement landscape. Second, it supplies working definitions for important concepts—such as harm, victimhood, and remediation—in the context of foreign bribery settlement agreements that have not been meaningfully addressed elsewhere. These definitions are broad and flexible to ease both their adoption and use by

9 See also JANCINTA ANYANGO ODOUR ET AL., LEFT OUT OF THE BARGAIN: SETTLEMENTS IN FOREIGN BRIBERY CASES AND IMPLICATIONS FOR ASSET RECOVERY 71 (2014).
10 Id.
13 See Section II.C.1–2.
remediation in foreign bribery settlement agreements. The factors comprising the framework have been derived from a study of enforcement actions and analogous areas of the law spanning multiple jurisdictions. They have also been tailored to account for the complicated practical nature of foreign bribery enforcement actions.

This Article proceeds as follows: Section II provides a high-level overview of the practical backdrop to remedial settlement distribution. It begins with an explanation of the modern foreign bribery enforcement landscape before addressing the numerous facets of foreign bribery practice that render remediation through settlement agreements resolutions inherently difficult. In doing so, Section II shows that any framework to guide remediation must be flexible enough to accommodate the importance of enforcement agencies’ prosecutorial discretion, as well as other vicissitudes of this area of practice. Section II concludes with an overview of current approaches to remedial settlement distributions in the U.S., U.K., and Canada. The U.S. and the U.K. are analyzed as they are particularly active in this space compared to other nations. Canada has also been selected, even though it is not an active enforcer of the foreign bribery offense, as it is the only jurisdiction of which the author is aware that has passed legislation specifically addressing remedial settlement distribution. Section II also establishes the relatively narrow scope of this Article. This Article does not consider the victims of securities fraud linked to foreign bribery or civil foreign bribery enforcement. Neither does this Article consider whether or how legislatures might approach remedial settlement distribution, or the provision of legal remedies to the victims of corruption through a legally enforceable right of action. Instead, it focuses solely on the exercise of prosecutorial discretion by enforcement agencies negotiating settlement agreements in foreign bribery cases. Section III introduces key terminology and concepts. It provides working definitions of harm and victimhood that enforcement agencies can easily adopt.

14 It should be noted, however, that the U.S. Supreme Court has recognized the importance of remediation within the context of the civil enforcement of U.S. foreign bribery laws by the U.S. Securities & Exchange Commission. See Liu v. SEC, 140 S. Ct. 1936 (2020).

15 Legislation has been introduced in the U.S. Congress that would apportion a percentage of all criminal FCPA fines and penalties to anti-corruption initiatives. See Countering Russian and Other Overseas Kleptocracy Act, H.R. 3843, 116th Cong. (2019); see also Abigail Bellows, Guest Post: Why the U.S. Congress Should Pass the CROOK Act, GLOBAL ANTICORRUPTION BLOG (July 7, 2020), https://perma.cc/67VM-ESW3. There is also legislation pending in the U.S. Senate that would mandate the transfer of certain FCPA civil settlements to pediatric research. See Gabriella Miller Kids First Research Act 2.0, H.R. 6556, 116th Cong. (2020).

16 Arguments have been made in favor of reforms to create a private right of action under the FCPA. See, e.g., Gideon Mark, Private FCPA Enforcement, 49 AM. BUS. L.J. 419 (2012). However, these proposed reforms have never appeared likely to be passed in either legislative house.
and adapt. Section III then proffers a conceptual breakdown of remediation, demarcating three distinct types: first, compensation, a loss-based remedy applicable to identifiable victims who have suffered ascertainable loss; second, reparations, which respond to the widespread and diffuse harms suffered by populaces en masse; and third, restitution, a gain-based form of remediation that strips ill-gotten gains from corrupt actors and awards them to victims. This Article focuses on compensation and reparations because these are the forms of remediation surfacing most often in practice. Sections IV and V then provide the guiding frameworks. These frameworks assume the form of multifactorial approaches, like those employed by the DOJ and SFO when determining whether to prosecute corporations that have breached foreign bribery prohibitions. The factors comprising the framework for compensation include: (i) whether a victim can be identified; (ii) whether a victim has suffered direct harm; (iii) whether that harm is ascertainable; (iv) whether there is a risk of repeat corruption; and (v) whether compensation is appropriate in the circumstances. The factors comprising the framework for reparations include: (i) whether victims have suffered indirect harm; (ii) whether there is a nexus between the act of bribery and the victims’ harm; (iii) whether there is a risk of repeat corruption; and (iv) whether reparations are appropriate in the circumstances.

II. BACKGROUND, CHALLENGES, AND CURRENT APPROACHES

A. Fundamental Aspects of Foreign Bribery Enforcement

1. The FCPA and the Foundations of the Global Foreign Bribery Regime

The global foreign bribery regime originated from the enactment of the Foreign Corrupt Practice Act of 1977 (FCPA). The FCPA was passed following a string of corruption scandals that came to light in the 1970s. Relevantly, numerous U.S. companies were found to have paid bribes to foreign public officials so as to rig the procurement process for government-awarded contracts in developing nations. The FCPA accordingly prohibited the offering and

---

17 It would be redundant to present a framework on both compensation and restitution. The only substantial difference would be that “the need to ascertain a victim’s loss” would be replaced by “the need to ascertain a corrupt actor’s gains.”


19 For a more comprehensive overview of the historical backdrop to the FCPA, see Mike Koehler, The Story of the Foreign Corrupt Practice Act, 73 OHIO ST. L.J. 929 (2012); Kevin Davis, Why Does the United States Regulate Foreign Bribery: Moralism, Self-Interest, or Altruism?, 67 N.Y.U. ANN. SURV. AM. L.
payment of bribes to influence the acts or decisions of foreign public officials.\textsuperscript{20} Foreign public officials who demanded or accepted bribes remained beyond the statute’s reach. Hence, the FCPA regulated the “supply-side” of foreign bribery.\textsuperscript{21} The FCPA also required companies to make and keep accurate books and records and to devise and implement adequate accounting systems.\textsuperscript{22} The statute was policed weakly until the early 2000s.\textsuperscript{23} The U.S. government then began to devote significant resources to its enforcement. In punishing FCPA breaches, the DOJ has since levied eighty-two fines ranging from US$10 million to US$100 million, twenty-five fines ranging from US$100 million to US$1 billion, and four fines exceeding US$1 billion.\textsuperscript{24} Businesses accordingly devote substantial resources to anti-corruption compliance policies and programs to avoid liability under the FCPA and the similar statutes that other nations have passed in its image.\textsuperscript{25}

The FCPA is buttressed by an international regime, centralized by the Organisation for Economic Co-operation and Development (OECD) Anti-Bribery Convention.\textsuperscript{26} The Convention, which became effective in 1999, requires parties to criminalize supply-side foreign bribery. Forty-three nations have acceded to or ratified it. The FCPA served as a blueprint for the OECD Convention, and it set the contours of foreign bribery enforcement generally. Therefore, understanding the provisions of the FCPA is key to understanding how the global foreign bribery practice functions. However, not every foreign bribery statute follows the FCPA’s structure. The U.K. Bribery Act 2010, for example,\textsuperscript{497} (2012). For a recount of the most notable corporate scandals, see H. Lowell Brown, The Extraterritorial Reach of the U.S. Government’s Campaign against International Bribery, 22 Hastings Int’l L.J. 407, 423–29 (1999). For an account of how the revelations of bribery were perceived at the time, see Charles R. McManis, Questionable Corporate Payments Abroad: An Antitrust Approach, 86 Yale L.J. 215 (1976) (noting that corruption scandals had “shaken foreign governments, rocked American corporate management, and tarnished the image of American private enterprise both at home and abroad”).

More specifically, the FCPA prohibits the payment or offer of payment either directly, indirectly, or through a third party, of money or “anything of value” to an official of a foreign government or political party, with corrupt intent, to obtain or retain business. 15 U.S.C. §§78dd–1–2 (2020).

An ad hoc committee of the Association of the Bar of the City of New York authored a report that was particularly influential in the drafting of the FCPA and, in fact, warned against this sort of “reaching out” by Congress. See Ass’n of the Bar of the City of N.Y., Report on Questionable Foreign Payments by Corporations: The Problem and Approaches to a Solution 5–6 (1977).

15 U.S.C.S. § 78m(a)–(b).


casts an even broader net by simply criminalizing all failures by corporations to prevent bribery.\textsuperscript{27}

Despite widespread ratification of the Convention, foreign bribery enforcement remains concentrated in a small number of nations. Transparency International classifies just six countries, other than the U.S., as “active enforcers” of foreign bribery laws: Germany, Israel, Italy, Norway, Switzerland, and the U.K.\textsuperscript{28}

\section*{2. The Extraterritorial Application of Foreign Bribery Laws}

The FCPA has assumed global prominence due to its extraterritorial application.\textsuperscript{29} In addition to covering “any person” who acts while in the U.S., the statute also applies to “issuers” and “domestic concerns” (as well the agents of issuers and domestic concerns) who make “use of the mails or any means or instrumentality of interstate commerce” in relation to an act of foreign bribery, regardless of where they are situated.\textsuperscript{30} “Issuers” are entities with securities registered in the U.S.\textsuperscript{31} or that are otherwise required to file reports with the U.S. Securities & Exchange Commission.\textsuperscript{32} An entity can meet the definition of “issuer” even if it is incorporated in a nation other than the U.S.\textsuperscript{33} “Domestic concerns” include U.S. citizens, nationals, and residents, as well as businesses organized in or with their principal place of business in the U.S., regardless of where they act.\textsuperscript{34}

The DOJ has interpreted these terms broadly. Liability can attach to foreign nationals and foreign entities for conduct that has only a slight U.S. nexus. For example, the DOJ has asserted jurisdiction simply because a transaction connected to a foreign bribery scheme resulted in money momentarily passing through a U.S. bank account or because a foreign bribery scheme involved electronic communication (such as an email) that passed through an internet server located in the U.S.\textsuperscript{35}

\textsuperscript{28} DELL \& MCDevitt, supra note 2.
\textsuperscript{29} See, e.g., BrANISLAV HOCK, EXTRATERRITORIALITY AND INTERNATIONAL BRIBERY: A COLLECTIVE ACTION PERSPECTIVE (2019); DAVIS, supra note 4.
\textsuperscript{31} Id. § 78c(a)(8). Furthermore, foreign issuers whose American Depository Receipts are listed on a U.S. exchange are “issuers” for purposes of the FCPA.
\textsuperscript{32} Id. § 78o(d).
\textsuperscript{34} 15 U.S.C. § 78dd-1.
Perhaps unsurprisingly, the extraterritorial application of the FCPA and foreign bribery statutes in other jurisdictions has invited allegations of imperialism.\(^{36}\) These allegations are naturally exacerbated by both the fact that corrupt malefactors often have minimal links to the enforcing nation, as well as the failure of enforcing nations to distribute the proceeds of foreign bribery settlements among the victims of foreign bribery. One U.K. civil society organization has identified a “growing realisation that it is unjustifiable for the U.K. government to financially benefit from fines levied against U.K. companies or individuals found guilty of corruption overseas, while not having suffered the damages of these acts, and with very little allocated to compensating the real victims.”\(^{37}\) Another U.S. commentator has remarked, “I am not sure where criminal fines should go when a French company bribes Costa Rican ‘foreign officials,’ but I am pretty sure that the answer should not be 100% to the U.S. Treasury.”\(^{38}\)

3. The Reliance on Non-Trial Resolutions and Prosecutorial Common Law

The vast majority of foreign bribery enforcement actions are resolved through settlement agreements. The most prominent type of agreement wielded by the U.S. and the U.K. is the deferred prosecution agreement (DPA). Enforcement agencies like the DOJ and the SFO use DPAs to resolve criminal investigations against corporate entities without trial provided the corporation cooperates with the agency. The terms of these agreements impose certain obligations upon the corporation. If the corporation does not comply, the enforcement agency may commence prosecution. Hence, the prosecution is “deferred” for a period of time set by the DPA, rather than obviated altogether as is the case when an enforcement action is resolved by a non-prosecution


\(^{37}\) RAID & AFREWATCH, DEMOCRATIC REPUBLIC OF CONGO: CONGO’S VICTIMS OF CORRUPTION 13 (2020) [hereinafter RAID].

\(^{38}\) Is ICE a Victim? And an Open Question!, FCPA PROFESSOR (May 25, 2011), https://perma.cc/N3XC-ZWXN. Professor Koehler’s comment is useful insofar as it prompts us to consider where the proceeds of foreign bribery cases ought to go. However, Professor Koehler has emerged as opposing remedial settlement distribution generally, decrying the practice as providing “feel good measures.” See Am I a Victim?, FCPA PROFESSOR (Apr. 15, 2015), https://perma.cc/32GA-KRGS.
agreement. The conditions imposed by DPAs typically require the corporation to agree to a specific set of facts, to pay a financial penalty, and to implement a program to achieve compliance in the future. To describe settlement agreements, this Article employs the broader term “non-trial resolution.” The term “non-trial resolution” was coined in a 2019 report by the OECD. A non-trial resolution is “any agreement between a legal or natural person and an enforcement authority to resolve foreign bribery cases without a full trial on the merits of the allegations either before or after indictment with sanctions and/or confiscation, irrespective of whether it is a conviction or a non-conviction mechanism.”

Corporate entities and enforcement agencies enter into non-trial resolutions to resolve allegations of foreign bribery without the reputational and financial risks of litigation. The U.S. has resolved 96% of its enforcement actions through non-trial resolutions, while the U.K. and Germany have each resolved 79% of their enforcement actions in this manner. Collectively, these three nations account for 80% of all foreign bribery enforcement and nearly 90% of all non-trial resolutions since the OECD Anti-Bribery Convention came into force. As of 2018, signatories to the OECD Anti-Bribery Convention had collectively levied US$14.9 billion in fines enforcing foreign bribery offences. Of this sum, 95% had been exacted through non-trial resolutions. This reliance on non-trial resolutions is a crucial, yet controversial, feature of the global foreign bribery regime.

A criticism often made is that to avoid the risk of trial, companies simply acquiesce and accept enforcement agencies’ aggressive and broad interpretations of foreign bribery laws, even if there is a strong liability defense. As a consequence of this, foreign bribery laws are rarely interpreted by courts, and an informal body of precedent has arisen based on the practices of enforcement agencies. This body of informal precedent, also referred to as “prosecutorial

---

39 For a brief overview of DPAs in the U.S. and U.K., see OECD, supra note 7, at 52.
40 Id. at 11.
41 Id. at 13.
42 Id.
43 Id. at 14–15.
44 See Nicholas Lord & Colin King, Negotiating Non-Contention: Civil Recovery and Deferred Prosecution in Response to Transnational Corporate Bribery, in CORRUPTION IN COMMERCIAL ENTERPRISE: LAW, THEORY AND PRACTICE 234 (Routledge, 2018) (offering a U.K. perspective and an explanation of how the increased use of DPAs can be likened to an “accommodation” of foreign bribery); David M. Uhlmann, Deferred Prosecution and Non-Prosecution Agreements and the Erosion of Corporate Criminal Liability, 74 MD. L. REV. 1295 (2013) (presenting a U.S. perspective).
46 See id. (“Against the backdrop of aggressive enforcement and the resulting multi-million-dollar fines and penalties is the undeniable fact that, in most instances, there is no judicial scrutiny of the FCPA enforcement theories. The end result is that the FCPA often means what the enforcement agencies
common law,” has cyclically entrenched expansive interpretations of foreign bribery laws, including their extraterritorial reach.\footnote{Remediation in Foreign Bribery Settlements, Hickey (Winter 2021) 377.} Another feature of the global regime relevant to this Article is the fact that almost all of the money raised through foreign bribery enforcement is retained by the treasury of the enforcing nation.\footnote{31 U.S.C. § 3302(b) is the statute responsible for the deposit of these funds into the U.S. Treasury. See also ODUOR ET AL., supra note 9, at 71 (“Between 1999 and July 2012, a total of about $4.2 billion was collected in criminal monetary sanctions. About 71 percent of the criminal sanctions were imposed in the form of fines. Confiscations and forfeitures made up about 26.3 percent of the total, with 2.4 percent from restitution or reparations and 0.3 percent imposed in legal or procedural costs.”).} This is so despite the fact that prominent enforcement agencies often refer to the harm foreign bribery causes vulnerable persons in the developing world as one of the justifications for enforcement.\footnote{U.S. DEP’T OF JUST., supra note 8 (“For let there be no doubt that corruption is not a victimless offense. Corruption is not a gentlemen’s agreement where no one gets hurt. People do get hurt. And the people who are hurt the worst are often residents of the poorest countries on the face of the earth, especially where it occurs in the context of government infrastructure projects, contracts in which crucial development decisions are made, in which a country will live by those decisions for good or for bad for years down the road, and where those decisions are made using precious and scarce national resources.”).} Remedial settlement distribution, as shown below, is the exception and not the rule.

4. The United Nations Convention Against Corruption

The OECD Convention is complimented by another international instrument: The United Nations Convention Against Corruption (UNCAC).\footnote{U.N. Convention against Corruption, adopted Oct. 31, 2003, 2349 U.N.T.S. 41 (entered into force Dec. 14, 2005).} The U.S. and the U.K. are both parties to this instrument. Chapter V of the UNCAC generally provides that state parties must repatriate embezzled wealth and property derived from corruption offences.\footnote{Id. arts. 53, 57.} However, this instrument has

\textit{say it means. Because of the ‘carrots’ and ‘sticks’ relevant to resolving a government enforcement action, FCPA defendants are nudged to accept resolution vehicles notwithstanding the enforcement agencies’ untested and dubious enforcement theories or the existence of valid and legitimate defenses.”).}

\textit{Bingham’s Michael Levy on the Rise of Prosecutorial Common Law, CORPORATE CRIME REPORTER (Feb. 7, 2011), https://perma.cc/CU5L-HWVF (“Companies don’t want to have to take a conviction. But companies don’t particularly have a strong incentive to fight on complex legal theories around jurisdiction and liability. What they care about is—what is going to be the dollar amount of the fine? Is it going to affect key employees? What is going to be the reputational damage? . . . The government on the other hand has a broader institutional interest. The government has consistently used the common law of settlement to push more and more aggressive interpretations of some of the laws it is interpreting, getting companies to agree to those aggressive interpretations, and then using the fact that a large company agreed that this was a violation of law to bring a similar case against other companies. . . . The fact that one company agreed that such and such was the law doesn’t make such and such the law. Only judges and courts can decide what the law is—not prosecutors and settling companies.”).}
not prompted the majority of enforcing nations to use the proceeds of non-trial resolutions negotiated in foreign bribery cases for remedial purposes. The money that companies pay as punitive fines is not property derived from corruption offences, and it is unlikely that the UNCAC has any bearing upon the way in which enforcement agencies are obligated to deal with the proceeds of DPAs and other non-trial resolutions.\textsuperscript{52} Even if the UNCAC could be construed in this manner,\textsuperscript{53} it is clear that this is not how the most prominent enforcer of the foreign bribery offense, the U.S., has interpreted it.\textsuperscript{54} The practices of the DOJ reveal a distinction drawn between, on one hand, the seized or forfeited proceeds of corruption, and, on the other, the punitive fines paid out by corporate entities resolving foreign bribery offences through non-trial resolutions.\textsuperscript{55} The fact that neither the UNCAC nor the OECD Convention seemingly control remedial settlement distribution reinforces the prosecutor-led nature of remedial settlement distribution.

B. Practical Barriers to Remedial Settlement Distribution

Numerous features of foreign bribery enforcement render remediation difficult. At the outset, it is important to acknowledge the role of prosecutorial discretion. Enforcement agencies must retain the discretion to deviate from the framework proposed in this Article when the situation demands it. It would be a mistake to base a framework on a set of mandatory rules and then contend that enforcement agencies should adhere to those obligations without exception. Without legislative backing, such an approach would come to nothing. Remedial settlement distribution emerged from the practices of enforcement agencies, and by and large it remains a fruit of prosecutorial discretion. A framework is, therefore, useful only if enforcement agencies are able and willing to adopt it.


\textsuperscript{53} Professor Kevin Davis has argued that the “spirit” of the UNCAC leans toward the sharing of the proceeds of foreign bribery non-trial resolutions. Davis, supra note 4, at 221. Davis does not, however, substantiate this argument with meaningful analysis. A similar argument is raised in RAID, supra note 37, at 15.

\textsuperscript{54} Jennifer Shasky, former supervisor of the DOJ’s Kleptocracy Asset Recovery Initiative, has even gone so far as to suggest that the DOJ’s asset repatriation efforts are not required under international law. Shasky is reported to have said that “there is no legal requirement to return the funds at all.” See Christopher M. Matthews, Fledging Kleptocracy Initiative Faces Challenges, Expectations, Just Anti-Corruption (Sept. 19, 2011), https://perma.cc/F8NT-NZ53.

\textsuperscript{55} See Larissa Grey et al., Few and Far: The Hard Facts on Stolen Asset Recovery (2014) (providing an overview of the difficulties in seeking the return of stolen wealth in accordance with the terms of the UNCAC; see also Section II.C.2 (providing a high-level overview of the different approaches taken by the Fraud Section and the Money Laundering and Asset Recovery Section of the DOJ).
Assuming that legislatures in prominent foreign bribery enforcing jurisdictions will not adopt a comprehensive approach to remedial settlement distribution in the foreseeable future, the most productive discussion that can be had is one that is aimed towards enforcement agencies and that shows deference to the role of prosecutorial discretion.

This Section pursues two lines of thought. First, it explains why prosecutorial discretion in foreign bribery enforcement actions cannot always be expected to accommodate a remedial agenda. Second, it shows that despite the difficulties in constructing a framework for remedial settlement distribution, there is reason for taking on this challenge and improving current practices.

1. The Demands and Implications of Prosecutorial Discretion

The futility of arguing for a framework based on mandatory rules that are not statutorily entrenched stems from three interconnected features of foreign bribery enforcement and the function of enforcement agencies. The first is that the negotiation of non-trial resolutions is an inherently and necessarily discretionary process. The negotiation of non-trial resolutions is subject to considerations that are innumerable and case dependent. These might include the strength of an enforcement agency’s evidence, the resources available to both parties, the risk involved in litigating a particular point of law, the skill of the negotiators representing each side, and so on. Second, the transnational character of many financial crime offenses invites political influences to come into play. Many anti-corruption agencies are politically accountable and cannot pursue a foreign bribery investigation when the executive has forbidden it, while some agencies are beholden to overarching political commands that are capable of influencing, undermining, and determining their actions. The third feature of foreign bribery practice complicating remediation is the limited competences and resources of anti-corruption agencies. Enforcement agencies are not adjudicative bodies equipped to calculate loss and formulate remedies for the harm caused by foreign bribery schemes. Nor are they able to assume the mandate of foreign aid agencies and administer complex systems that ameliorate the harms that corruption can cause to societies at large.


57 See R (on the Application of Comer House Research and Others) v. Director of the Serious Fraud Office, [2008] UKHL 60 (appeal taken from Eng.) (illustrating an example of the U.K. government exercising its power to halt a foreign bribery investigation initiated by the SFO on what arguably were national security grounds); see also Susan Rose-Ackerman & Benjamin Billa, Treaties and National Security, 40 N.Y.U. J. INT’L L. & POL. 437 (2008).


59 See Section IV.
The corollary of these three factors is that agencies need to maintain a degree of discretion to respond to factual peculiarities that arise case-by-case and to operate within the confines of particular political agendas and their own resources and capabilities. Enforcement agencies may therefore have to deviate from the framework proposed below for the sake of resolving an investigation. The degree and nature of possible deviations are not discussed further. This is because it is impossible to anticipate what might warrant deviation in future cases. What is important to note is that the framework proposed by this Article is intended to operate subject to prosecutorial discretion.60

2. Justifications for Improving Extant Practice

The practical challenges of pursuing remedial settlement distribution outlined above must be considered in light of the fact that enforcement agencies have already begun to practice remedial settlement distribution, while numerous governments have committed, at a policy level, to pursuing remediation in foreign bribery cases. Consequently, the aforementioned challenges do not justify neglecting to develop a framework to guide remedial settlement distribution.

This Article does not delve deeply into normative arguments for or against remedial settlement distribution. Debating the merits of the practice is beyond its scope. However, for the sake of providing broader context to the underlying policy debates, it is worth at least canvassing some of the primary arguments for and against remedial settlement distribution.

Those in favor of remedial settlement distribution typically rely on policy and ethical considerations. Perhaps unsurprisingly, there is a growing sentiment among senior personnel of enforcement agencies,61 civil society organizations,62 and NGOs63 that some portion of the proceeds of foreign bribery non-trial

60 There are of course additional difficulties, such as the risk of remediation monies being repurposed for corrupt ends, diplomatic and security risks associated with providing remediation monies to certain countries, and difficulty in formulating remedies for the diffuse types of harm caused by corrupt practices. These risks are considered below, as they are dealt with as considerations that come into play when determining whether and how to pursue remedial settlement distribution. The difficulties that have been raised here pertain to the general functions of enforcement agencies.

61 See, e.g., Richard Vanderford, Victims of Other Companies’ Foreign Bribery Should Come Forward to Seek Cash, DOJ Says, MLEX GLOBAL ADVISORY (Dec. 3, 2019) (“On December 3, 2019, Daniel Kahn, the Principal Deputy of the Criminal Fraud Section at DOJ, went a step further and invited purported victims of FCPA cases to make a claim: ‘If there is a victim of a crime, [DOJ] of course wants those victims to come forward. That is our primary objective, to ensure that victims are made whole.’”); Baker, supra note 8 (“The ideal outcome, where-ever it is possible, is for the money secured through asset recovery to be returned to victims, using that term in the widest sense.”).


63 ODUOR, ET AL., supra note 9. Note also the attempt of a Nigerian NGO to convince the Securities and Exchange Commission to divest a portion of settlement proceeds to assist the citizens of
resolutions should be diverted to victims in developing nations. The rationale supporting this sentiment can be expressed in numerous ways. One view is that it is ethical to remediate the harm caused by corruption in the developing world because corruption is pervasive and citizens in developing nations are vulnerable to its corrosive impact. Another position, which goes a step further, is that it is unethical for Western nations to extract billions of dollars in fines and penalties through non-trial resolutions for conduct that has occurred in developing nations, and that has harmed governments and citizens in developing nations, only to retain that money as public revenue.\(^{64}\) Another argument is that diverting the proceeds of foreign bribery enforcement away from enforcement agencies and their respective governments and toward developing nations in which bribery occurs would reduce the incentive of enforcesing nations to enforce these laws overzealously in a rent-seeking fashion, and instead incentivize developing nations to police foreign bribery.\(^{65}\)

Buttressing these claims, it is clear that there are no other available methods for addressing the harm caused by corruption. Foreign bribery non-trial resolutions have emerged as practical mechanisms that are at least capable of remediating some of the harm done to victims of corruption in the developing world in certain circumstances.\(^{66}\) The global asset recovery regime has not emerged as a completely effective means of repatriating stolen assets to governments, let alone as a solution to the economic and social havoc to civilians caused by corrupt practices.\(^{67}\) Further, private rights of action have proved largely ineffective, for numerous reasons, at providing individual victims with remedies.\(^{68}\) These include a lack of access to courts for individuals in developing nations and difficulties posed by legal standards that arise as preconditions to the commencement of litigation, such as standing, venue, and jurisdiction.\(^{69}\) Difficulties in pursuing remediation through courts are further intensified due to the necessity of proving causation and of satisfying legislative definitions of

---

\(^{64}\) Spahn, \textit{supra} note 36; RAID, \textit{supra} note 37, at 13.

\(^{65}\) Turk, \textit{supra} note 12.

\(^{66}\) ODUOR \textit{et al.}, \textit{supra} note 9, at 2 ("The reality is that, in the majority of settlements, the countries whose officials were allegedly bribed have not been involved in the settlements and have not found any other means to obtain redress.").

\(^{67}\) See generally GRAY \textit{et al.}, \textit{supra} note 55, at 19; KEVIN STEPHENSON \textit{et al.}, \textit{Barriers to Asset Recovery: An Analysis of the Key Barriers and Recommendations for Action} (2011).

\(^{68}\) See Mark, \textit{supra} note 16 (arguing in favor of including a private right of action in the FCPA and pointing out numerous shortcomings in those few rights and remedies available to private individuals harmed by foreign bribery).

victimhood and standing that were not fashioned for the foreign victims of corruption.\textsuperscript{70} Taking all these considerations into account, it is worthwhile to consider less conventional avenues of redress.

There are, of course, strong ripostes to these views, including that enforcement agencies are government-funded entities enforcing national laws and their governments are entitled to retain whatever money is raised through enforcement.\textsuperscript{71} Indeed, it was never the purpose of the FCPA to cater to the victims of foreign bribery. In passing the FCPA, Congress was primarily concerned with the economic and political effects of foreign bribery on U.S. interests.\textsuperscript{72} Proponents of this view might also argue that the nations enforcing foreign bribery prohibitions should not share the proceeds of their enforcement efforts with other nations that have left them to act alone in these efforts.\textsuperscript{73} Another argument is that enforcement agencies are already assisting developing nations by fighting corruption on a transnational basis.\textsuperscript{74} Detractors might also allege that funneling capital back into corruption-prone states is counterproductive and provides corrupt governments with more resources to support illicit activities (this particular issue is dealt with in depth below).

Additional arguments, both for and against distributing the proceeds of foreign bribery non-trial resolutions, can be made. However, it is beyond the scope of this Article to consider whether one nation might be obliged, legally or morally, to share proceeds. What is ultimately important for the purposes of this Article is that there is a growing sentiment among influential stakeholders that remediation is an agenda worth pursuing and that enforcement agencies and governments are beginning to recognize and support this agenda. As such, this Article simply adopts the view that it is better to instill a degree of principle into the practice than to leave the practice entirely unprincipled. The merits and demerits of remedial settlement distribution itself present a distinct, and much broader, issue that is beyond the scope of this Article.

\textsuperscript{70} Stephenson, supra note 62.


\textsuperscript{72} Koehler, supra note 19, at 929.

\textsuperscript{73} The Chairman of the U.S. Securities and Exchange Commission, Jay Clayton, has emerged as a proponent of these views, noting that the U.S. is essentially alone in its enforcement efforts, while some other nations attempt to game the system by paying bribes to gain an edge in foreign markets over U.S. companies that do not pay bribes. Jay Clayton, Chairman, U.S. Sec. & Exch. Comm’n, Remarks to the Economic Club of New York (Sept. 9, 2019), https://perma.cc/K9BQ-MSHE.

\textsuperscript{74} Id.
This Article has a narrow purpose: posing suggestions for the improvement of enforcement agency techniques regarding remedial settlement distribution. This Article proceeds upon the premise that, if remedial settlement distribution is to occur, it should be subject to objective and palpable criteria to allow for the development of a body of informal precedent and best practices over time. The development of such a body of precedent will increase consistency of outcomes in this space by allowing enforcement agencies to refer to past decisions for guidance. Current practices are marred by uncertainty: there are no tools agencies can refer to for the purpose of determining whether a particular non-trial resolution should include remediation. Similarly, there exists no guidance on identifying victims, ascertaining loss, or managing the risk that monies put toward remediation might be repurposed for corrupt ends. The adoption of a guiding framework will go a long way toward curing the inconsistency and lack of transparency that has burdened remedial settlement distribution in its infancy. Furthermore, the adoption of a framework to guide remediation in non-trial resolutions is important for the sake of accountability. If adopted, the framework set out below would prevent enforcement agencies from ignoring remediation where it is warranted and from pursuing it where it is not. At the very least, it would require them to justify their reasons for pursuing or refusing to pursue remediation. The importance of accountability cannot be understated. Foreign bribery non-trial resolutions involve substantial sums of money, and that money should not be subject to disposal in a manner that is unprincipled and without justification.

In sum, the creation of a clear yet flexible framework will improve consistency of outcomes and deliver a degree of much needed principle and accountability into an area of foreign bribery practice that implicates the interests of those harmed by corruption.

C. Current Approaches to Remedial Settlement Distribution

The remainder of this section outlines the approaches of the U.K., U.S., and Canada to remedial settlement distribution. The U.K. has been selected because it has pursued remediation for corruption victims in the developing world through foreign bribery non-trial resolutions more than any other jurisdiction. The U.S., on the other hand, is worth discussion because it is the most active enforcer of the foreign bribery offense and has pursued a remedial agenda through non-trial resolutions on several occasions. Meanwhile, Canada is neither an active enforcer of the foreign bribery offense, nor does it have a history of pursuing remediation in foreign bribery enforcement. However, Canada is the only jurisdiction the author has encountered that has a dedicated statutory regime geared toward remediation in foreign bribery settlements. While other prominent enforcers of
the foreign bribery offense, such as Germany, have successfully obtained modest forms of remediation, they are not discussed here due to language limitations.75

1. United Kingdom

U.K. government agencies, primarily the SFO, the Department for International Development (DFID), and the Foreign and Commonwealth Office (FCO), are largely responsible for pioneering remedial settlement distribution in foreign bribery cases.76 The SFO investigates and prosecutes foreign bribery and negotiates the terms of non-trial resolutions, while the DFID and occasionally the FCO formulate and administer reparations programs in developing nations in accordance with the terms of those agreements. The SFO entered into a string of non-trial resolutions between 2010 and 2017 in which some form of remediation was negotiated for the benefit of a government or populace harmed by either foreign bribery or some other form of corrupt conduct.77 The DFID, the FCO, the SFO, and various other instrumentalities of the U.K. government assisted in accomplishing these remediation efforts. The SFO has also pursued remediation in criminal prosecutions as well as criminal and civil forfeiture cases, which are instructive insofar as they employ remedial techniques that can be adapted for non-trial resolutions.78 The SFO practice in this area is particularly illuminating because DPAs negotiated by the SFO are given legal force only if they are judicially approved, and every DPA approval judgment published to date has considered whether remediation ought to be awarded. Remedial settlement distribution in the U.K. is unique in that it is accompanied by a small amount of relevant judgments. The U.K DPA regime came into effect in February 2014 and provides that judges should approve DPAs only if they are likely to be in the interest of justice and their terms are fair, reasonable, and proportionate.79 That regime also specifies that a DPA “may impose” a requirement to “compensate

75 See OECD, IMPLEMENTING THE OECD ANTI-BRIBERY CONVENTION: PHASE 4 REPORT: GERMANY 34, 89–95 (2018). In bribery cases, German courts may suspend individual defendants’ prison sentences and instead impose various other conditions, including a requirement that the defendant make specified contributions to charity. Id. at 34. While courts imposed a handful of such conditions in the last decade, the amounts were relatively small. See id. at 89–95 (noting required charitable contributions ranging from €22,000 to €175,000).

76 The cooperation of various U.K. government instrumentalities is reflected in the U.K. compensation principles. SERIOUS FRAUD OFFICE, GENERAL PRINCIPLES TO COMPENSATE OVERSEAS VICTIMS (INCLUDING AFFECTED STATES) IN BRIBERY, CORRUPTION AND ECONOMIC CRIME CASES (2018).

77 A complete account of DPAs entered into by the SFO can be found on the SFO website. Deferred Prosecution Agreements, SFO, https://perma.cc/2KT8-PXLK.

78 Civil recovery orders, which permit forfeiture of the proceeds of crime in the absence of a criminal conviction, have also been a point of controversy in England and Wales. See Jennifer Hendry & Colin King, How Far Is Too Far? Theorising Non-conviction-based Asset Forfeiture, 11 INT’L J.L. IN CONTEXT 398 (2015).

79 Crime and Courts Act 2013, c. 22, § 45, sch. 17 (UK).
victims of the alleged offence” or “donate money to a charity or other third party.”  Also in 2014, a prosecutorial manual and a sentencing guideline were published, encouraging the SFO and U.K. courts to consider awarding remediation wherever possible. To date, U.K. government agencies have sought approximately £33 million in remediation and have obtained a total of approximately £602 million through non-trial resolutions.

Five cases exemplify the various ways that the U.K. has dealt with the question of remediation in corruption cases. The first involved the corrupt conduct of U.K. weapons manufacturer BAE Systems Plc. The SFO and BAE entered into a settlement agreement in February 2010. The settlement agreement pertained to accounting malpractice in the sale of a radar system to the Tanzanian government. Clause 5 of the agreement is particularly relevant. It provided for reparations to be paid by BAE. Specifically, BAE was to “make an ex gratia payment for the benefit of the people of Tanzania in a manner to be agreed between the SFO and [BAE]. The amount of the payment shall be £30 million less any financial orders imposed by the court.” DFID played a key role in orchestrating this scheme, brokering the arrangement and monitoring its execution. Having a presence in Tanzania, DFID was able to play a monitoring role to ensure that the funds were used for their intended purposes. Second, in 2014, U.K. print company Smith & Ouzman, along with four of its directors, was found guilty at trial of bribery-related offenses involving public officials in Kenya, Ghana, Mauritania, and Somaliland. The case did not involve a non-trial resolution, as it proceeded to judgment, and Smith & Ouzman and key directors

80 Id. § 5(3)(b)–(c).
81 See SENTENCING COUNCIL, FRAUD, BRIBERY AND MONEY LAUNDERING OFFENSES: DEFINITIVE GUIDELINE 45.
83 This case was accompanied by significant controversy that is beyond the scope of this Article. See UK Wrong to Halt Saudi Arms Probe, BBC NEWS (Apr. 10, 2008), https://perma.cc/636Q-SCD6; Lords Says SFO Saudi Move Lawful, BBC NEWS (July 30, 2008), https://perma.cc/FTA6-6RNA.
84 Id.
85 Id.
were convicted. In its sentencing remarks, the court found that neither compensation nor reparations was appropriate because there was a risk that money paid in compensation might be repurposed for corrupt ends.\(^89\) However, though remediation was not ordered as part of the sentencing phase, the SFO, the DFID, and the FCO later identified a method of delivering reparations to Kenyan citizens: using part of the money extracted through fines and penalties to purchase eleven ambulances for public hospitals in Kenya.\(^90\)

Third, in 2015, Standard Bank Plc, a U.K. bank, self-reported that one of its subsidiaries had paid bribes in connection with a public contract obtained in Tanzania. The contract in question granted the subsidiaries the right to act as underwriters for the Tanzanian government for purposes of a sovereign note placement. After the subsidiaries won the right to act as underwriters, one of them entered into a sham consulting agreement to provide a kickback to a consulting company owned by Tanzanian public officials. This agreement stipulated that one percent of all proceeds raised during the underwriting process would be advanced to the shell company for nonexistent consulting services. The note placement ultimately raised US$600 million, so US$6 million was advanced to the shell consulting company. After the payment was discovered, Standard Bank entered into a DPA with the SFO. Because the Tanzanian government would have received the US$6 million paid to the shell consulting company but for the corrupt scheme, it was ordered that this money would be paid in compensation to the Tanzanian government (plus interest).\(^91\)

Fourth, in July 2016, U.K.–based design and manufacturing company Sarclad Ltd. entered into a DPA with the SFO after Sarclad’s management self-reported suspicious activities that it had identified in its operations in Asia. It eventually came to light that Sarclad had entered into at least twenty-eight contracts with foreign governments between 2004 and 2012 that had been procured through bribery. The bribes had been paid throughout numerous Asian countries via a complex web of third-party intermediaries. A total of £17.24 million was paid to Sarclad as a result of the twenty-eight contracts. Compensation was deemed inappropriate. In a decision that has attracted criticism,\(^92\) the court

\(^89\) Id.

\(^90\) Max Golbart, £2m Smith & Ouzman Fine Funds African Development, PRINTWEEK (Mar. 21, 2017), https://perma.cc/2QGE-TYEW.


\(^92\) The facts of Sarclad and the reasons offered for denying remediation in that case were analogous to another DPA concluded by the SFO with Rolls Royce Plc. The reasoning in Rolls Royce has been criticized on the basis that it shows a failure to engage meaningfully with the complexities that need to be addressed if remediation is to become a staple of anti-corruption enforcement. See Matthew Stephenson, Guest Post: The UK’s Compensation Principles in Overseas Corruption Cases—a New Standard for Aiding Victims of Corruption?, THE GLOBAL ANTICORRUPTION BLOG (July 5, 2018),
found that the SFO had not been able to identify victims due to the complexity of the bribe schemes. Moreover, the use of intermediaries had made it impossible to identify victims or quantify loss. Criticism is warranted inasmuch as this case shows that the same criminal restitution schemes used to remediate the harm to victims of nonfinancial crimes are unlikely to be able to address the type of harm suffered by the victims of corrupt practices. Finally, in March 2018, Canadian energy company Griffiths Energy used a front company to bribe Chadian diplomats by offering them significant discounts on its shares. The wife of a former Chadian public official purchased 800,000 shares at less than CA$0.001 each and sold them for a significant profit. The sale of the shares generated £4.4 million, and the SFO obtained property-freezing and forfeiture orders against those proceeds. The DFID then orchestrated the investment of these funds in infrastructure and development projects in Chad.

Each of these cases reflects the commitment of the SFO to achieving some form of redress for the victims of corruption. In June 2018, the U.K. government formally cemented this commitment by adopting a policy of remedial settlement distribution with the publication of \textit{General Principles to Compensate Overseas Victims (Including Affected States) in Bribery, Corruption and Economic Crime Cases} (also known as \textit{Compensation Principles}). However, despite its title, this document does not set out a principled approach to remediation. It reads as a list of overarching instructions, couched in qualifying terms, that create a sense of low modality. \textit{Compensation Principles} provides, among other things, that various U.K. enforcement agencies, including the SFO, “will consider the question of compensation in all relevant cases.” In cases in which “compensation is

\url{https://perma.cc/LMC9-SUUA} (“The picture that emerges is that compensation is not straightforward, even when the government commits to pursuing whatever legal means to achieve it. While the Code of Practice for UK Deferred Prosecution Agreements (DPAs) states that it is ‘particularly desirable’ for compensation to be paid, and despite the fact that under a DPA it should be easier for a prosecutor to get a company to agree to pay compensation, only one of the three UK DPAs concluded so far for foreign bribery has included compensation to affected countries or victims. In the other two (including one involving Rolls Royce), no compensation was provided, for two related reasons: First, the cases were ‘too complex’ involving bribes paid across multiple jurisdictions, and, second, it was not obvious who the victims were. . . . [T]he UK enforcement bodies need to develop mechanisms that enable compensation determinations to be made even in complex cases. The current legal landscape, whereby compensation can only be given in ‘simple’ cases, has left the UK in the somewhat bizarre situation whereby the more widely a company bribes, the more global its wrongdoing, and the more it uses intermediaries to pay the bribes, the less likely it is to have to pay out compensation to countries or individuals affected by its wrongdoing. Complexity should not be an insurmountable obstacle to compensation.”) (emphasis added).


94 Saleh v. Dir. of the Serious Fraud Office [2017] EWCA (Civ) 18 (appeal taken from Eng.), \url{https://perma.cc/RY4R-TS94}.

95 See RAID, supra note 37, at 15, for a list of recommendations to the SFO aiming to improve the operation of the Compensation Principles.
appropriate,” those agencies are to “use whatever legal means are available to secure it.” The agencies are also to “work collaboratively” with other U.K. government entities to identify who should be regarded as potential victims overseas[,]… assess the case for compensation, obtain evidence which may include statements in support of compensation claims, ensure the process for payment of compensation is transparent, accountable and fair, [and] identify a suitable means by which compensation can be paid to avoid the risk of further corruption.\footnote{SERIOUS FRAUD OFFICE, supra note 76; see also HM GOV’T, United Kingdom Anti-Corruption Strategy 2017–2022: Year 1 Update 26 (2018).}

Unfortunately, the principles do not provide any guidance regarding which cases will be “relevant” for the purposes of remediation, when remediation will be “appropriate,” or the means that might be used to secure remediation. Further, Compensation Principles does not significantly alter preexisting practices—it directs the SFO, DFID, and FCO to do what they were already doing. Nevertheless, the existence of Compensation Principles reflects a palpable commitment by the U.K. government to remedial settlement distribution. Most importantly, they are proof of the way in which the agencies’ practices gained traction and eventually led to a shift in government policy.

2. United States

The FCPA Unit of the DOJ’s Fraud Section does not typically practice remedial settlement distribution in its enforcement of the FCPA, but there have been some exceptions. For example, in September 2018, the DOJ entered into a non-prosecution agreement with Brazilian oil company Petrobas following numerous breaches of the FCPA. The terms of the agreement required Petrobas to pay a criminal penalty of approximately US$853 million. Of that sum, approximately US$682 million would be paid to a Brazilian agency, Ministerio Publico Federal.\footnote{See Letter from Sandra Moser, Acting Chief, Fraud Section, Crim. Div., Dep’t of Just., to F. Joseph Warin, Gibson, Dunn & Crutcher LLP (Sept. 26, 2018).} In a press release, Petrobas revealed that the money paid to this agency would be “deposited by Petrobras into a special fund in Brazil to be used in strict accordance with the terms and conditions of the consent agreement, including for various social and educational programs to promote transparency, citizenship and compliance in the public sector.”\footnote{Petrobras Reaches Coordinated Resolutions with Authorities in the United States and Agreement to Remit Bulk of Associated Payments to Brazil, PETROBAS (Sept. 27, 2018), https://perma.cc/N3H4-2C84. It must be noted that this scheme was not executed as smoothly as had been planned. The highest appellate court in Brazil at first found this attempt at remediation to be unconstitutional. Since then, the money has been diverted away from general infrastructure projects and toward fighting the COVID-19 pandemic. See Kevin Abikoff & Aline Osorio, Corruption Settlements, Coronavirus and the Road Paved with Good Intentions, FCPA PROFESSOR (Apr. 2, 2020), https://perma.cc/33NL-SMY2.} In contrast with the U.K.’s
efforts in this area, the DOJ does not have the same institutional commitment to remediation. Rather than orchestrate the method of payment itself, the DOJ simply redirected enforcement proceeds to the anti-corruption agency of the state in which bribes had been paid. This is unsurprising. Unlike the SFO, the DOJ has no enshrined responsibility to consider the question of remediation in foreign bribery cases. Nevertheless, the DOJ’s practice in the Petrobas enforcement action reveals that the world’s most active and influential anti-corruption enforcement agency is willing to pursue remediation in certain instances.

The DOJ’s relative inactivity in this field should also be understood in light of the existence of a separate statutory regime that provides for the payment of remediation through the judicial system. Those who allege they have sustained losses due to a foreign bribery conspiracy have the option of seeking to recoup their losses under the Mandatory Victims Restitution Act of 1996 (MVRA). Certain requirements under this statute (especially victimhood status) make recovery difficult, such that only a small handful of foreign governments have successfully claimed remediation. Indeed, the MVRA has been criticized as imposing too high a bar for recovery by only a narrow class of claimants for it to meaningfully contribute to efforts to remediate the harm to victims of corruption. Moreover, the statute does not provide for those who seek to rectify societal harm or harm suffered by citizens in developing nations.

It is important not to discredit the U.S.’s commitment to remediating harm to victims of corruption. Under the novel Kleptocracy Asset Recovery Initiative, the DOJ’s Money Laundering and Asset Recovery Section works in partnership with other U.S. enforcement agencies to forfeit the proceeds of

---

99 See 18 U.S.C. § 3663(a)(1)(A). Although the remedy provided under this statute is termed “restitution,” it calculates remediation using a loss-based formula, thereby likening it to the concept of compensation discussed below. Note that this provision does not authorize restitution for a substantive breach of the FCPA, but it can provide restitution for losses incurred over the course of a conspiracy to violate the FCPA. Section 3663(a)(1)(A) provides restitution for breaches of offenses contained under Title 18 of the U.S. Code, but the FCPA is an offense under Title 15. However, the general conspiracy offense, 18 U.S.C. § 371, encompasses “any offense against the United States,” including conspiracies to violate the FCPA, and thus opens the possibility of obtaining restitution under the MVRA. Some corporate entities have tried unsuccessfully to attain victim status under the MVRA and other related statutes. See, e.g., United States v. Alcatel-Lucent France, SA, 688 F.3d 1301 (11th Cir. 2012).


101 Stephenson, supra note 52; Spalding, supra note 11, at 1412; see also Shane Frick, “Ice” Capades: Restitution Orders and the FCPA, 12 RICH. J. GLOB. L. & BUS. 433, 437 (2013) (noting that the statutory regimes for victim compensation in the U.S. “fail to address FCPA victims’ needs in almost every conceivable scenario”).

102 For an informative discussion of the initiative, including an assessment of its success, see Pablo J. Davis, “To Return the Funds at All”: Global Anti-Corruption, Forfeiture, and Legal Frameworks for Asset Return, 47 U. MEMPHIS L. REV. 291 (2016).
corruption offenses and, where appropriate, return those proceeds to benefit the people harmed.\footnote{Kleptocracy Asset Recovery Rewards Act, HR 389, 116th Cong. (2019).} Put another way, the DOJ repatriates assets after it has successfully sought the forfeiture of the proceeds of crime through the U.S. judicial system, even though it does not typically redistribute the proceeds of FCPA settlements. The Kleptocracy Asset Recovery Initiative was founded in 2010, and, as of 2016, had returned approximately US$63 million in bribery proceeds and embezzled funds to the victims of corruption.\footnote{Loretta E. Lynch, Attorney General, Remarks at the Organization for Economic Co-Operation and Development Anti-Bribery Ministerial Meeting, U.S. DEPT OF JUSTICE (Mar. 16, 2016), https://perma.cc/3XU6-977C.} Asset recovery, which focuses on the receipt or demand-side of foreign bribery rather than the supply-side, often involves the use of non-trial resolutions and the payment of reparations. The efforts of the DOJ in this field are therefore instructive with regard to how FCPA settlements might be drafted for the purpose of remedial settlement distribution in foreign bribery cases. Indeed, several cases concluded as part of the Kleptocracy Asset Recovery Initiative are discussed in depth below.

Of course, the amount of money repatriated under the Kleptocracy Asset Recovery Initiative pales in comparison to the fines and penalties that the DOJ has levied in FCPA cases.\footnote{See generally ODUOR ET AL., supra note 9.} But despite its relative inactivity in seeking remediation through its enforcement of the FCPA, the U.S. remains a global leader in the field of asset repatriation and remediation in corruption cases generally.

3. Canada

Canada does not enforce foreign bribery prohibitions to the same extent as either the U.S. or the U.K.\footnote{Whereas the U.S. had concluded 207 non-trial resolutions to impose sanctions on legal persons in connection with a foreign bribery scheme and the U.K. had concluded 11, Canada had concluded just 3 such resolutions. OECD, supra note 7, at 107 n.202.} However, unlike those two countries, Canada’s federal legislature has adopted a statutory framework for non-trial resolutions that explicitly contemplates remediation in certain cases involving economic crime, including foreign bribery.\footnote{For an overview of the procedure that preceded this development, see id. at 36–37.} In late 2018, Canada amended its federal Criminal Code to introduce “remediation agreements.”\footnote{The Criminal Code defines “remediation agreement” as “an agreement, between an organization accused of having committed an offence and a prosecutor, to stay any proceedings related to that offence if the organization complies with the terms of the agreement.” R.S.C. 1985, c C-46, § 715.3.} Remediation agreements, like other non-trial resolutions, allow corporations to report certain economic crimes

\footnote{\textcopyright{} Chicago Journal of International Law.}
and to enter into agreements to resolve cases without risking conviction.\(^{109}\) Canadian remediation agreements are similar to the U.K. DPA model in that they require judicial approval before going into effect.\(^{110}\) Moreover, a remediation agreement submitted for judicial approval must state what remediation the corporation is required to make to victims, or alternatively, must contain a statement by the prosecutor showing why remediation is not appropriate.\(^{111}\) The controlling legislation (the Criminal Code) provides that remediation agreements are intended to denounce corporate wrongdoing, hold organizations accountable, “contribute to respect for the law,” “encourage voluntary disclosure of [ ] wrongdoing,” “provide reparations for harm done to victims or to the community,” and “reduce the negative consequences of the wrongdoing” for persons associated with a corporate offender who did no wrong, such as employees, customers, and pensioners.\(^{112}\)

Canada’s new remediation agreements will not necessarily pave the way for a more principled approach to this practice. The legislation provides a list of factors for prosecutors to take into account in deciding whether to enter into a remediation agreement,\(^{113}\) but it does not give any guidance on how the recipients of remediation should be determined, or how remediation should be distributed. The conceptual framework proposed by the legislation is also unclear. In remedial settlement distribution, a clear conceptual approach to notions of harm is essential to achieve any degree of cogency. It is also noteworthy that the term “victim” has not received any useful explanation in the legislation. “Victim” simply means those who have “suffered physical or emotional harm, property damage or economic loss,” including persons outside of Canada.\(^{114}\) One reading of the new Canadian legislation is that it is deficient in laying the basis for any kind of principled approach to remedial settlement distribution. Another view, however, is that the Canadian federal legislature recognized that concepts of harm and victimhood in the remedial settlement distribution context are particularly amorphous and that it is best to leave it to enforcement agencies to construe these notions on a case-by-case basis until the practice evolves. To date, no remediation

---

\(^{109}\) It should be noted that remediation agreements came into effect after entities within SNC-Lavalin Group Inc. came under scrutiny for engaging in sophisticated foreign corruption schemes. SNC-Lavalin thereafter engaged in lobbying efforts to secure the introduction of a DPA regime. See MARIO DION, OFF. OF THE CONFLICT OF INT. & ETHICS COMMISSIONER, TRUDEAU II REPORT 21 (2019).

\(^{110}\) Criminal Code, R.S.C. 1985, c C-46, § 715.37(6). There are numerous features unique to the Canadian regime, however. DPAs also require consent of the Attorney General and the prosecutor must be satisfied that there is a reasonable prospect of conviction with respect to the offence. *Id.* § 715.32.

\(^{111}\) *Id.* § 715.34(1)(g).

\(^{112}\) *Id.* § 715.31(a)–(l).

\(^{113}\) *Id.* § 715.32.

\(^{114}\) *Id.* § 715.3(1).
agreements have been entered into, and so it is unclear how these non-trial resolutions will be wielded in practice. Nonetheless, it is apparent that Canada has emerged as a country that recognizes the importance of remedial settlement distribution, and it is possible that Canada’s efforts on this front will continue to develop.

III. TERMINOLOGY AND CONCEPTS

Because remedial settlement distribution developed from the practices of enforcement agencies, it has never been supported by a comprehensively articulated policy rationale or any kind of theoretical substructure. This is one of the reasons that a review of cases involving remedial settlement distribution reveals a disjointed and incoherent approach to remediation. Accordingly, this Article considers several basic yet foundational concepts that can be used to support future research and practice in this field. These include, first, a conception of harm and victimhood, and second, an approach to conceptualizing the different kinds of remedial responses in non-trial resolutions.

A. Direct and Indirect Harm

Foreign bribery in the developing world is capable of harming citizens and governments.\footnote{It is not the purpose of this Article to provide a complete account of the manifestations of this harm nor to distinguish which forms of bribery ought to be thought of as harmful. Neither does this Article explore the ethicality of bribery, whether some bribes might in some cases achieve utilitarian benefit or the role of subjectivity in characterizing bribes. That has been done elsewhere. For a helpful overview of the literature detailing ways in which foreign bribery can cause harm to individuals, companies, governments, public institutions, and societies at large, see Dávid-Barrett, supra note 1. See also Jeffrey R. Boles, The Two Faces of Bribery: International Corruption Pathways Meet Conflicting Legislative Regimes, 35 Mich. J. Int’l L. 673, 678–80 (2014) for an overview of the way in which foreign bribery can cause harm, and David Kennedy, The International Anti-Corruption Campaign, 14 Conn. J. Int’l L. 455 (1999), for a discussion of the role of subjectivity in the ethicality of bribery.} The explanation of harm provided here is intended only to underlie working definitions for enforcement agencies considering whether to include remediation in a non-trial resolution. The balance of this Section describes how acts of foreign bribery can cause harm to citizens and governments and then provides a simple taxonomy to describe this harm.\footnote{This Article focuses only on the harm caused by bribes that fall within the purview of the foreign bribery offence. This excludes most forms of petty corruption and embezzlement, as well as most instances where the bribe-payer and recipient are of the same nationality and operate within the same country. For an overview of the harm caused by corrupt practices more generally, see Int’l Council on Hum. Rts. Pol’y, CORRUPTION AND HUMAN RIGHTS: MAKING THE CONNECTION (2009), https://perma.cc/H8TN-M5CK.}

Harm caused by foreign bribery can be classified as “direct” or “indirect,” and those who suffer harm can accordingly be classified as either “first-order” or...
“second-order” victims. Direct harm is that which an enforcement agency is satisfied would not have occurred but for the payment of a particular bribe or the existence of a particular bribery scheme. When agencies are satisfied there exists a sufficient link between foreign bribery and harm, harm can be classified as “direct.” It is this type of harm that enforcement agencies are typically addressing when an individual or an entity is made to pay compensation to a specific victim. “Indirect harm” describes the “trickle-down” effects of foreign bribery that manifest because economic conditions and institutional instability render acts of corruption especially harmful. Typically, this is the harm suffered by citizens en masse and societies at large. It is this harm that enforcement agencies are attempting to address when non-trial resolutions provide for charitable donations, the purchase of assets for public benefit, or the injection of capital into public infrastructure. The recognition of each type of harm for purposes of remedial settlement distribution is naturally contingent on prosecutorial discretion. In the absence of a third-party decision-maker or court to assess harm and designate victims, the prosecutor alone makes these determinations. Under these

---

117 In the absence of legislative guidance on or judicial analysis of victimhood, this Article has opted for broad definitions of “victimhood” that enforcement agencies will be able to adapt to particular cases. This approach has found favor in practice, with senior personnel at the SFO stating that the broadest possible definition of “victimhood” is to be preferred. See Baker, supra note 8 (“All economic crime has victims . . . They may be the citizens of those states that fail to thrive due to rampant corruption and theft by their leaders . . . . The ideal outcome, where-ever it is possible, is for the money secured through asset recovery to be returned to victims, using that term in its widest sense.”).

118 A but-for standard has been employed in the context of discerning whether compensation was appropriate under a DPA negotiated by the SFO. See discussion supra Section II.C.

119 The literature on this topic has long recognized the concept of “social harm.” See Juanita Olaya, Kodjo Artisso & Anja Roth, Repairing Social Damage Out of Corruption Cases: Opportunities and Challenges As Illustrated in the Alcatel Case in Costa Rica (SSRN Working Paper, Dec. 6, 2010), https://perma.cc/S9RN-YEKP (“Social damage is the loss experienced in aspects and dimensions of the collective or the community relevant to the law.”).

120 Elizabeth Dávid-Barrett and Mihály Fazekas provide a recent overview discussing how both corruption generally and bribery within public procurement can cause far-reaching harm. Given that public procurement accounts for on average 29% of total general government expenditure in OECD countries (2013 data), and closer to 50% of public spending in developing countries, [corrupt] practices can cause serious damage to the economy and to public confidence in institutions. Favoritism in the allocation of public contracts can lead to higher prices, reduced value for money, the provision of low-quality or unsafe works, goods and services, and reduced competition. It is also likely to harm democracy since, by distributing resources according to particularistic ties, partisan favoritism disadvantages parties that lack connections and thus weakens political competition. Clientelism may even reverse the conventional relationship of democratic accountability, with politicians holding supporters to account for their behavior.

circumstances, prosecutors are not bound by the narrower concepts of causation that determine rights and duties under common law. Anti-corruption enforcement agencies do not undertake detailed causal analysis in deciding whether to pursue remediation, so causation in this context must be understood as clothed in prosecutorial discretion.

Direct harm is often facially apparent and easily subject to quantification. Cases involving direct harm are likely to involve compensation directed toward the government of the nation whose public official requested or received a bribe. For example, when a government agent or contractor pays a bribe using money that belonged to the government or to which the government was entitled, the government has sustained a loss equal to the value of that bribe. Alternatively, when a bribe is paid to a government employee to avoid paying import duties or some other form of taxation, then that government has been deprived of whatever capital it would have received had the tax been paid. A third example arises when a service provider wins a government contract by paying a bribe to a rogue government official, and it can be shown that the bribe-payer ultimately provided defective or overpriced services. In these examples, harm can be classified as direct because there is a self-evident connection between the bribe being paid and the government being deprived of capital that an enforcement agency may well recognize. As evinced below, cases involving direct harm present fewer practical hurdles.

---

121 It should be noted that while some countries do involve the judiciary in the approval of DPAs, those countries’ courts do not afford sustained analysis to questions of remediation. For example, see supra Section II.C for a discussion of Standard Bank, in which Lord Leveson found that harm would not have occurred “but for” the relevant acts of bribery but did not undertake the kind of causal analysis typical of decisions addressing whether a particular act caused damage or loss and rights under the common law.

122 While corporations might also suffer direct harm, they are not the focus of this Article because non-trial resolutions have not been used to ameliorate harm to corporations. The fact that the corporate competitors of bribe-paying companies might suffer harm through FCPA violations is a longstanding observation. The original Senate Bill of the FCPA provided a right of action to harmed competitors. International Contributions, Payments, and Gifts Disclosure Act, S. 3379, 94th Cong. § 10, 122 Cong. Rec. 12,607 (1976), https://perma.cc/Z8MF-25AN (“Any person who can establish actual damage to his business resulting from illegal . . . contributions, payments, or gifts, made by a competitor and who has not made such illegal payments himself in a relevant time period, may maintain a cause of action against that competitor.”).

123 See supra Section II.C (discussing Standard Bank).

124 See, e.g., United States v. Kay, 359 F.3d 738 (5th Cir. 2004).

125 See supra Section II.C (discussing Sarladi).

126 Although non-trial resolutions have not yet been used to award compensation to individual citizens for direct harm, it is possible that citizens might also suffer direct harm if an enforcement agency is satisfied that a specific act of foreign bribery caused them harm that was not “trickle-down.”
It is more difficult to articulate the nature and extent of indirect harm sustained by citizens, governments, and societies at large. This is because indirect harm is diffuse. It is typically expressed in terms of its intangible impact, such as the inefficiencies that result from public spending being distorted by bribes, and the more abstract harm that results as a byproduct of diminished institutions. Put another way, citizens suffer and the public interest is compromised when public spending decisions are influenced by the interests of bribe-payers and bribe-taking public officials. For example, individual citizens of developing nations might suffer indirect harm when an act of foreign bribery causes a misdirection or depletion of public resources that otherwise could have been made available for areas of public outlay that contribute to human development (healthcare, education, social welfare, etc.). This misallocation of public resources might occur when bribe-paying service providers ingratiate themselves to public officials over time, monopolizing a particular area of government spending and overpricing their services. Citizens might also sustain indirect harm because inefficient redirection of resources results in inferior public services. This occurs when an entity that bids for a public infrastructure contract pays a bribe, is successful in its bid, and then provides services inferior to those that would have been provided by an unsuccessful bidder that did not pay a

---

127 For an overview of the fraught nature of attempting to calculate the harm caused by corruption on a macro scale, see Matthew Stephenson, It’s Time to Abandon the “$2.6 Trillion/5% of Global GDP” Corruption-Cost Estimate, THE GLOBAL ANTICORRUPTION BLOG (Jan. 5, 2016), https://perma.cc/6HUV-QEPZ.

128 In his opening statement in Liu v. Sec. & Exch. Comm’n before the U.S. Supreme Court earlier this year, Deputy Solicitor General Malcolm Stewart stated that there is “no obvious universe of individual victims from an FCPA violation.” Transcript of Oral Argument at 35, Liu v. Sec. & Exch. Comm’n, 140 S. Ct. 1936 (2020) (No. 18-1501). While Stewart was speaking in the context of investors, his comments are arguably relevant to FCPA violations generally.

129 See generally SUSAN ROSE-Ackerman, CORRUPTION AND GOVERNMENT: CAUSES, CONSEQUENCES, AND REFORM (1999); Vito Tanzi & Hamid Davoodi, Corruption, Public Investment and Growth 1–26 (IMF, Working Paper No. 97/139, 1997). See also Dávid-Barrett, supra note 1, for a general theory of how bribery’s distortion of public decision-making causes harm. The OECD reported in 2014 that the purpose of fifty-seven percent of all instances of foreign bribery involved the public procurement process. OECD, PREVENTING CORRUPTION IN PUBLIC PROCUREMENT (2016), https://perma.cc/FM6E-M5FQ.

130 See Dávid-Barrett, supra note 1, at 128 n.13. See OECD, CONSEQUENCES OF CORRUPTION AT THE SECTOR LEVEL AND IMPLICATIONS FOR ECONOMIC GROWTH AND DEVELOPMENT (2015), https://perma.cc/4545-473U (providing a discussion of how corruption in the procurement process can lead to misallocated resources, higher expenses, and lower quality goods and services).

bribe.\textsuperscript{132} Finally, citizens can also sustain indirect harm from worsened economic outcomes associated with pervasive levels of corruption. This harm can be expressed in terms of its trickle-down effects on society at large. These trickle-down effects manifest by diminishing foreign investment,\textsuperscript{133} undermining the legitimacy of government institutions,\textsuperscript{134} contributing to fiscal deficits, and increasing income inequality.\textsuperscript{135}

Populations might also sustain harm in ways unrelated to inefficiencies that result from the redirection of public spending. This is evidenced by the Och-Ziff Capital Management Group bribery scandal.\textsuperscript{136} Between 2008 and 2012, American investment and hedge fund manager Och-Ziff Capital Management and related entities paid bribes to the Congolese government to procure a mining license.\textsuperscript{137} Before the license could be awarded, the Congolese government initially had to strip Canadian mining company First Quantum Minerals Ltd. of the license it held. The abrupt closure of the mine had a devastating economic impact on local communities that were suddenly deprived of income.\textsuperscript{138} Before the mine’s closure, the World Bank had invested in First Quantum’s mining activity through the International Finance Corporation.\textsuperscript{139} In return for the World Bank’s investment, First Quantum was obligated to provide social and environmental benefits to local communities, including the delivery of clean water, the provision of healthcare and education, and the alleviation of air pollution.\textsuperscript{140} After First Quantum was stripped of its mining license, World Bank involvement and all related infrastructure and social programs ceased.\textsuperscript{141} Tens of thousands of citizens in local communities were affected.\textsuperscript{142} Och-Ziff Capital Management and one subsidiary

\textsuperscript{132} See Wim Wensink & Jan Maarten de Vet, Identifying and Reducing Corruption in Public Procurement in the EU (2013), for a discussion of case studies showing how corruption leads to suboptimal decision-making and cost overruns in the procurement process.


\textsuperscript{136} See generally RAID, supra note 37.


\textsuperscript{138} RAID, supra note 37, at 11–12.


\textsuperscript{140} RAID, supra note 37, at 11.

\textsuperscript{141} Id. at 7–8.

\textsuperscript{142} See id. at 11–38, for a description of how the closure of the mine and withdrawal of the World Bank aggravated conditions that led to the impoverishment and in some instances death of citizens.
were found to have breached the FCPA and entered into a DPA with the DOJ, but no remediation was provided to Congolese citizens. Shareholders of one of Och-Ziff’s related entities, however, were able to claim victimhood status and seek remediation under the MVRA. The SFO is still investigating this matter, and Congolese citizens have come forward and identified themselves to the SFO as victims. The Och-Ziff Capital Management bribe scheme is just one example of how foreign bribery can cause widespread harm.

It is not only citizens who sustain indirect harm from the distortion of public spending. Governments might, for example, sustain either direct or indirect harm when their resource allocation has been diverted or influenced by a foreign entity paying a bribe to a rogue public official through the procurement process. In this situation, the government has been deprived of a degree of autonomy—which naturally varies based on the extent to which corruption pervades the national government and economy. The government might also sustain harm in the form of inefficient resource expenditure or reputational loss. This harm can, of course, be contrasted with the harm that flows from entities paying bribes to influence lawmaking and policy formulation. The phrase “state capture” describes the practice of bribes being paid to influence the lawmaking process. A government may sustain harm if the bribe-payer redirects an entire government agenda that is ultimately against the national interest.

Admittedly, this description of the harm suffered by governments is simplistic, and the distinction between direct and indirect harm resulting from resource misallocation is difficult to draw without sustained and detailed economic analysis. Indeed, several factors might complicate the nature of the harm suffered by governments. For instance, it is difficult to argue that a government is a victim when the bribe-seeking public official is not rogue and the corrupt behavior is a common or expected practice in a particular

147 For a discussion of the role of state capture in creating social harm in the Philippines, see Hannah Isabella Chan et al., Civil Action Against Corruption: Empowering the Filipino People in a Captured State Situation, in ISSUES IN COMBATTING TRANSNATIONAL CORRUPTION, supra note 87. See also Dávid-Barrett & Fazekas, supra note 120, for a discussion of how corruption can derail democratic public spending by hijacking policy at its formulation, implementation, or monitoring.
administration. Moreover, the extent to which it is appropriate to distinguish between the harm sustained by citizens and the harm sustained by their governments can also be disputed. It is entirely arguable, if not facially apparent, that citizens also sustain harm when their government’s agenda has been influenced or completely commandeered by bribe-payers.

The harm described here as indirect is not only diffuse, but also vague and involves conceptions of loss that are difficult, if not impossible, to quantify. This is because conceptions of societal harm that rely on a trickle-down analysis cannot be applied to isolate the harm caused by a specific bribe or bribery scheme. Put another way, the above descriptions of indirect harm conflate the harm caused by a given act of foreign bribery with the harm caused by other forms of corruption and institutional instability generally. Foreign bribery typically causes harm in developing nations when it contributes to a broader culture of corruption permeating the state. This is why a singular bribery scheme within a developing country may cause more harm than an identical bribery scheme in a developed country. It is therefore difficult in developing economies to delineate the harm that has been caused by one form of corruption from the harm that has been caused by another. This may be problematic if one takes the view that enforcement agencies purporting to pursue remediation against individual bribe-offerors should seek only to remediate the exact harm caused by particular acts of bribery.

Notwithstanding the difficulty (or impossibility) of separating harm caused by specific acts of foreign bribery from harm caused by other more general and pervasive forms of corruption, the remedial practices of enforcement agencies reveal a willingness to conflate these different types of harm and to pursue remediation for social harm generally. The prevailing thinking, as evinced by the cases discussed above, is that it is better to pursue remediation for the benefit of those harmed by bribery than it is to forgo remediation because it is sometimes impossible to calculate the harm caused by a specific bribe. This approach can be easily praised or decried. On one hand, it necessarily attributes a degree of social or economic harm to singular acts of bribery when that harm was actually caused by the amalgamation of many past acts of corruption. In this sense, the

148 See, e.g., Republic of Iraq v. ABB AG, 768 F.3d 145, 163 (2d Cir. 2014) (refusing to order restitution in favour of the Iraqi government for harm that flowed from the corrupt practices of that government).


150 See supra Section II.C.1 (discussing SFO enforcement actions).
remedial response can be criticized as being disproportionate to the harm. On the other hand, it might be argued that bribe-payers operating within the developing world know or should know that their corrupt conduct compounds with prior corrupt acts and institutional instability and is all the more harmful as a result. In support of this view, there are simple and strong policy reasons for extracting capital from corrupt actors for remedial purposes and providing remediation to citizens harmed by bribery in the developing world. This reasoning has found favor with enforcement agencies in the exercise of their discretion.\footnote{See discussion of U.S. and U.K. enforcement actions, supra Section II.C.}

Another complication of accepting that remedial settlement distribution is an appropriate response to indirect harm is that questions naturally arise as to when harm will be too remote to be considered to have been indirectly caused by a particular act of foreign bribery. An extremely expansive reading would posit that every act of foreign bribery undermines the rule of law, diminishes the institutional integrity of government, and subverts public policy to some degree.\footnote{See, e.g., Dávid-Barrett, supra note 1, at 132 (“Every bribe paid to influence a public official to divert from following the rules associated with her office demonstrates that the rules are not consistently applied in line with the law. This is true whether the bribe is paid by Oxfam or by an arms dealer, whether the bribe is big or small, and whether the bribe secures a place at the front of the customs queue or a million-dollar contract. The fact that the public official decides to violate the rules in order to secure some private advantage is unjust.”).}

On this view, every single act should sound in some form of remediation, and every state in which bribery has occurred and every citizen of that state would be a victim. This Article does not support this kind of expansive approach to remediation and acknowledges that the term “indirect harm” implies a degree of ambiguity and requires enforcement agencies to choose what harm is too remote and what harm is not. For as long as remedial settlement distribution is practiced pursuant to prosecutorial discretion rather than statutory mandate, enforcement agencies will make that choice for themselves on the basis of the factual matrix at hand.

B. Categories of Remediation

The bulk of commentary in this field employs the term “compensation” to describe all transfers of capital to victims.\footnote{See supra Section II.C.1 (discussing the U.K. government’s Compensation Principles); see also Spalding, supra note 11; Delphia Lim et al., Access to Remedies for Transnational Public Bribery: A Governance Gap, in ISSUES IN COMBATTING TRANSNATIONAL CORRUPTION, supra note 87, at 3–23.} This term is unhelpful. It conflates the provision of remediation in instances where distinct entities have suffered direct harm with the provision of remediation in instances where large groups of people have suffered indirect harm. Practice shows that the appropriate remedy for direct harm in most cases is the advancement of capital equivalent to the loss occasioned by the bribery. To describe this form of remediation, this Article uses

\footnote{See supra Section II.C.1 (discussing the U.K. government’s Compensation Principles); see also Spalding, supra note 11; Delphia Lim et al., Access to Remedies for Transnational Public Bribery: A Governance Gap, in ISSUES IN COMBATTING TRANSNATIONAL CORRUPTION, supra note 87, at 3–23.}
the term “compensation.” The appropriate remedy for indirect harm in most cases is the advancement of capital through charitable endeavors, infrastructure investment, the purchase of public assets, or some other publicly accessible means. To describe this type of remediation, this Article uses the term “reparations.” A third kind of remediation that is not typically pursued through non-trial resolutions, but that could be adopted with relative ease, involves the advancement of capital equivalent to the profit made by the bribe-payer. To describe this form of remediation, this Article uses the term “restitution.” “Remediation” is an umbrella term that encapsulates compensation, reparations, and restitution. The balance of this Section explains the nature of each of these forms of remediation. Because restitution has not been employed in practice, the focus is on compensation and reparations.

Compensation in this context is informed by the same concept of corrective justice that underlies compensatory damages in tort law. The defining feature of such damages is duality. That is, damages pair the wrongdoer with the victim. The contours of the remedy are determined by the relationship between the wrongdoer and the victim, to the exclusion of all other considerations. In this sense, the remedy aims to restore the normative equilibrium that existed between wrongdoer and victim before the wrong took place. Compensation, therefore, should theoretically aim to correct the harm occasioned by foreign bribery by placing victims in their pre-wrong state, thus restoring whatever degree of normative equality existed between the relevant parties before the wrong occurred. Compensation in this context is a loss-based form of remediation. It is not to be calculated with regard to the value of the bribe, the benefit derived by the bribe-offeror, or the need to punish the bribe-offeror. Of course, this notion of compensation is a creature of private law theory, and its applicability here is limited for numerous reasons. In negotiating remediation terms in foreign bribery settlements, it is unlikely that it will always be feasible to place the victims in the position they were in before the wrong was committed, whether due to the insolvency of the bribe-offeror, the fact that the exact amount of loss sustained cannot be ascertained or even quantified, or the dynamics at play in the negotiation process. Prosecutorial discretion remains the only arbiter for determining the value, form, and recipient of compensation in foreign bribery enforcement

154 Some texts employ the word “restitution” instead of “compensation.” This paper does not adopt this approach. Remedies that operate pursuant to a compensatory rationale are loss-focused, meaning that they respond to the loss of a plaintiff or a victim and attempt to undo that loss. See Jeff Berryman, The Compensation Principle in Private Law, 42 Loy. L.A. L. Rev. 91, 103 (2008). Restitution, on the other hand, also takes into account the gain of the wrongdoer and describes those remedies whereby the wrongdoer is compelled to “give up” their ill-gotten gain.


156 Id. at 70.

157 Id.
actions. Nevertheless, this notion of compensation has value insofar as it provides a benchmark: a loss-based concept to guide compensation in remedial settlement distribution.\footnote{Id. at 60–61.}

The distribution of reparations to the victims of indirect harm is best understood in light of traditional notions of distributive justice. Distributive justice provides that an appropriate response to a particular wrong with widespread consequences is to divide the remedy among the populace that was wronged.\footnote{Id.} In situations where citizens have sustained harm due to acts of bribery, the remedy does not take the form of a direct transfer of wealth from the wrongdoer to the wronged but rather a distribution of wealth through publicly available means (a practice explored in greater detail below).\footnote{See supra notes 75–105 and accompanying text.} Finally, restitution in this context refers to a gain-based method for determining remediation. It is measured with reference to the bribe-offeror’s ill-gotten gains. The adaptability of restitution to remedial settlement distribution is entirely possible, as enforcement agencies often extract fines and penalties on a disgorgement basis—meaning that the fine or penalty is equivalent to the wrongdoer’s ill-gotten gains.

The terms “compensation,” “restitution,” and “reparations”—while considered appropriate by the author—need not necessarily attach to the concepts articulated above. What is important to appreciate is that remediation can be conceived as being either loss-based, gain-based, or responsive to trickle-down societal harm. This demarcates the ways in which a non-trial resolution can remediate the harm to victims of foreign bribery: first, by advancing compensation to identifiable victims; second, by using public means to disburse reparations to victims who have suffered indirect harm;\footnote{The pairing of direct harm with compensation and indirect harm with reparations may not always be practical. An enforcement agency may find it necessary to advocate for a term in a non-trial resolution that provides victims with reparations, notwithstanding the fact that those victims suffered direct harm. This would likely be the case when the provision of compensation would be unfeasible from an administrative standpoint due to the sheer number of victims harmed. Reparations may be the appropriate or the only practical vehicle for remediation simply because an enforcement agency cannot calculate the loss of each specific victim or organize a transparent and efficient way of providing compensation. In this instance, the enforcement agency might decide that it is better to provide reparations—such as infrastructure investment, donation to a charity to which victims have access, etc.. The concepts here provide the foundation for a cogent and cohesive approach to remedial settlement distribution, but in the absence of being entrenched in statute, they must remain malleable so that enforcement agencies will be willing to apply them. See supra notes 75–105 and accompanying text.} and third, by requiring an entity that has entered into a non-trial resolution to give up ill-gotten gains to victims.
IV. A FRAMEWORK FOR COMPENSATION

The balance of this Article outlines a framework for the provision of compensation and reparations in foreign bribery non-trial resolutions. The relevant factors that enforcement agencies should take into account in considering whether to seek compensation in a non-trial resolution are as follows (direct harm is not addressed below, as it has already been discussed):

- Whether there is an identifiable victim;
- Whether that victim has suffered direct harm;
- The extent to which that harm is ascertainable;
- Whether there is a risk of repeat corruption and, if so, whether that risk can be managed; and
- Whether compensation is appropriate in the circumstances.

It should be noted that these factors could also be applied to support a framework for restitution. Indeed, it would be redundant to present separate frames for restitution and compensation because the only distinction would be in the third factor: for restitution, “the extent to which illicit gains are ascertainable” would be substituted for “the extent to which that harm is ascertainable.”

A. The Identifiability of the Victim

A non-trial resolution can provide for compensation only when a specific victim can be identified.162 This is the necessary corollary of compensation being measured against the loss of discrete individuals or entities. Indeed, the case law that has emerged from DPA approvals in the U.K. confirms that whether a victim can be identified is a necessary precondition to the provision of compensation. Two cases mentioned above serve as illustrative examples: Standard Bank and Sarclad.

In Standard Bank, Standard Bank’s Tanzanian subsidiaries were successful in their bid for a contract that permitted them to act as underwriters on behalf of the Tanzanian government for a sovereign note placement. After the subsidiaries won that contract, one of them entered into a consultancy agreement with a shell company directed by public officials. The agreement stipulated that the Tanzanian subsidiary would pay one percent of all funds raised as part of the note placement to the consulting company.163 This one percent commission, which under the circumstances was revealed to be a bribe, diverted money that would otherwise

---
162 The international asset recovery community has long recognized the importance (and difficulty) of identifying the victims of corruption as a precondition to providing remedies. See U.N. Secretariat, Good Practices in Identifying the Victims of Corruption and Parameters for their Compensation: Note by the Secretariat for the Open-ended Intergovernmental Working Group on Asset Recovery (Aug. 4, 2016), https://perma.cc/9SN6-XMBD.
have been paid to the Tanzanian government to the consulting company.\(^{164}\) The note placement raised US$600 million, so US$6 million was paid to the consulting company.\(^{165}\) Thus, the bribe resulted in the Tanzanian government being deprived of US$6 million. There was no room for doubt about the identity of the victim. The DPA approval judgment unsurprisingly reflects minimal deliberation on this front. The presiding judge and president of the Queen’s Bench Division, Lord Justice Leveson, simply found that the Tanzanian government would not have incurred the loss “but for” the bribes paid by the Tanzanian subsidiary.\(^{166}\) This was enough to justify compensation in the circumstances.\(^{167}\) Indeed, \textit{Standard Bank} might be termed an “easy case” with regard to discerning the identity of a victim to be paid compensation.

\textit{Sarclad} falls at the opposite end of the spectrum. As noted above, the DPA in that case pertained to a web of sophisticated bribe schemes spanning multiple jurisdictions over a long time. Lord Leveson, who had again been tasked with presiding over the SFO’s DPA approval application, provided the following analysis of why compensation was not appropriate:

\begin{quote}
[Seventeen] of the [twenty-eight] implicated contracts were with entities based in a country in Asia with which there is neither a request for mutual legal assistance nor an established mechanism or practice in place for payments of compensation orders to the authorities. Other bribes . . . involved agents based in or working in relation to other countries in Asia and elsewhere in respect of which the same difficulties arise. Further, the amounts of the bribe payment are not always confirmed in the evidence and neither is any rise in the contract price to accommodate it (which would generate the loss). Finally, the SFO is not able to demonstrate whether and, if so, in what sum, the various Sarclad agents actually paid bribes to named or unknown individuals. Taken together, these factors amount to it not being possible to positively identify any entities as victims who may be compensated.\(^{168}\)
\end{quote}

This excerpt illuminates more than the obvious fact that unidentifiable victims cannot be compensated. The reasoning highlights four considerations that indicate whether a victim can be identified: (i) whether there had been a request for mutual legal assistance;\(^{169}\) (ii) whether the victim(s) resided in a state with established mechanisms or practices for payments of remediation; (iii) whether the amounts of the bribes paid were confirmed in evidence and whether any other

\(^{164}\) Id. [10].

\(^{165}\) Id. [8].

\(^{166}\) Id. [40], [51].

\(^{167}\) Id. [41].


\(^{169}\) OECD, Convention on Combating Bribery of Foreign Public Officials in International Business Transactions, art. 9, ¶ 1, Nov. 21, 1997, https://perma.cc/U9K4-LWZA (providing that mutual legal assistance entails the provision of “prompt and effective legal assistance . . . for the purpose of criminal investigations and proceedings”).
evidence indicated loss; and (iv) whether it could be shown that bribes of a specific amount had been paid by particular agents to particular individuals. These considerations are not presented as creating a mandatory threshold for the payment of compensation, but as factors that might make identifying a victim more practicable on a given set of facts. These are worth discussing further.

The first two considerations are relatively straightforward. In certain cases, a victim might be identified only if the victim’s identity is put forward by a foreign enforcement authority. The existence of established mechanisms for payments of compensation orders would doubtlessly facilitate such payment. It should be noted, however, that the existence of such mechanisms bears on the practicality of delivering compensation rather than the antecedent issue of whether a victim can be identified.

The third consideration focuses on whether there is evidence to prove a victim’s loss, such as evidence showing the value of bribes or a decline in the value of services provided under an agreement tainted by foreign bribery. Naturally, evidence of a specific victim’s loss indicates the existence of that victim. However, Lord Leveson’s suggestion that the value of bribes or a decline in the value of services indicates a victim’s identity is misplaced, as neither of these pieces of evidence speaks to the identifiability of victims in a precise sense. These two indications are better taken into account when determining whether a loss is ascertainable. Furthermore, there is no reason why the value of the bribe should have any bearing on the amount given to a victim. Indeed, if the point of compensation is to respond to harm sustained by a victim, then the value of the bribe is irrelevant unless the bribe correlates to the loss in some way. Finally, Lord Leveson indicated that victims could not be identified because there was a sparsity of evidence showing whether specific bribes had been paid to specific foreign officials. Put another way, while the SFO’s investigation had revealed a broad and pervasive culture of bribery, Sarclad’s lack of internal records and documentation ultimately made it impossible to show that a particular entity had sustained direct harm and would be due compensation. This finding provides further support for the assertions made above that the payment of compensation for direct harm requires some kind of connection between an act of bribery and a particular victim’s loss.

At bottom, Lord Leveson’s comments indicate that when considering whether to award compensation, the first point of inquiry is whether a victim can be identified. In some instances, the identity of the victim will be apparent because

---

170 Lord Leveson of course uses the term “compensation,” but that term as used in Sarclad and Standard Bank is cognate with the term “remediation” as used in this article. Lord Leveson does not distinguish between remediating individuals for ascertainable loss and remediating the public at large.

171 Sarclad, EWHC (QB) U20150856, [53].
of the nature of the bribe scheme, as in *Standard Bank*. In more difficult cases, enforcement agencies might refer to any one of numerous case-dependent considerations in attempting to discern the identity of a victim.

**B. Whether Harm is Ascertainable**

Because compensation is a loss-based form of remediation, it is necessary to be able to ascertain some degree of loss before compensation can be included in a non-trial resolution. The case law supports this view. *Standard Bank* and *Sarclad* again reflect opposite extremes. The extent of loss in *Standard Bank* was facially evident. The Tanzanian government had been deprived of exactly the amount given to Standard Bank’s bribe-paying subsidiary.172 In *Sarclad*, Lord Leveson found that the degree of factual complexity meant that loss could not be ascertained at all.173

There is little doubt that loss should be established before an award of compensation is made. A more contentious issue is the extent to which loss needs to be determined. This Article rejects the view that it will, in all cases, be necessary to ascertain loss with absolute certainty. Enforcement authorities, at their discretion, might be content to employ means of measuring approximate loss. For example, an enforcement agency might decide that compensation should reflect the devaluation in goods or services provided under a contract tainted by bribery. In cases involving complex, prolonged bribery schemes, enforcement agencies may elect to compensate whatever loss can be ascertained or accurately estimated. Such is an advantage of the discretionary nature of the process of negotiating non-trial resolutions.

The proposition that loss need not be calculated with absolute certainty finds support in relevant areas of U.S. and U.K. corporate criminal law. In *R v. Alstom Power*, a case heard before the U.K.’s Southwark Crown Court, French energy company Alstom Power pleaded guilty to several offenses in relation to a bribe paid to Lithuanian public officials in the energy sector.174 The Lithuanian government formally requested compensation so that it could reimburse the accounts of Lithuanian banks defrauded as part of Alstom’s corrupt scheme. The court ordered that £10,963,000 be paid in compensation to the Lithuanian government, notwithstanding the fact that all parties agreed that loss could not be calculated with absolute certainty.175 Similarly, in cases under the MVRA in which victims seek remediation for the harm caused by financial crime, U.S. federal courts need only make a “reasonable estimate of the [actual] loss” given the

---

172 *Standard Bank*, EWHC (QB) U20150854, [9]–[15].
173 *Sarclad*, EWHC (QB) U20150856, [53].
175 *Id.* at 6.
information available. Enforcement agencies are of course not obligated to borrow from the approaches of courts in corporate criminal cases, but the willingness of courts to award compensation without being able to ascertain harm with absolute certainty indicates that enforcement agencies could function with a similar modus operandi when negotiating non-trial resolutions.

Ultimately, the extent to which loss can be ascertained may bear on the appropriateness of awarding compensation, but it should not mandate that compensation be paid or not paid—unless, of course, the relevant enforcement agency is unsure of whether any loss was sustained. It is also worth noting that some enforcement agencies might elect to bypass ascertaining a victim’s harm and order compensation pursuant to some other calculus. U.S. lawmakers and academics have considered the payment of remediation calculated with regard to either the value of the bribe, the bribe-payer’s gain, or the value of business lost by a victim. However, this Article takes the view that remediation should respond to a victim’s loss so as to function as a kind of remedy. Other calculations might be used as alternatives where necessary, depending on the practices of a particular agency.

C. The Risk of Repeat Corruption

A standard objection to remedial settlement distribution is that the practice is mired in the inherent risk that distributed capital might be repurposed for corrupt ends. This Article suggests that neither compensation nor reparations should be provided if an enforcement agency is satisfied that there is too great a risk that remediation monies might be repurposed in this way. The risk of repeat corruption should hold substantial weight in an enforcement agency’s decision to pursue remedial settlement distribution. Anti-corruption enforcement agencies generally vie to punish corrupt practices and deter would-be malefactors from engaging in future corrupt acts. It is nonsensical to think that they might exercise their discretion to empower corrupt governments or expend time and resources

176 United States v. Gallant, 537 F.3d 1202, 1252 (10th Cir. 2008).
177 For example, see the discussion of the Foreign Business Bribery Prohibition Act, H.R. 3531, 112th Cong. (2011), in Spalding, supra note 11, at 1419.
178 This view was articulated to the author by several current and former senior prosecutors with relevant experience, and it can also be evidenced by scrutinizing the non-trial resolutions the DOJ has entered in which the DOJ repatriated the proceeds of corruption to the citizens of developing nations. See also Matthews, supra note 54 (quoting the former head of the DOJ’s Kleptocracy Asset Recovery Initiative as saying, “We have to be flexible and nimble in finding ways to responsibly repatriate . . . We don’t want to give funds back to a potential corrupt government, or to the people who stole the money in the first place.”). Indeed, the need to manage the risk of corruption, through audits and other means, in the dispersal of capital—in ways that are akin to reparations as described in this Article—in developing nations is well noted. See, e.g., NORAD, JOINT EVALUATION OF SUPPORT TO ANTI-CORRUPTION EFFORTS: TANZANIAN COUNTRY REPORT (2011).
to assist those who cannot reasonably be considered victims. Indeed, MVRA precedent supports this view. Foreign governments seeking compensation under the MVRA in the wake of foreign bribery enforcement actions have been rejected on the basis that corruption was effectively rampant throughout the entire administration.\textsuperscript{179} U.K. courts have shown a similar recalcitrance to award compensation where there is a risk that it might fall into the wrong hands.\textsuperscript{180} However, enforcement agencies should, if possible, go one step further than identifying risk at face value and determine whether this risk can be managed or reduced. Practice suggests that if the risk of repeat corruption can be reduced or managed, remediation will still be appropriate.\textsuperscript{181}

In some cases, enforcement agencies will not have to explore methods to curb this risk because it will be minimal at the outset. For example, if an enforcement agency is deciding whether to include compensation in a non-trial resolution for the benefit of a foreign government and the bribe-receiving public official in that case was not acting within a prevalent culture of public corruption, and was thus a rogue public official, then the risk of repeat corruption will likely be minimal.\textsuperscript{182} Assessing the risk of repeat corruption in cases where compensation is being considered will be relatively straightforward compared to cases involving reparations. In the former, enforcement agencies need only assess the risk associated with a discrete and identifiable victim. The requisite due diligence is therefore limited in scope. In addition, the enforcement agency may be able to cooperate with the victim and come to terms regarding auditing and transparency about the way in which the compensation money is used. Where reparations are concerned, however, the task might be more difficult. The enforcement agency has to assess the risk that repeat corruption remediation

\textsuperscript{180} See Smith & Ouzman Ltd: First Corporate Convicted for Overseas Bribery to Pay £2.2m, CMS LAW NOW (Jan. 11, 2016), https://perma.cc/Y28F-UJRD.
\textsuperscript{181} In practice, techniques for managing the risk of repeat corruption are drawn from reparations cases and so they are addressed in greater detail below in Section IV.B. These techniques have been employed in two cases involving the DOJ. The James Giffen prosecution eventually led to the creation of the BOTA Foundation, a charitable body that operated within Kazakhstan, and a settlement agreement involving then Vice President of the Republic of Equatorial Guinea Teodoro Obiang Mangue. For an overview of the BOTA Foundation and the circumstances leading to its creation, see Michael Steen, Kazakh “Oil Bribe” Millions to Go to Poor Children, REUTERS (May 4, 2007), https://perma.cc/W9PL-6PQV; WORLD BANK, FINAL SUPERVISION REPORT OF THE BOTA FOUNDATION (2015), https://perma.cc/E2AH-FZQU. For an overview of the Nguema Obiang case, see Press Release, U.S. Dep’t of Just., Second Vice President of Equatorial Guinea Agrees to Relinquish More than $30 Million of Assets Purchased with Corruption Proceeds (Oct. 10, 2014), https://perma.cc/4P6C-NDEL.
\textsuperscript{182} This might be the case when a bribe scheme involves low-level government officials who lacked influence at higher tiers of government, such as customs officials. See United States v. Kay, 359 F.3d 738 (5th Cir. 2004).
monies are dispersed through publicly accessible means, which is likely to involve more intermediaries, more recipients, and an increased likelihood of repeat corruption.\footnote{Id. The terms of the non-trial resolutions in the James Giffen and Teodoro Obiang Nguema Mbasogo cases, see supra note 181, regarding the way in which reparation monies could be spent, reflect, first, that enforcement agencies take the risk of repeat corruption very seriously and, second, that enforcement agencies are capable of dealing with this risk in a variety of ways.} The risk of repeat corruption in reparations cases is addressed below.

D. The Appropriateness of Compensation

Finally, as remedial settlement distribution is a byproduct of prosecutorial discretion, and its pursuit is a strain on the limited competences and resources of enforcement agencies, it is realistic to expect that enforcement agencies might decline to pursue compensation at the outset when doing so would be highly impracticable or otherwise unnecessary. The most obvious example is when the victim is a wealthy state or entity that has not indicated any interest in receiving compensation or has refused to accept compensation.\footnote{A country may also deny remediation on the basis that accepting it would constitute an acknowledgment of corruption. The SFO has experienced this in attempting to provide compensation to the Ghanaian government. See IN’L DEV. COMM., FINANCIAL CRIME AND DEVELOPMENT, 2010–12, HC 847, at 38 (UK); see also LINKLATERS, PUNISHING CORPORATE OFFENDERS FOR BRIBERY AND CORRUPTION OFFENCES: THE SENTENCING GUIDELINE IN ACTION 4–5 (2016), https://perma.cc/R54F-AEEQ (noting that one of the reasons that compensation was inappropriate in the Smith & Ouzman case discussed above was that neither the Kenyan nor Mauritanian governments had requested compensation).} Another example is when the victim is not barred by jurisdictional hurdles from pursuing compensation through an established statutory framework. Lastly, an enforcement agency negotiating a non-trial resolution might find that remediation is more or less appropriate depending upon whether local enforcement agencies have been cooperative. While this consideration may not have anything to do with the questions of harm and victimhood that ought to guide remediation, it is possible agencies enforcing foreign bribery laws will take this into account when deciding how to exercise prosecutorial discretion; it is for that reason worth mentioning.

This final consideration is admittedly broad. If enforcement agencies were to simply decline to pursue remediation on the basis that it is inappropriate, without further justification, then any commitment to any kind of guiding framework would be pointless. However, this consideration has been included because, as remedial settlement distribution is a fledging practice, it is necessary to include one open-ended consideration to allow enforcement agencies to take the unique facts of each case into account. This is to say, until enforcement agencies have developed experience in assessing the appropriateness of compensation in foreign bribery non-trial resolutions, it would be unrealistic to
expect them to commit to a narrow list of considerations without any express justification to respond to factual oddities or political and social realities.

V. A FRAMEWORK FOR REPARATIONS

Articulating a framework that can be applied consistently and cogently to guide the provision of reparations is an inherently difficult task. As explained above, indirect harm is diffuse and often intangible. In many cases, it may be impossible to design or implement a remedial scheme that correlates to the harm caused by a particular bribe scheme.  

Other than recognizing the need to manage the risk of repeat corruption, enforcement agencies have not applied any palpable criteria to guide the provision of reparations. The available materials do not suggest that enforcement agencies have used any indicia to determine the value of reparations, how they are provided, or to whom they should be provided. Enforcement agencies have seemingly taken the view that, because bribery harms society generally, it is appropriate that corrupt entities pay reparations to society generally, whether that be through charitable donations, infrastructure investment, or some other means. The balance of this Article contemplates how that approach might be improved, exploring whether an open-ended multifactorial approach might provide a more principled and consistent framework for providing reparations. These factors—the first of which was discussed above—are the following:

- Whether an act of foreign bribery has caused indirect harm;
- Whether there is a nexus between the bribery and harm;
- Whether there is a risk of repeat corruption and, if so, whether that risk can be managed; and
- Whether reparations are appropriate in the circumstances.

A. Nexus Connecting Foreign Bribery and Harm

An important issue for enforcement agencies to consider when determining whether and how to include provision for reparations in a non-trial resolution is the extent to which there needs to be some kind of nexus between the indirect harm and the reparations scheme. For example, if bribes paid to secure a

---

185 Indeed, it has been argued that it would be impossible to formulate fair and effective legally enforceable rights that would grant redress for indirect or societal harm. See, e.g., Sakshi Aravind & Tanmay Dangi, Shades of Corruption: Thoughts on Why Compensation May Not be a Viable Solution, in ISSUES IN COMBATTING TRANSNATIONAL CORRUPTION, supra note 87, at 82–83.

186 This approach is derived from the DOJ’s former enforcement practices pertaining to U.S. environmental regulations. The DOJ routinely used to enter into “supplemental environmental projects” with companies that had been found to have caused tangible environmental harm. The purpose of these projects was to ensure “that any harm or threatened harm to victims or the
monopoly over a state’s public health sector lead to the provision of defective services in that sector, recognizing a nexus requirement could mean that reparations be put toward the benefit of hospitals, medical supplies, or the health sector generally. It is beyond the scope of this Article to mount a comprehensive argument regarding whether enforcement agencies should insist upon a nexus requirement as a precondition for including reparations in non-trial resolutions. Instead, this Article will outline reasons both for and against the adoption of a nexus requirement and then suggest a middle ground as the best approach.

Insistence upon a nexus between indirect harm and reparations ensures a degree of parity between harm and remediation. A nexus requirement assumes a pivotal function: ensuring that reparations operate as a type of remedy, it links the harm that is the subject of the non-trial resolution to the reparations. If enforcement agencies do not insist upon a nexus, it is arguable that reparations play a role that is more similar to foreign aid or charity.

However, there are legitimate reasons why an enforcement agency might choose to exercise its discretion to ignore an obvious nexus between foreign bribery and harm. One reason is that it might be counter to the interests of the enforcement agency’s government to provide reparations in a particular way. For example, if an act of foreign bribery caused harm to the military sector of a developing nation with unstable institutions, the decision to focus reparations elsewhere would be a logical one for that enforcement agency.

187 International criminal law provides a loose analogy. Judges of the International Criminal Court have ruled that reparations can be ordered only with respect to “harm suffered as a result of crimes if such crimes have been found to have taken place and the person standing trial for his or her participation in those crimes is found guilty.” Prosecutor v. Ruto, ICC-01/09-01/11-2027-Red-Corr, Public Redacted Version of Decision on Defence Applications for Judgments of Acquittal, ¶ 149 (Apr. 5, 2016). The analogy is loose because enforcement agencies are not bound by the same constraints as judges (constraints such as the existence of a conviction, for example). Nonetheless, the quoted reasoning shows that other bodies tasked with providing remediation to large groups for incalculable harm have insisted upon there being some connection between harm and victims. It is not enough that victims simply exist.

188 The BAE Systems settlement agreement negotiated by the SFO, discussed above, could be seen as an example. In that case, the bribe scheme involved the military sector, and reparations were paid to the education sector. See Joe Murphy, Let’s Pause Before We Dole Out Dollars, FCPA PROFESSOR (May 14, 2015), https://perma.cc/TS4A-HYF5 (warning that funneling capital into failing or unstable states could be perilous and warning of “unelected enforcement officials making policy
Remediation in Foreign Bribery Settlements

Hickey

act of foreign bribery caused detriment to a portion of society or an area of outlay that was not as in need of immediate capital injection as other portions of society or areas of outlay, then it again may make sense to ignore any nexus. One can of course argue that such considerations are unrelated to remediation and that remedial settlement distribution should focus on identifying and alleviating harm, but that position ignores the fact that enforcement agencies are highly unlikely to ignore political and practical realities.

The prevailing approach in anti-corruption cases is to disregard whether any nexus exists. To date, non-trial resolutions have simply mandated that the bribe-payer funnel capital into the country in which the bribes were paid. This may involve purchasing assets for the public benefit, supporting public infrastructure initiatives, donating money to charities, or creating a foundation. However, a point of stark contrast and a potential alternative method is found in the DOJ’s past approach to negotiating non-trial resolutions for breaches of environmental regulations. Professor Spalding noted that the DOJ often negotiated remediation schemes as part of non-trial resolutions involving corporate offenders that breached environmental regulations. These remediation schemes were “closely connected” to whatever environmental damage formed the basis of the non-trial resolution. Environmental breaches are, of course, an inapposite point of analogy because environmental harm is necessarily physical, so remediation can be tied to the geographic area in which the harm occurred. The

decisions about how to spend large pools of funds” before noting that determining victimhood “is not simply an administrative task; there will be important policy decisions to be made, including matters of foreign policy”).

The SFO followed this approach in the Chad Oil case discussed above. There, the SFO transferred the recovered money to DFID to “identify key projects to invest in that will benefit the poorest in Chad.” See SFO Recovers £4.4m from Corrupt Diplomats in ’Chad Oil’ Share Deal, SFO: NEWS RELEASE (Mar. 22, 2018), https://perma.cc/B2JV-RYPH. Furthermore, where the DOJ has opted to pursue infrastructure investment as part of the Kleptocracy Asset Recovery Initiative, the DOJ has opted to identify areas in which infrastructure investment is most needed.


Spalding, supra note 11, at 1417.
same cannot always be said for indirect harm flowing from corrupt practices. Nonetheless, this aspect of environmental regulation provides a useful and insightful point of comparison.

There is precedent for a form of remediation that does not require monies to be paid into the state in which bribery occurred. Instead, money can be paid to general anti-corruption initiatives.\textsuperscript{196} For example, as part of the settlement agreement entered into between the World Bank and Siemens AG, a German manufacturing giant, for bribery paid in connection with a World Bank development project in Russia, Siemens agreed to dedicate US$100 million over fifteen years to support anti-corruption work.\textsuperscript{197} Siemens subsequently launched the “Siemens Integrity Initiative,” inviting nongovernmental organizations and international organizations, associations, and universities to apply to Siemens for funding for projects that would “promote business integrity and fight corruption.”\textsuperscript{198} There is also support for this type of approach in the literature.\textsuperscript{199} Matthew Turk has suggested that money paid as part of FCPA disgorgement fines that cannot be transferred to the state in which the bribery occurred should instead be given to the OECD to bolster its anti-bribery efforts.\textsuperscript{200} Professor Spalding has similarly suggested that fines could be directed towards local organizations capable of investigating corruption.\textsuperscript{201} Spalding makes the point that the DOJ could determine the amount of money to dedicate to remediation in FCPA settlements in the same way it calculates fines.\textsuperscript{202} In the U.S., there is even support for this

\textsuperscript{196} International criminal law may provide a degree of guidance. International criminal law scholars and courts have dealt with the concept of “transformative reparations,” which are dedicated to changing the societal structures that perpetuated a particular breach of international criminal law. See Andrea Durbach & Louise Chappell, \textit{Leaving Behind the Age of Impunity: Victims of Gender Violence and the Promise of Reparations}, 16 INT’L FEMINIST J. POL. 543 (2014). However, it must be kept in mind that reparations of this nature can carry the potential to further neocolonial agendas and usurp democratic mandates in the state in which they are carried out. See Leila Ullrich, Faculty of Law, University of Oxford, Can Reparations Transform Societies? The Practice of “Transformative Justice” at the International Criminal Court, Presentation at Oxford Transitional Justice Research Seminar (Mar. 17, 2016), https://perma.cc/K9DE-MK6K.

\textsuperscript{197} Press Release, World Bank, Siemens to Pay $100 Million to Fight Fraud and Corruption as Part of World Bank Group Settlement (July 2, 2009), https://perma.cc/D8UU-MUWQ.

\textsuperscript{198} Press Release, World Bank, Siemens Launches US$100 Million Initiative for Anti-Corruption (Dec. 9, 2009), https://perma.cc/9J97-AKMH.

\textsuperscript{199} See, e.g., Lim et al., supra note 153, at 17 (“Rather than regarding FCPA fines as national revenue, an institutional mechanism could be established to channel the monies collected towards public purposes, pursuant to a broad conception of ‘remedy.’”).


\textsuperscript{201} Spalding, supra note 11, at 1417.

type of approach at a legislative level. Legislation has been introduced in both the House of Representatives and the Senate that would see a portion of FCPA settlements placed into an “anti-corruption fund.” That fund would then be used to assist developing nations in their respective fights against corruption. The proposed legislation would not remediate the harm to victims of corruption, but it is an example of the U.S. political system showing its readiness to repurpose the proceeds of FCPA settlements for relatively benign and constructive purposes. Ignoring a nexus requirement has one distinct advantage: enforcement agencies are not left to attempt to put a number on how much money should be paid in reparations or to identify a portion of society most impacted by an act of foreign bribery. Conversely, there is nothing to tie these payments to the nation in which bribery occurred, let alone to a specific act of bribery. Given the inherent difficulties intertwined with the risk of repeat corruption in orchestrating the payment of reparations, these types of payments could well find favor with enforcement agencies and the other government agencies who take it upon themselves to organize reparations schemes.

Ultimately, this Article expresses a preference for enforcement agencies to first determine whether a nexus exists and only then consider whether reparations can be provided to some other area of public outlay if either no nexus can be identified or there is some other compelling reason to not provide reparations to whichever segment of society was harmed. Alternatively, if it is impossible or impracticable, for whatever reason, to pay reparations into the state in which the bribery occurred, enforcement agencies might consider ordering payment to some kind of general anti-corruption agenda, as occurred in the Siemens case, as a last resort. This way, if adhering to a nexus requirement or paying reparations back into the state in which bribery occurred is infeasible, enforcement agencies can at least try to remedy indirect harm by seeking some other form of remediation. What is most important, however, is that some conception of nexus be articulated, including the circumstances in which that nexus requirement might be disregarded, so that at the very least the provision of reparations is made predictable, consistent, and coherent to the greatest extent possible.

203 The legislation introduced in the House would apportion five percent of all FCPA fines and penalties to anti-corruption initiatives, while the legislation introduced in the Senate would redirect US$5 million from every case in which fines and penalties exceed US$50 million. See H.R. 3843, supra note 15; S. 3026, 116 Cong. (2019).

204 For a more in-depth discussion, see Abigail Bellows, Revamping U.S. Anti-Corruption Assistance, Am. Interest (June 15, 2020), https://perma.cc/JMC2-UKEH.

205 Margaret Urban Walker has argued that transformative reparations that do not respond to any particular type of harm can substitute broader social goals for the need to remediate specific victims, which in turn diminishes the victim-centric function of reparations. See Margaret Urban Walker, Transformative Reparations? A Critical Look at a Current Trend in Thinking about Gender-Just Reparations, 10 INT’L.J. TRANSITIONAL JUST. 108, 110 (2016).
B. Risk of Repeat Corruption

Assessing the risk of repeat corruption may prove more complex in reparations cases, as the agency administering reparations has to satisfy itself that funds dispersed through a publicly accessible apparatus will not become repurposed for corrupt ends. The U.S. and U.K. exhibit different approaches to managing this risk. Non-trial resolutions negotiated by the DOJ tend to impose positive and negative obligations on the parties to the agreement, stipulating how reparations monies can and cannot be spent. The SFO, on the other hand, does not include these details in non-trial resolutions, leaving reparations schemes to be administered by other U.K. government agencies.

As part of its Kleptocracy Asset Recovery Initiative, the DOJ has implemented reparations schemes by donating money to charity, investing money in infrastructure projects, and creating a charitable organization. One instance in which the DOJ was involved in the creation of an independent foundation to deliver reparations is particularly instructive, as it sets out how those charged with managing the risk of repeat corruption might do so successfully. The settlement agreement, negotiated in 2007 following the seizure and forfeiture of more than US$115,000 over the course of investigating alleged breaches of the FCPA and other federal crimes, contained the following relevant stipulations: the foundation would be administered by a carefully selected multinational panel; money could be withheld from the foundation if the foundation’s independence became compromised; stringent procedural safeguards attached to the release of funds from the foundation; the foundation was subject to extensive auditing and reporting requirements; and the DOJ (and other organizations involved) retained the right to demand the return of funds from the organization in the event that its independence became compromised. The foundation supported three programs: conditional cash transfers to increase access to health, education, and

---

206 This problem has arisen in the context of India’s policy of corporate social responsibility. In India, the Companies Act of 2013 obligates high-earning and high-value companies to donate to charities. The process of donation creates a risk of bribery, exposing companies to liability under both Indian national bribery laws and foreign bribery laws such as the FCPA. Companies subject to India’s Companies Act of 2013 are therefore essentially tasked with managing the risk of repeat corruption. This is typically achieved through preemptive means, such as extensive due diligence. For more information, see John E. Turlais & David W. Simon, Where India’s Companies Act Meets the FCPA, FOLEY (May 18, 2017), https://perma.cc/XD4C-4CGZ.

207 For further background on this case, see Spalding, supra note 11, at 1418.


social welfare for impoverished families; conditional grants to local and international charities and nongovernmental organizations to promote children’s health and the well-being of orphans; and a tuition-assistance program.\textsuperscript{210} In this particular case, there were attempts to fraudulently extract money from the foundation, but they all failed when they were detected as a result of the auditing and transparency measures in place.\textsuperscript{211} And while substantial costs were incurred in ensuring accountability and transparency, reparations were delivered successfully.\textsuperscript{212} Similar techniques were also used by BAE Systems in administering reparations to the Tanzanian education system. BAE Systems insisted that reparations money would be paid out only in tranches, subject to the approval of a supervisory committee that would review the way the money was used.\textsuperscript{213}

In another Kleptocracy Asset Recovery Initiative case in 2014, the DOJ entered into a settlement agreement stipulating that the assets of a corrupt politician from a developing nation would be auctioned, and the resultant proceeds would be donated to a charity in that politician’s state.\textsuperscript{214} That settlement agreement contained the following features to minimize the risk of repeat corruption in dispersing reparations: the reparations monies could be spent only “for the benefit of the people” of the relevant state; the reparations monies could not be used to “make any payments or provide any form of consideration” to an enumerated list of third parties associated with the corrupt politician; and, lastly, the charitable body that received the reparations monies had to publish “an accounting of its expenditures of the funds and the results of those expenditures on an annual basis until such funds are fully expended.”\textsuperscript{215}

The DOJ has not wavered from its approach of using the terms of non-trial resolutions to solidify the contours of reparations schemes. In February 2020, the DOJ orchestrated the return of approximately US$308 million to Nigeria through the terms of a trilateral agreement with the governments of the Federal Republic of Nigeria and the Bailiwick of Jersey.\textsuperscript{216} That agreement included measures to ensure transparency and accountability, including administration of the funds and

\begin{footnotes}
\footnote{IREX, \textit{supra} note 193.}
\footnote{Id. at 54.}
\footnote{See Nicholas \textit{et al.}, \textit{supra} note 87, at 42.}
\footnote{Press Release, \textit{U.S. Dep’t of Just.}, \textit{supra} note 181.}
\end{footnotes}
projects by a designated agency, independent audits, and monitoring.\textsuperscript{217} The agreement also precluded the expenditure of funds to benefit alleged perpetrators of corruption.\textsuperscript{218}

In the U.K., the risk of repeat corruption has been managed primarily by the DFID and the FCO. These agencies have, in the past, been tasked with exploring opportunities to donate and invest monies exacted through non-trial resolutions. This approach relies less on the drafting of the non-trial resolution because an entity separate from the enforcement agency is able to dedicate its competences to managing the risk of repeat corruption.\textsuperscript{219} An advantage of this approach can be found in the\textit{Smith & Ouzman} case mentioned previously. In that case, the sentencing judge noted that the SFO had not been able to identify any safe recipient of remediation monies, so remediation of any kind was inappropriate.\textsuperscript{220} However, U.K. government agencies, including the DFID, were able to orchestrate a reparations scheme involving the purchase of ambulances. The advantage of this approach to reparations is that dedicated government agencies may be able to achieve outcomes not possible for an enforcement agency seeking to curb the risk of repeat corruption through the terms of a non-trial resolution. The disadvantage is obviously one of resources and funding.

An alternate approach to relying on the terms of the non-trial resolution or involving other government instrumentalities may be to outsource certain work to civil-society organizations. As noted in a report prepared by the University of Dar es Salaam in Tanzania, for the purposes of evaluating reparations administered in the BAE Systems case, “civil society is more likely to ensure that the funds obtained through compensation for corruption benefit the victims. A high number of civil society organizations work directly with people adversely affected by corruption such as impoverished communities.”\textsuperscript{221} It is therefore possible that enforcement agencies might consider whether there is a reliable civil-society organization that might play some role in transparently administering reparations.

It is beyond the scope of this Article to discern which approach is superior. Indeed, the appropriateness of either may be determined by external factors such as the availability of funds to dedicate to managing risk. What is important is that enforcement agencies consistently refer to the risk of repeat corruption as a precondition for pursuing reparations, explain their reasoning regarding why that risk could or could not be managed in a given case, and articulate their strategies

\textsuperscript{217} Id.
\textsuperscript{218} Id.
\textsuperscript{219} The DFID’s role in the BAE Systems case provides a useful example. See generally Nicholas et al., supra note 87.
\textsuperscript{220} LINKLATERS, supra note 184, at 5.
\textsuperscript{221} See Nicholas et al., supra note 87, at 49.
for managing repeat corruption. This will ensure that enforcement agencies continue to develop competences in this field.

C. Whether Reparations are Appropriate in the Circumstances

As noted above, the rubrics to underscore remediation presented in this Article are a first attempt at laying the groundwork for a fledging practice entrenched in prosecutorial discretion, so the inclusion of an open-ended consideration is necessary. By taking into account whether reparations are appropriate, enforcement agencies can refer to peculiar circumstances that could not be foreseen. As also noted above, the inherent vagueness of this consideration should not undermine the internal coherence of remedial settlement distribution as a practice if enforcement agencies clearly articulate why reparations are appropriate or inappropriate. Indeed, if enforcement agencies consistently refer to the same consideration as rendering remediation unnecessary, then that particular consideration may find its way into later articulations of official multifactorial approaches adopted by enforcement agencies. That is to say, remedial settlement distribution is a developing practice, and the inclusion of this final consideration will, if wielded correctly, aid in its development.

The most prominent example that may arise in the context of reparations is whether the victim state has a need for the reparations. If the victim of a foreign bribery scheme is a developed nation, it is highly likely that remediation would not be considered by the victim state to be necessary. This is because the trickle-down impact of a single foreign bribery scheme in the developed world is unlikely to cause any harm, even diffuse or indirect harm, to citizens. Further, the willingness of an international organization, foreign government, or a local civil society organization to assist in administering or monitoring a reparations scheme might also count toward the appropriateness of reparations.

VI. CONCLUSION

The remediation of the harm suffered by the victims of foreign bribery, and of corrupt practices more generally, is a conceptually dense and practically fraught field that has received little attention in the literature to date despite its growing prominence. This Article has limited its focus to foreign bribery non-trial resolutions and attempted to reduce some of the more complex issues into a set of working definitions and a framework that can be readily referred to by those

---

222 The importance of enforcement agencies creating a body of informal precedent to develop the competencies of agencies assessing the risk of repeat corruption cannot be overstated. Those involved with administering reparations to the Tanzanian school system following the BAE System cases discussed above complained of “a lack of precedents” to guide the process. See id. at 40 (citing INT’L DEV. COMM., supra note 184, at 10).
agencies involved with the remedial settlement distribution, in addition to academics and policy-makers operating in this growing area.
A Transnational Law of the Sea
Josh Martin*

Abstract

It is widely accepted that we are presently struggling to govern the vast expanse of the ocean effectively. This Article finally gets to the real cause of much of the failures of the law of the sea: Westphalian sovereignty. In particular, it evidences that certain features of our obstinate model of public international law—such as sovereign exclusivity, equality, and territoriality—can be linked with a large majority of the governance “gaps” in the global ocean context. It thereby exonerates the falsely accused Grotius’s mare liberum doctrine and flag state regulation, which both still continue to receive an unmerited level of condemnation. This Article also argues that worldwide searches for new integrated systems of ocean management are, in fact, a search for a new paradigm of governance, well-known among lawyers, but yet to be thoroughly analyzed in the law of the sea context, that of transnational law and governance. The study supports this conclusion by showing that two principal features of a transnational law of the sea—in the form of multi-stakeholder participation and multi-level governance—have already proven essential in ameliorating many of the routine weaknesses in our present international system of ocean governance.

* For their encouraging, thoughtful, and instructive feedback on previous drafts of this article, the author would like to extend sincere thanks to Professor Michael Schmitt, Professor Richard Barnes, Professor Hitoshi Nasu, Professor Mike Williams, Professor Helena Wray, Dr. Edwin Egede, and Professor Andrea Lista. He would further like to thank his master’s degree mentor of many moons ago, Professor Craig Barker, for originally piquing his interest in transnational law and for inspiring him to continue academic research.
# Table of Contents

I. Addressing the Legal System, Not Just the Legal Rules.................................................. 421

II. Introducing Transnational Law......................................................................................... 423
   A. Westphalianism ........................................................................................................... 423
   B. Transnational Law ..................................................................................................... 426
   C. Lex Maritima .............................................................................................................. 429

III. Weaknesses of the International Law of the Sea ............................................................ 432
   A. Sovereign Exclusivity ................................................................................................. 432
   B. Sovereign Equality ...................................................................................................... 442
   C. Sovereign Territoriality ............................................................................................... 446
   D. Westphalian Ocean Management: A System of Recurrent “Gaps” ......................... 449

IV. The Scapegoats of Our Failing Ocean Law ................................................................. 453
   A. Mare Liberum: Res Communis or Res Nullius? ....................................................... 453
   B. The Fault of “Flag” States or Flag “States”? ............................................................. 458

V. A Transnational Law of the Sea ...................................................................................... 462
   A. Understanding Integrated Ocean Management as a Demand for a
      Transnational Law of the Sea ................................................................................. 463
   B. A Transnational Law of the Sea: Multi-Stakeholder Inclusivity and Multi-Level Governance ......................................................................................... 466
      1. Multi-stakeholder inclusivity ................................................................................ 467
      2. Multi-level governance ......................................................................................... 469

VI. Conclusion: Looking Beyond the Horizontalism .......................................................... 473
I. ADDRESSING THE LEGAL SYSTEM; NOT JUST THE LEGAL RULES

The law of the sea is failing. Wherever one looks, whether it be ecosystem damage and biodiversity decimation, overfishing, seabed trawling, coral destruction, human rights abuses, human trafficking, piracy, smuggling, crime, wreck looting, noise pollution, land-based pollution, vessel-source pollution, health and safety failures, or major maritime disasters, one can witness recurring deficiencies in regulatory oversight. In consideration of the ocean’s ecological, social, economic, and cultural value, this increasing visibility of poor regulatory management has led to a recent proliferation of research dedicated to improving our protection of the seas. For example, in 2001, an expert international committee submitted a report that at the turn of the twentieth century, “the state of the world’s seas and oceans [was] deteriorating.” Furthermore, “most of the problems identified decades ago have not been resolved, and many are worsening.” Twenty years later, things still have not changed.

As James Harrison said in 2017:

[As the twentieth century progressed, the rapid industrialization of the oceans has meant that any lingering belief that the seas were “inexhaustible” gave way to a growing sense of crisis. This trend has continued to the extent that, today, there are warning signs that the oceans are at tipping point, owing to the impacts of pollution and other environmental stresses caused by anthropogenic activity.]

This has led to calls from every corner of the international community to transform our approach to ocean management away from the traditional “zonal” system of ocean management toward a more integrated and inclusive system. The development of such “Integrated Ocean Management” (IOM) processes, in various forms, can be increasingly witnessed on local, national, regional, and global

---


2 Group of Experts on the Scientific Aspects of Marine Environmental Protection, A Sea of Troubles 1 (Geoffrey Lean et al. eds., 2001).

3 Id.


5 See Section V.
scales. Additionally, many international, governmental, and non-governmental organizations are increasingly exploring the need for a new IOM paradigm, including the International Union for the Conservation of Nature, the U.N. Food and Agricultural Organization, the U.S. National Oceanic and Atmospheric Administration, the World Wildlife Fund, the U.N. Educational, Scientific and Cultural Organization (UNESCO), and the Conference of the Parties under the 1992 Convention on Biological Diversity. Research is also becoming more active in the search for the precise meaning of such “integrated” systems of ocean management, sporadically and fragmentedly spotting the need for ecosystems-based approaches, stakeholder participation, regional governance, or proposals for a more holistic system of regulation across ocean space.

Despite these well-intended efforts, humankind is still failing to get a grip on ocean governance. This Article demonstrates that the real root cause of this failure is our devotion to a Westphalian system of international law. It suggests that it is not—as has been previously suspected—the fault of Hugo Grotius’s 1609 thesis propounding a supposed “Freedom of the Seas” or the widespread use of flag state regulation per se. Instead, responsibility lies squarely with a dogmatic reliance

---

6 See generally Scott, supra note 4.
7 World Commission on Protected Areas, Incorporating Marine Protected Areas into Integrated Coastal and Ocean Management: Principles and Guidelines (Charles Ehler et al. eds., 2004).
16 Gabriela A. Oanta, Protection and Preservation of the Marine Environment as a Goal for Achieving Sustainable Development on the Rio+20 Agenda, 16 Int’l Comm. L. Rev. 214, 219 (2014); see also Section V.
on a horizontal and inter-state system of international law in the ocean context. As will be evidenced, the failure of the legal system to effectively steward the oceans can really be principally linked to three different traits of Westphalian sovereignty: *exclusivity, equality,* and *territoriality.* This Article also provides a detailed explanation and introduction to transnational law, including an account of the dormant *lex maritima,* in order to (1) highlight the richness of this field and (2) point to its lacking connection with the search for a new paradigm of ocean governance. Finally, it concludes with the view that we should reduce our emphasis on sovereign rights and duties and instead recognize the prevalent and longstanding calls to legal plurality, multi-scalarity, multi-stakeholderism, and post-nationalism.

Section II introduces the concept of transnational law as it stands in contrast to the Westphalian system of public international law, highlighting the worldwide movement to recognize and encourage law beyond the state, as well as the blurring of “public” and “private” sources and systems of law. Section III then critically examines the public international law system that has governed ocean management up to now, illustrating the ways in which Westphalianism can be linked to many failures of ocean stewardship. Section IV at last exonerates the scapegoats for failed ocean governance over the past decades—Grotius’s *Mare Liberum* and the notion of flag state regulation. In fact, this Article shows that Westphalianism itself caused both well-meaning doctrines to fail in practice. Section V connects the search for new “integrated” models of governance in the marine environment with the need for a model that dispenses with strict notions of national sovereignty and horizontal intergovernmental relations. This Section demonstrates that two features of an effective transnational system of ocean management are: (1) the expanded governance role for multiple and varied stakeholders and (2) the arrangement of diverse normative frameworks across multiple geographical scales. The Article concludes by calling for a greater recognition of the need for a transnational law of the sea.

II. INTRODUCING TRANSNATIONAL LAW

A. Westphalianism

Westphalian sovereignty refers to our familiar system of *inter-national* law resolved during thirty years of negotiations over the Münster and Osnabrück treaties, concluded between numerous European nations in 1648, and effectively ending the European wars of religion—a point in history known as the “Peace of Westphalia.” But most recognize that Westphalianism was not necessarily created

---

at this moment, but rather organically developed over many centuries around the
world.\textsuperscript{19} The overall effect was the formal legitimation of the nation-state as the
exclusive legal authority for all territorially internal matters and the understanding
that all nation-states are equal and unitary for all external matters.\textsuperscript{19}

A corollary to achieving temporary peace across Europe within this
emergent multi-state system was the widespread constitutionalization of polities
and quelling of competing internal claims to power.\textsuperscript{20} International governance
between the 17th and 20th centuries thus centered on both the strengthening of
zonal political boundaries and the demarcation of sovereignty within territories—in
which monolithic states possessed absolute independence to determine their
own internal laws free from outside influence and interference.\textsuperscript{21} The evermore
rigid political borders around states and the modeling of states as entirely unitary
and equal has, over the centuries, given birth to our modern system of inter-national
law—founded upon the principle of sovereign territorial independence, the
conclusion of positive international treaties, and the resolution of customary
norms between nations.\textsuperscript{22} Today, we remain firmly within this Westphalian system
of international law and thinking.\textsuperscript{23}

National sovereignty has, therefore, been at the heart of our understanding
of law and jurisprudence for several centuries. Most legal philosophers in this time
have explicated law’s basis as being positively determined by a higher authority,
whether through the canon of natural law, or as achieved functionally through the
formal use of power.\textsuperscript{24} Even as recently as the 1960s, both H.L.A. Hart’s famous
primary and secondary rules and Hans Kelsen’s Grundnorm theories espoused that
law’s ultimate source of power is derived from its production through

\begin{flushleft}
\textsuperscript{18} Croxton, supra note 17, at 570; John Gerard Ruggie, Continuity and Transformation in the World Polity:
\textsuperscript{19} See generally Francis Harry Hinsley, Sovereignty (2d ed., Cambridge Univ. Press 1986); Dieter
Grimm, Sovereignty: The Origin and Future of a Political and Legal Concept (Belinda Cooper trans., 2015).
\textsuperscript{20} Croxton, supra note 17; Philipott supra note 17, at 73–150.
\textsuperscript{21} Henry Kissinger, World Order, 11–48 (2014); Marie-Laure Djelic & Kerstin Sahlin, Rerordering
the World: Transnational Regulatory Governance and Its Challenges, in The Oxford Handbook of
\textsuperscript{22} James Turner Johnson, Sovereignty: Moral and Historical Perspectives 91 (2014);
Croxton, supra note 17; Kissinger, supra note 21; Leo Gross, The Peace of Westphalia 1648–1948, 42
\textsuperscript{23} Thomas G. Weiss & Sam Daws, The United Nations: Continuity and Change, in The Oxford
Handbook on the United Nations 30–40 (Thomas G. Weiss & Sam Daws eds., 2018); John
802 (2003).
\textsuperscript{24} Expressed, for example, by early scholars such as Thomas Hobbes, John Locke, Jeremy Bentham,
John Austin, and Immanuel Kant.
\end{flushleft}
foundational law-making infrastructure, invariably found in the nation-state. Such positivist accounts of law remain or have become—unconsciously, at least—widely subscribed to among people today; most of us do not believe that any law is “law” unless a nation-state formally posits it through domestic legislative and judicial processes. However, as discussed below, the multiple and ever-intensifying processes of globalization over the past half-century have triggered a revisiting of the positivist and nationalist account. Indeed, as will be shown, in an increasingly transnational world, there has been movement toward more pluralistic and sociological accounts of law, taking a broader view of law’s underlying quality as a norm.

Whether the identification of non-state law is achieved, for example, by Brian Tamanaha’s Labelling System, Armin von Bogdandy’s Systems Theory, Gunther Teubner’s Autopoietic Theory, William Twining’s Levels Theory, or Grahf-Peter Calliess’s Running Code Theory, the direction of recent thinking has been predominantly consonant: not all law does, nor should, originate from the nation-state. Law can be written or unwritten, and its sources can be local, communal, religious, supranational, or global, as well as public, private, or hybrid, and can possess normativity on a spectrum between the extremities of hard and soft. In the globalization context, such accounts of global legal pluralism open the possibility that we can be subject to multiple legal obligations, many of which can originate within or without the domestic legal system of the nation-state. Naturally, there remain traditionalists who dispute this perspective. In some ways,

29 See generally Global Law Without a State (Gunther Teubner ed., 1997).
the battle lines are presently drawn between the legal positivists, who maintain an account of law built upon “national” sovereignty, and legal pluralists, who recognize and, in many cases, actively seek to expand additional laws “beyond the state.”34 The pluralist account has certainly been welcomed in the globalized era.35

B. Transnational Law

Over the past half-century, the pluralist account of law has contemporaneously led to the advancement of a new discourse in transnational law.36 Transnational law, as widely understood, seeks to comprehend and even encourage the complex configuration of legal rules and norms, both within and without the state.37 Its appeal lies in its pluralist recognition that numerous actors beyond the nation-state can be a source of legal norms or the principal shapers of the legal system.38 Whether the law takes the form of supranational regulation, global standards derived by non-state bodies, industry self-regulation, private dispute settlement, or community norms, all trans-nationalists recognize that individuals are often subject to rules through a variety of compliance-inducing forces, such as positive, moral, communal, virtual, physical, internal, or natural.39 Yet, it is important to recognize that international law and order still remains state-centric. The vast majority of legal objects obey the legal rules of a singular domestic legal system, whether by territorial situation, contractual choice, or imposition through private international law, and most actors rationalize themselves as so subject.40 As such, transnational law is more of an emerging


36 See, e.g., TRANSNATIONAL LAW: RETHinking EUROPEAN LAW AND LEGAL THINKING (Miguel Maduro et al. eds., 2014).

37 Id. See generally NEGOTIATING STATE AND NON-STATE LAW, supra note 33.


40 See generally Dalhuisen, supra note 26.
diversification and fragmentation in sources and subjects of law, rather than a
persuasive account of *lex lata*.41

In 1956, in his visionary introduction to transnational law, Philip Jessup
defined transnational law as “*all* law which regulates actions or events that
transcend national frontiers. Both public and private international law are
included, as well as other rules which do not wholly fit into such standard
categories.”42 This definition, though an early formulation, still commands an
impressive level of subscription among thinkers today.43 It also exemplifies the
potential vastness of the subject.44 Based on this idea, this Article defines
transnational law as the global legal system that recognizes and promotes all legal
norms of public, private, and hybrid origin, that vary between hard and soft and
that apply to the multiplicity of actors interacting across and between multiple
governance levels—local, national, regional, and global. In other words, both
transnational law and governance emphasize the multi-faceted, multi-level, and
multi-stakeholder nature of the global civil order and of the challenges facing it,
as well as seek to look beyond a state-based approach to law and accountability.

By contrast, the Westphalian account of international law has created an
unfortunately limited dualistic account of law: public international law, which
covers agreements between states, and municipal law, which encompasses national
law within states.45 This has resulted in the apparent neglect of a growing number
of legal norms—outside multilateral treaties and national law—which carry
normative force without sole reliance on state power, such as: (1) industry self-
regulation and standards; (2) supranational law; (3) standards and rules developed
by international governmental, non-governmental or epistemic bodies; (4)
cooperation in law development and enforcement between public and private
partners; (5) other local, religious, or global standards; and (6) all forms and scales
of “governance,” between hard and soft.

Through internal enforcement mechanisms—such as peer pressure, media
scrutiny, economic sanctions, loss of trade access, diminution in consumer
demand, and loss of network access—transnational rules can drive high levels of
direct compliance by stakeholders, often being witnessed in the form of industry


44 See, e.g., Carrie Menkel-Meadow, *Why and How to Study Transnational Law*, 1 UC Irvine L. Rev. 97, 103–05 (2011); Cotterrell, supra note 38, at 501.

standards, certification schemes, internal adjudicative processes, community agreed rules, or other non-state derived laws. To great effect, nation-states can also loan their domestic enforcement architecture to external private, regional, or global legal systems in co-regulatory processes which bolster the external network’s internal power of enforcement.\footnote{See, e.g., United Nations Convention on the Recognition and Enforcement of Foreign Arbitral Awards, June 10, 1958, 330 U.N.T.S. 243; André Nollkaemper, Compliance Control in International Environmental Law: Traversing the Limits of the National Legal Order, 13 Y.B. INT’L ENV’T L. 165, 173–75 (2003).}


\begin{footnotesize}
\begin{enumerate}
\item See generally Lorenzo Casini, Down the Rabbit-Hole: The Projection of the Public/Private Distinction Beyond the State, 12 INT’L J. CONST. L. 402 (2014).
\item See generally Michele M. Betsill & Harriet Bulkeley, Transnational Networks and Global Environmental Governance: The Cities for Climate Protection Program, 48 INT’L STUD. Q. 471; Diane Stone, Global Public Policy, Transnational Policy Communities, and Their Networks, 36 POL’Y STUD. J. 19 (2008).
\end{enumerate}
\end{footnotesize}
resolution in transnational settings; ongoing demands and uses for regulatory harmonization;\textsuperscript{55} mounting interconnectedness of global society, enabling decisions in one state to impact other states’ internal interests;\textsuperscript{56} and the generally observed decline in the role of the nation-state, as traditionally understood, in resolving cross-border challenges.\textsuperscript{57}

C. Lex Maritima

Transnational law’s appeal lies particularly in the romantic notion of finding overlapping international “communities” or networks who subject themselves to a self-crafted legal system built around internal legal rules (self-defined or based on community custom) and external legal rules (state, supranational, and global laws), replete with their own dedicated machinery for internal norm resolution or enforcement (including by negotiation, arbitration, or adjudication).\textsuperscript{58} The most famous such global system is the supposed medieval lex mercatoria, or merchant law. This system has arguably been revived in the modern context with commercial customs and usages, state and non-state in origin, to which transnational commercial actors subject themselves today.\textsuperscript{59} This is largely supported by the almost unquestioning recognition of the legality of arbitration awards and the very high level of internal compliance with them,\textsuperscript{60} as well as new methods of internal coercion across the business community,\textsuperscript{61} with strong reputational, mutual obligation, and internalization mechanisms occasioning


\textsuperscript{58} See generally Transnational Legal Orders, supra note 45.


\textsuperscript{60} Ralf Michaels, The True Lex Mercatoria: Law Beyond the State, 14 IND. J. GLOB. LEGAL STUD. 447, 455 (2007).

compliance. Similarly, one can look at how the sovereign state has acquiesced its role in deciding transnational commercial matters and readily recognizes and enforces applications of non-state law. This picture of a transnational legal system, most idealistically found in the commercial context, has also found expression in numerous other visions of global legal communities, the most interesting of which for present purposes is the lex maritima, or maritime law.

The lex maritima envisages that, long before nation-states appropriated the law of the sea and transcribed it into domestic legislation, much maritime activity was self-governed by the maritime community themselves. The networking of mariners across continental ports in previous centuries arguably necessitated the development of mariners’ own systems of rules and customs. These customs and rules were often enforced internally or via available town councils, merchant courts, and guild consuls. Indeed, it seems well accepted that prior to the embedding of the Westphalian ideology from the 17th century, many of the maritime community’s rules had derived from widely shared codes and customary principles, such as the Lex Rhodiana, Rôles d’Oléron, Laws of Wisby, and the Consolata del Mare. For example, one historian noted how maritime law was regarded as a universal and “common system of law,” given that “[t]here was . . . in those days nothing strange in laws that were not national.” Therefore, the extent to which these maritime codes actually formed a unified common law, in preference to local


63 Stone Sweet, supra note 59, at 638; Michaels & Jansen, supra note 34, at 872; THOMAS E. CARBONNEAU, CASES AND MATERIALS ON INTERNATIONAL LITIGATION AND ARBITRATION 293 (2005).


65 Such as the lex sportiva (sports law), lex informativa (information law or cyber law), lex constructionis (construction law), lex financiaria or argentaria (finance law), and lex petrolea (oil law).


70 Senior, supra note 68, at 260.
custom and decentered regulation, has for some time been a question of academic interest.

More recent and detailed historiographical scholarship on the matter, however, has put considerable doubt on whether such a common maritime law ever existed.\textsuperscript{71} Indeed, many have correctly pointed out that the dearth of social bonds and interdependencies between regional actors in the pre-globalization age makes it inevitable that divergent customs and interests would have undermined any efforts at establishing unified laws across continents.\textsuperscript{72} A great deal of research has, therefore, attempted to disprove the transnational account by disproving the historical account.\textsuperscript{73} Yet, whether a common maritime law existed in the medieval period does not detract from the essential hypothesis that unified systems carry normative advantages in denationalized contexts. As Ralf Michaels has aptly summarized, “whether there ever was a true lex mercatoria . . . [is] relatively secondary.”\textsuperscript{74}

Along this more precise line of inquiry, the commentary is far more unanimous. For example, a system of maritime community-led law should carry additional advantages of efficiency through the utilization of stakeholder resources, lower transaction costs, and strong compliance incentivization.\textsuperscript{75} By penalization, suspension, or ostracism of community members, it would also be possible to effectively punish rule-breakers and to improve trade access by utilizing reputational mechanisms, trust-building, and clearing houses.\textsuperscript{76} Furthermore, there is an argument that such communities of mariners would hold greater esteem toward legal rules that were crafted and enforced by and among


\textsuperscript{73} See, e.g., Kadens, supra note 71; Justice Mustill, The New Lex Mercatoria: The First Twenty-Five Years, 4 ARB. INT’L 86 (2014).

\textsuperscript{74} Michaels, supra note 60, at 449.


\textsuperscript{76} See generally Paul R. Milgrom et al., The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs, 2 ECON. POL. 1 (1990); Avner Greif et al., Coordination, Commitment, and Enforcement: The Case of the Merchant Guild, 102 J. POL. ECON. 745 (1994); Bruce L. Benson, It Takes Two Invisible Hands to Make a Market: Lex Mercatoria (Law Merchant) Always Emerges to Facilitate Emerging Market Activity, 3 STUD. EMERGENT ORDER 100 (2010).
themselves. In other words, by eschewing the strict role of the nation-state, a transnational legal system has a greater capacity to handle the complex, reflexive, and multi-level interactions in the globalized or transnational context.

III. WEAKNESSES OF THE INTERNATIONAL LAW OF THE SEA

It is becoming increasingly clear that there is something wrong with the system of the international law of the sea, rather than merely the content of the laws themselves. However, while law of the sea scholars have made ad hoc or casual references to issues such as zonality, territorial sovereignty, and state compliance, commentary has rarely pointed at the Westphalian system of international law as the root cause of failed ocean management. This section suggests that three integral and interlinked manifestations of the Westphalian system—sovereign exclusivity, sovereign equality, and territorial sovereignty—represent the fundamental weaknesses in our international law of the sea.

A. Sovereign Exclusivity

Sovereign exclusivity refers to the unrestricted authority of states to assume absolute rule over their own subjects. It regards nation-states as entirely unitary systems, where everything that relates to regulatory governance of a nation’s citizens is under the self-determination of a discrete and centralized authority. In the maritime context, as elsewhere, this exclusive sovereignty manifests itself by nation-states freely deciding whether to enter into international treaties. Negotiations, therefore, habitually lead to diluted, ambiguous, and hortatory commitments between states. What is more, assuming that a recalcitrant state even agrees to enter into a resulting treaty, it still possesses discretion as to the

---

78 See Section V.
79 See Section III.A.
interpretation and implementation of the treaty against its own citizens in the domestic realm. Further, for most global public goods—such as the protection of the ocean environment—states often have more to gain individually and less to lose by weak compliance.

Although rarely linked, these manifestations of sovereign exclusivity are important criticisms of the present model of ocean management. First, the majority of international ocean management treaties suffer from the trade-off between invoking strong commitments and the need for widespread ratification. A good example is the United Nations 1982 Convention on the Law of the Sea (LOSC). Although a remarkable achievement in terms of comprehensive and consensus-based treaty making, the LOSC did not receive support and ratification from most key maritime powers until an implementation agreement in 1994—in effect, a rewrite of Part XI—neutered the original vision of fairly sharing the resources of the deep seabed. Formal treaties, therefore, also end up with weak and precatory language, such as requiring states to “cooperate” or that they “should” follow a course of action. Often, the only way to get states to enter into international commitments is by adopting hollow language or developing


For example, a duty of “cooperation” is also incumbent on states under the LOSC with regard to protecting living resources in the high seas, LOSC, supra note 86 art. 117, and protecting the marine environment, LOSC, supra note 86 art. 197. This has led to research which has puzzled over the precise obligations of states. See, e.g., Yoshinobu Takei, Filling Regulatory Gaps in High Seas Fisheries: Discrete High Seas Fish Stocks, Deep-sea Fisheries, and Vulnerable Marine Ecosystems 52–68 (2013).

For example, the Convention on the Protection of the Underwater Cultural Heritage, art. 7, ¶ 3, Nov. 2, 2001, 2562 U.N.T.S. 3 (entered into force Jan. 2, 2009) [hereinafter UNESCO Convention], merely requires that coastal states “should” inform flag states when their sunken warships are discovered in the coastal state’s territorial waters.
international “soft law.”90 Naturally, however, such rules have numerous difficulties, including weaknesses in enforceability and the lack of vigor in compliance.91 Even through a constructivist lens—that sees the gradual hardening of norms by facilitated learning and coordination92—compliance can be damagingly poor for extended time periods. This is often the case until a media fallout from a catastrophic event that at last foments regulatory motivation and action, which is all too frequently after the event.93

Second, states can hold treaty negotiations for ransom, driving forward the hegemonic and politicized nature of ocean law. It is no coincidence that the most powerful maritime nations tend to espouse legal rules that are most closely aligned with international custom.94 Such multilaterally defined laws usually favor those nations found higher in the pecking order of global power. The excessive reliance upon flag state enforcement has suited the most powerful flag states95—just as a “first-come, first-served” system of managing resources in the high seas has suited the most industrialized nations.96 A driving factor in the sudden expansion of state claims in the aftermath of World War II is perhaps that the U.S., U.K., Russia, France, Japan, Canada, and Australia are each in the worldwide top ten of


95 See Section IV.

96 See generally Nico Schrijver, Managing the Global Commons: Common Good or Common Sink?, 37 THIRD WORLD Q. 1252 (2016).
exclusive economic zone (EEZ) size.\textsuperscript{97} The subsequent conclusion of an international convention that permits these states to exclusively extract the wealth of resources hundreds of miles offshore, but with little meaningful legal responsibility to steward the protection of their EEZ’s natural environment, is perhaps unsurprising. As Gregory Shaffer says in another context, international law has “failed to constrain power when power chose to belittle and ignore it, and it served to legitimize power when power deigned to deploy it.”\textsuperscript{98} Seen in this light, the burgeoning naval strength of China in the Southwest Pacific and its growing friction with both the LOSC and the rule of law is as unsurprising as it is predictable.\textsuperscript{99}

An essential result of this politicization of the law of the sea and of the freedom of states to reject or dilute international agreements is the inability to compel or coerce states into assuming additional obligations or burdens.\textsuperscript{100} With its flawed reliance upon states primarily agreeing to be bound by consent, international law allows for commitments between states, which maximizes the opportunity to externalize losses and minimize economic risks from ocean-based activities.\textsuperscript{101} The most visible example is the continual revocation of the system of flag state enforcement. This system for regulating ocean stakeholders—relying on the exclusive enforcement of a flag state’s national legal rules within its domestic courts—is widely felt to be a poor system of ocean supervision and accountability.\textsuperscript{102}

---


Winter 2021 435
requiring that there be a “genuine link” between vessels and the states where they are registered, which is discussed further below—and the fact that most flag states are distant from and indifferent to the true activities of vessels bearing their flag has led many to argue that sole reliance upon flag state enforcement is a formula for failure.\footnote{Mansell, supra note 102; Dorota Englender et al., Cooperation and Compliance Control in Areas Beyond National Jurisdiction, 49 Marine Pol'y 186, 186 (2014). See generally Patricia Birnie, Reflagging of Fishing Vessels on the High Seas, 2 Rev. Eur. Comty. Int’l Env’t L. 270 (2006); Tamo Zwire, Duties of Flag States to Implement and Enforce International Standards and Regulations—And Measures to Counter Their Failure to Do So, 10 J. Int’l Bus. & L. 297 (2011).} Indeed, many “flags-of-convenience” specialize in maximizing the internalization of financial gains and the externalization of environmental or health and safety harms.\footnote{See Rayfuse, supra note 102, at 25. See generally Tina Shaughnessy & Ellen Tobi, Flags of Inconvenience: Freedom and Insecurity on the High Seas, 5 J. Int’l L. & Pol’y 1 (2007); Tony Alderton & Nik Winchester, Globalisation and De-regulation in the Maritime Industry, 26 Marine Pol’y 35 (2002); Tony Alderton & Nik Winchester, Regulation, Representation and the Flag Market, 4 J. Mar. Res. 89 (2002).}

Another manifestation of sovereign exclusivity is the complete freedom of states to interpret, implement, and enforce resulting treaties. Such commitments between states are not only weak in compliance pull, but create international agreements that are deliberately vague and ambiguous.\footnote{See generally supra notes 82, 88–91, 93 and accompanying text.} Examples abound in the maritime context, including phrases such as “maximum sustainable yield” in both Article 119 of the LOSC\footnote{LOSC, supra note 86, art. 119(1)(a).} and Article 5 of the 1995 U.N. Fish Stocks Agreement,\footnote{Agreement for the Implementation of the Provisions of the United Nations Convention on the Law of the Sea of 10 December 1982 Relating to the Conservation and Management of Straddling Fish Stocks and Highly Migratory Fish Stocks, art. 5(b), opened for signature Dec. 4, 1995, 2167 U.N.T.S. 88, (entered into force Dec. 11, 2001).} and also “purposes of scientific research” in the 1946 International Convention for the Regulation of Whaling.\footnote{International Convention for the Regulation of Whaling, art. 8, Dec. 2, 1946, 161 U.N.T.S. 72 (entered into force Nov. 10, 1948).} Such equivocal phrases are intentionally included to provide sufficient latitude in self-interpretation and self-discipline, so as to incentivize objecting or free-riding states to join the treaty regimes.\footnote{See generally Dražen Pehar, Use of Ambiguities in Peace Agreements, in LANGUAGE AND DIPLOMACY 87 (Jovan Kurbalija & Hannah Slavik eds, 2001).} While this practice is often coined “constructive ambiguity,” an alternative term could be “destructive ambiguity,” given how states often flout such well-intentioned phrases and interpret them in a self-interested manner that is destructive to the wider community.\footnote{See generally Italy Fischhendler, When Ambiguity in Treaty Design Becomes Destructive: A Study of Transboundary Water, 8 Glob. Env’t Pol. 111 (2008); Susanne Therese Hansen, Taking Ambiguity Seriously: Explaining the Indeterminacy of the European Union Conventional Arms Export Control Regime, 22 Eur. J. Int’l Rel. 192 (2016).}
This unrestricted freedom of states to interpret, implement, and enforce the laws governing their citizens is at the heart of the struggling system of ocean stewardship. Its weakness is perhaps most vividly manifested in the unconditional freedom of states to self-interpret and enforce the genuine link requirement for registering vessels under Article 91 of the LOSC. As flag states assume the central responsibility for managing offshore operations, it is vital that those operations possess a meaningful relationship with the supervising flag state and, more so, that they are residents of or hold identifiable assets in that country against which sanctions can be enforced. Unfortunately, open registry states—states providing flags-of-convenience—are almost entirely free to self-interpret the genuine link requirement according to their own standards. Ironically, international efforts to close this critical loophole through the U.N. 1986 Convention on Conditions for Registration of Ships failed because the intended addressees—by ultimately using this very same freedom to act autonomously—were free to reject the treaty. In 1999, the International Tribunal on the Law of the Sea revisited this loophole in the M/V “Saiga” case. Unfortunately, the Tribunal held that the strength of a genuine link between vessel and flag state is not a matter that can be contested by others (outside the flag state itself) nor a question of the quality of state regulatory oversight, but rather is purely an administrative question of whether the flag state has been formally appointed as the registered flag state.

This failure of unencumbered internal sovereignty goes much further. For example, it enables offshore tax havens, money laundering, asset moving, forum-shopping, and the creation of impenetrably complex, multi-front company

---


structures across multiple jurisdictions. Thus, the freedom of states to craft their own regulations fails in a system that allows ocean stakeholders—legal or natural persons of a truly transnational quality—to freely select a jurisdiction to hide assets, register front companies, hear foreign claims, align their environmental standards, access markets, and pay taxes.\(^\text{117}\)

What is more, each national legal system is free to interpret and implement their commitments across all sectors in an endless variety of ways. One method is to create a complex “horrendogram” of multiple overlapping and conflicting policies that make it even more challenging to identify clear norms and ensure their observance.\(^\text{119}\) This intensive fragmentation of law creates not only a great uncertainty of ocean law, but also gives wide berth for different interpretations. The further result of this lack of certainty is that state obligations are rarely opaque, leaving room for despondency when it comes to implementation.\(^\text{120}\)

Given that international commitments are arranged horizontally between political sovereigns, their subsequent implementation relies on a complex, costly, and arguably cumbersome system of interstate bilateral and diplomatic enforcement. In other words, an “injured” state needs to invest valued political resources—including civil service time, finance, and goodwill—to direct enforcement against an evidently “culpable” state.\(^\text{121}\) All transnational ocean users

---


118 See generally Victor Galaz et al., Tax Havens and Global Environmental Degradation, 2 NATURE ECOLOGY & EVOLUTION 1352 (2018); Jane Marc Wells, Comment, Vessel Registration in Selected Open Registries, 6 MAR. L. 221 (1981); Pink Leon, SWEATSHOPS AT SEA: MERCHANT SHAMEN IN THE WORLD’S FIRST GLOBALIZED INDUSTRY, FROM 1812 TO THE PRESENT (2011).


must therefore petition their own nation-state to take on their litigative mantle, creating a constrained and indirect route between two “foreign” stakeholders.\(^{122}\)

Nollkaemper refers to an uncited case which perfectly illustrates this quandary.\(^{123}\) Here, fishermen in the North Sea brought a claim against the German government alleging that a permit authorizing a factory to dump acid in the North Sea—which killed and deformed many of the fish stocks they relied on—was in breach of the London Convention\(^{124}\) and the Convention for the Prevention of Marine Pollution by Dumping from Ships and Aircraft, ratified by Germany.\(^{125}\) Despite clear evidence of contravention of these agreements, however, the Federal Administrative Court of Hamburg rejected the claim on the basis that these international agreements were between states and did not create private rights for the ocean users themselves.\(^{126}\)

The horizontal nature of sovereign equal-state relations also effectively minimizes available sanctions and further neutralizes the effectiveness of the adjudicatory process. This emphasis on state interests and state responsibility provides states with the freedom to discount the external interests of the international community or the internal interests of one’s national community. It also leads to a problematic mismatch in the allocation of governance authority and, often, to self-centered decision making in areas with transnational impacts.\(^{127}\) Certainly, there are movements in the right direction toward new,\(^{128}\) cosmopolitan,\(^{129}\)

---


\(^{123}\) This case is referred to in Nollkaemper, supra note 46, at 177–78.


\(^{126}\) Nollkaemper, supra note 46, at 177–78.


\(^{129}\) See supra note 122.

---

Winter 2021 439
interdependent, relational, responsible post-Westphalian, and contingent or conditional forms of sovereignty. But these developments are a slowly emerging byproduct of globalization and gradual universal integration. They do not, therefore, excuse traditional norms of nonintervention and firm sovereign boundaries as culprits for our presently failing global environmental stewardship.

What is more, at their core, these alternative forms of sovereignty are merely idealized or aspirational concepts of international relations, rather than depictions of the true allocation of legal authority that we rely on today. This is particularly the case with international agreements to divide up the ocean’s wealth. Thus, they only reflect the slow and tired process of states reacting to political crises that routinely occur after the event, once disasters finally attract sufficient media coverage to generate political currency. The Westphalian expectation of exclusivity of national jurisdiction, along with its intense distrust of systems of shared responsibility, therefore, forces the hand of the law of the sea and its arbiters toward maintaining the status quo.

Critically, this stringent doctrine of sovereign exclusivity promotes the widespread norm of nonintervention. Robert Jackson even described nonintervention as a grundnorm of Westphalian sovereignty and, certainly, the


customary norm that flagrant rule-breakers can only be interdicted under narrow circumstances is a feature that provides our oceans with an aura of lawlessness. In the EEZ, the number of circumstances under which a coastal state can intercept or regulate nonstate vessels is somewhat greater than that of the High Seas, but it is still limited to specific and discrete conflicts that center on the coastal state’s economic interests. Consequently, many wider security issues are left out, such as military operations, organized crime, and environmental crime. Regulation is further restricted to prescriptive rules set at the international table, such as those negotiated by the International Maritime Organization.

This guarding of flagged vessels roaming the oceans from any interference is not the underlying notion of the mare liberum, as is frequently misunderstood. Rather, it is a wholly Westphalian idea that national governments exclusively govern their respective citizens with no other nation or institution permitted to intervene or share supervision. This system results in flag states undertaking regulatory “supervision” from jurisdictions with no practical connection to activities and located thousands of miles away. It also permits flag states wide latitude in the design and enforcement of the standards against which their fleet are monitored, thereby bringing a vital source of income to that state. As a result, there are countless reports of underenforcement and poor supervision by flag states. These systems of flag state supervision are so defective that


142 See Section IV.

143 See generally supra note 137 and accompanying text.

144 See generally LANGEWIESCHE, supra note 1; GEORGE, supra note 117.

145 See High Seas Task Force, supra note 85, at 52–54.

oceangoing vessels have been referred to as “neglectful,” “outlaws,” “lawless,” “mobile pockets of sovereignty,” “sovereign islands,” “delinquent,” and “a law unto themselves.” Moreover, given strong links between organized crime and poor flag state supervision, rogue vessels are frequently synonymized with piracy.

B. Sovereign Equality

Like sovereign exclusivity, sovereign equality holds that all states are self-governing and unitary. It is principally concerned, however, with the horizontal nature of state relations. As an important principle for preventing a world ordered by military or economic power, the equal treatment of states accords identical legal rights and responsibilities to each state. In reality, however, the strict interpretation of sovereign equality and the routine treatment of all states as equals further propagate the consent-based order of international law. This is because such an order removes the opportunity to create any over-arching authority and, with it, any capacity to compel or coerce noncompliant states into

148 See generally Langewiesche, supra note 1.
150 High Seas Task Force, supra note 85, at 35.
producing global public goods. Indeed, such a principle could secure for states the equality of responsibility, opportunity, and rights to self-governance despite actual asymmetries in these areas. Ironically, sovereign equality therefore arguably derails distributive justice and sustains illiberal democracies by permitting all nations to enjoy equal authority, even if their internal systems are corrupt or harmful to social and environmental interests.

Horizontalism—while intended to minimize anarchy and hegemony—still results in ineffective enforcement powers, thus ensuring that international law centers around the same power politics and is habitually undermined by its realist limitations. Recent and notorious examples of this freedom of the system’s intended subjects to reject unfavorable interpretations of the law include: Japan’s continued refusal to follow the rulings of the International Whaling Commission, the Russian Federation’s refusal to recognize the compulsory jurisdiction and ruling of the Permanent Court of Arbitration in the Arctic Sunrise case in 2015, and China’s refusal to recognize the arbitration panel’s compulsory jurisdiction and ruling 2016, which rejected China’s amassing territorial claims in the South China Sea.


159 See generally Roth, supra note 157; Michael Blake, Distributive Justice, State Correlation, and Autonomy, 30 PHIL. & PUB. AFF. 257 (2001).


Horizontalism’s equality requirement also opens space for conflict over the potential hierarchization of international norms,\textsuperscript{164} thus potentially bulwarking defenses against anything beyond a narrow interpretation of peremptory or \textit{erga omnes} norms intended to provide for universal responsibility. In its external manifestation, as with exclusivity, this equality requirement also gives states the power to ritually contest the jurisdiction, or worse—the legitimacy of external institutional processes.\textsuperscript{165} As a result, states also habitually prefer to avoid politically transparent and expensive adjudicatory processes and, instead, resolve matters through drawn out and obstacle-ridden diplomatic channels.\textsuperscript{166} This not only breeds uncertainty and indecision, but also reduces the opportunities to clarify or develop international legal jurisprudence.\textsuperscript{167} This lack of adjudication is then compounded by the narrow focus of states upon economic or political interests when justifying the pursuit of international claims, which further limits the opportunity for hearing and advancing the rules of responsibility for producing global goods, such as those protecting the international environment.\textsuperscript{168}

A horizontal system of \textit{inter-national} relations also promotes a system of constant competition between states. One state’s inalienable right to undertake a course of action free from interference permits that state to freely produce externalities, which can only be absorbed by another state. Given the harmful interoperation of free riding and the prisoner’s dilemma in addressing collective action challenges, states operating as equal bargaining agents usually treat international relations as a zero or negative-sum game. At the same time, selfish decision making can still be rewarded, and altruism—causing short-term socioeconomic loss to one’s own citizens—risks punishment.\textsuperscript{169} A system of \textit{international} relations in which states can, consciously or unconsciously, externalize losses and maximize gains provides the perfect environment for regulatory “races


\textsuperscript{168} \textit{See} sources cited \textit{supra} note 163.

to the bottom,” where states are forced to compete for limited available resources.\textsuperscript{170} It also locks states into zero-sum games where they fear that a strategy of abatement, such as from the presently unsustainable subsidization of industrial-scale fishing, will lead to considerable economic losses for them in exchange for gains to free riders.\textsuperscript{171} In other words, the Westphalian system is built entirely around a false belief of independence, while the shared use of a globalized and transnational ocean can only ever be interdependent.

The drive to attract ship registrations, processing fees, company registrations, and legal fees incentivizes competition among states to offer more effective flags or ports of convenience.\textsuperscript{172} The more that a state can externalize losses—such as ensuring that environmental degradation takes place overseas or that foreign citizens are unable to pursue economic claims against its nationals—the more it can profit.\textsuperscript{173} States locked in this competition for maritime business are treated as equals, regardless of whether they actually possess the necessary resources, expertise, or regulatory infrastructure to properly supervise their flagged vessels or enforce legislation. In fact, in the majority of cases, they do not.\textsuperscript{174}

Underenforcement and turning a blind eye, therefore, become the norm for popular ports and flag state regulators.\textsuperscript{175} As research has consistently shown, states locked in such negatively-reinforcing spirals will find it immensely difficult to break out of such patterns of behavior within a consent-based legal system.\textsuperscript{176} Only by hundreds of ongoing interactions can actors engage in “repeat games,” thus building up the trust and goodwill that enables them to agree to more meaningful rules or better systems of enforcement. Nevertheless, when each state

\begin{flushright}
\textsuperscript{172} \textit{See generally} supra note 146 and accompanying text; Shaughnessy & Tobin, supra note 104; Wells, supra note 118.\textsuperscript{173} \textit{See generally} supra note 101.
\textsuperscript{175} \textit{See generally} supra note 169 and accompanying text.
\end{flushright}
has the equal and unrestricted freedom to externalize losses—particularly by avoiding “unequal” regulatory oversight from a higher order or a collective of foreign states—the temptation to free ride or withdraw from efforts at regulatory integration can continue to undermine collaborative efforts. This is particularly true where states lack true political or economic incentives to constrain their own sovereign freedoms.  

These harmful effects of horizontalism are pervasive—even between politically friendly nations. For example, regional fisheries management organizations were specifically designed to remove comparative trade-offs between states in a regional context and to ensure coordination and the collective raising of regional standards. Yet, even here, there is strong international competition and pathologies of free riding. Indeed, Donald Rothwell reports how—even between regional neighbors—the achievement of effective cooperation patterns is still entirely contingent on:

[the] overall political relationship between the States concerned, cultural and socio-economic divergences, the presence or absence of pervasive territorial or maritime disputes, the significance accorded to and prioritizing of oceans management by individual States, the effective implementation of regional instruments by individual States, the nature and extent of sea-based activities and financial resources and capacity.

Indeed, the rejection by U.K. voters of regionally integrated collective gains in the 2016 Brexit referendum demonstrates just how vividly much of society still perceives their entitlement to self-government and self-advancement under the veil of national sovereignty, even after evidence of sustained collective gains between politically friendly nations.

C. Sovereign Territoriality

Territorial sovereignty—the idea of segregated “zones” upon ocean space—is another symptom of Westphalianism, which has already been recognized as a

---

177 See generally supra note 171.
key regulatory weakness in ocean governance. It is also possible, as with the other two characteristics of sovereignty, to see territoriality as overlapping and interlinked with the other traits. For example, through sovereign equality, all states demand reciprocal rights to any claimed resources in the ocean, which ultimately leads to a global allocation of resource zones. Similarly, through sovereign exclusivity, there is a necessary presumption that one state must be positioned to assume regulatory jurisdiction over each subject matter, with the result that all states have carved up the entirety of the ocean in pursuit of a fair allocation of juridical responsibility for every factual circumstance. While this zonal approach has grown predominantly by creeping unilateral claims to offshore resources, it has also been viewed as a strategic opportunity to propertize all ocean space, with the hope that coastal states will internalize environmental degradation and so guard “their” environmental assets in offshore regions. This strategy, however, ultimately failed given that states could now focus on exploiting their newly acquired resources in these distant offshore spaces, while conveniently treating the protection of the environment “out there” as an externality.

The global patchwork of maritime zones, which results from this inter-state territoriality, has created hundreds of diverse regulatory systems that are cut-off and distinct from neighboring zones. Further, it results in interactions between transnational actors that—moving casually and fluidly across the entire ocean space—must necessarily take place through tired inter-national lines, which also leads to forum-shopping between enforcement agencies and regulatory systems. Ocean ecosystems, humans included, therefore witness constant regulatory gaps and overlaps, despite taking little practical notice of artificially-constructed


183 Barnes, supra note 91, at 236–37, 241–42; High Seas Task Force, supra note 85, at 41; TANAKA, supra note 15, at 8.


185 See generally supra note 116 and accompanying text.

political borders.\textsuperscript{187} Perhaps the biggest driving force behind the development of transnational law in other fields has been the desire to avoid the cost, complexity, unpredictability, apprehension, and inefficiency associated with private international law.\textsuperscript{188} Unless a cross-border claim is particularly strong and carries a significant payoff, it is rarely worth the risk and cost of pursuing.\textsuperscript{189} This failure of law between borders is further compounded by: (1) the near-phantom legal nature of many maritime actors, who operate within multi-front and multinational companies;\textsuperscript{190} (2) the lack of transparency of national actors and agencies, whose efforts to produce global goods can be safely shrouded in hortatory language; and (3) the use of a complex system of multifarious national legal systems, which are randomly allocated by technocratic and idiosyncratic conflict of law rules.

Furthermore, given the absolute freedom of states to self-regulate and reject foreign interference, the use of civil liability regimes—aiming to facilitate cross-border enforcement by harmonization and reciprocation—also fail on account of their ritual rejection or wholesale dilution.\textsuperscript{191} Ocean stakeholders can, therefore, avoid the force of private liability between overseas actors. Moreover, as highlighted earlier, transnational stakeholders must also rely on foreign states to implement effective public and private legislation and have no path to appeal to foreign governments for deficient regulation and enforcement.\textsuperscript{192} The result could be a sense of detachment and disassociation of regulatory actors from the regulatory systems under which they find themselves, and the disassociation of those regulatory systems from the actors.

These permitted cutoffs between regulatory systems in the ocean environment propagates a “Not-In-My-Backyard” attitude in which regulators naturally focus on investing in protection over internal or local interests and are


\textsuperscript{190} See generally supra note 117 and accompanying text.


\textsuperscript{192} See generally supra note 122 and accompanying text; supra note 125 and accompanying text.
tempted to disregard external matters. Perhaps most fundamentally, the artificial fragmentation of ocean space into pockets of national interest prevents an efficient integration of capacities. As such, the considerable security resources, data, equipment, skills, and rules needed to achieve effective ocean management and protection are habitually and inefficiently duplicated side-by-side—rather than harmonized together in effective and coordinated regional systems between all actors and agencies. Thus, even if states work together to achieve collective action, the underlying belief system built around exclusive sovereign “rights” and “ownership” in each maritime zone continues to undermine any sense of joint and several responsibility. This is at the heart of demands for a more integrated, regionally-coordinated, and ecosystems-oriented model of ocean governance.

D. Westphalian Ocean Management: A System of Recurrent “Gaps”

Many of the findings surrounding the ocean’s reliance upon Westphalian intergovernmentalism, therefore, highlight an endemic recurrence of “gaps.” First, we witness recurrent knowledge gaps. This relates to the lack of communication between stakeholders and regulators, meaning that local or private actors’ knowledge or interests have not been effectively incorporated into regulatory decision making. It also refers to the lack of informed decision making by market actors, the lack of data exchange, resource pooling, and surveillance cooperation between regulators and enforcement agencies, the lack of accurate scientific data, and the need for marine stakeholders to effectively cooperate


197 See generally supra note 194 and accompanying text.

and communicate regarding each other’s relative activities, resources, and interests.¹⁹⁹

Second, the geographical gaps inherent in ocean governance are widely known, and—as discussed in the context of sovereign territoriality—the fragmented nature of ocean management is a familiar opprobrium. Not only do these gaps relate to the weakness and inefficiency of completely segregated regulation,²⁰⁰ but also to the fluid movement of persons, species, and other objects between political borders;²⁰¹ the lack of regulatory regimes in many regions around the world;²⁰² and the failure of multilevel coordination and cooperation between regulatory systems.²⁰³

Third, closely interrelated, but less obvious, are the prevalent regulatory, incentive, normative, and compliance gaps in transnational ocean management. Regulatory gaps occur where legislation or regulatory processes are lagging or suffer from poor coordination. For example, there is a particular concern with divided schemes of regulation where disconnected organizations and actors operate within regulatory silos.²⁰⁴ Critics also regularly point to the lack of coordination and regulatory cooperation between international organizations, regulators, national legislatures, and enforcement agencies.²⁰⁵ What is more, we witness a languid pace of regulation, especially when it is negotiated through political inter-state

¹⁹⁹ See generally Oran R. Young et al., Solving the Crisis in Ocean Governance: Place-Based Management of Marine Ecosystems, Environment, 49 ENV’T 21 (2007); Beverley Clarke et al., Enhancing the Knowledge-Governance Interface: Coasts, Climate and Collaboration, 86 OCEAN & COASTAL MGMT. 88 (2013).

²⁰⁰ See generally supra notes 184–194 and accompanying text.


bargaining processes that fail to keep pace with constantly shifting and intensifying human activities in the ocean or with new threats and opportunities.\textsuperscript{206}

Fourth, and related, normative gaps describe deficiencies in accountability and the widespread lack of mandatory compliance-inducing obligations, with a recurring reliance upon ambiguous language or hortatory commitments, including within unenforceable rules or industry self-regulation.\textsuperscript{207}

Fifth, compliance gaps refer to the issue of compliance by states with their international commitments and with the poor level of implementation and enforcement by national regulators when dealing with externally valued public goods.\textsuperscript{208} It also refers to the lack of compliance by stakeholders with national law, as they freely select between, or distance themselves from, traditional regulatory structures.\textsuperscript{209}

Sixth, closely related to normative and compliance gaps are the incentive gaps, which acknowledge the gap between a desired regulatory object and the incentive on the part of political actors to comply. Only by enabling consensus-made law to actually reign over consent-based self-enforcement, along with the eradication of horizontalism by the building of more powerful enforcement mechanisms, can regulators be incentivized to invest in much-needed cooperation and to break down the harmful interoperation of the prisoner’s dilemma and free riding, as discussed above.\textsuperscript{210}

The consent-based, horizontal, and zonal system of public international law has thus weakened the effectiveness of ocean management for so long that an entirely new approach to ocean law and governance is needed. In order to bridge all these “gaps,” such a new approach—which is discussed in Section V—must operate as a truly multi-stakeholder system which is arranged with true power reallocated across the global, regional, local, and transnational scales. In other words, we need a true and real attainment of a transnational law of the sea.

\textsuperscript{206} Ardron et al., supra note 202, at 101; Gjerde et al., supra note 1, at vii.


\textsuperscript{210} See supra note 169 and accompanying text.
Despite all the evidence that Westphalianism is centrally to blame for our struggling law of the sea, this has been hardly noticed or noted in existing academic discussions. In 2006, Richard Barnes, David Freestone, and David Ong did tell us that:

the underlying emphasis of prescriptive and enforcement authority [of the law of the sea] is in the hands of individual States. This reflects the more fundamental nature of international law as a horizontal legal system in which States are sovereign equals under no higher authority than that of international law.211

In 2010, Thomas Gammeltoft-Hansen and Tanja Aalberts also observed that the present-day conception of the ocean as a liberal free-for-all creates a space in which “participating states may successfully barter off and deconstruct responsibilities by reference to traditional norms of sovereignty and international law. Thus . . . the Mare Liberum [sic] becomes the venue for a range of competing . . . disclaims to sovereignty.”212 Similarly, Barnes only briefly noted in 2015 that while the flag state system, to him at least, has not failed, it “is far from effective”213 and was “facilitated by the emergence of the modern political State after the peace of Westphalia.”214

In 1992, Philip Allott provided a welcomed and perceptive criticism of both the Mare Liberum and Mare Clausum doctrines. He proposed the need for a new socially inclusive system: the Mare Nostrum, or “Our Sea.”215 Importantly for Allott, the law of the sea had become a system which:

while still seeming to the uninitiated to be a process for the collective formation of social objectives, turned itself in practice into a system that gives effect neither to universal social objectives of all humanity nor to the social objectives of all human individuals collectivized through the state systems. It came to be dominated by an independent dialectic at the median level between the two, the level of relations between so-called states. Humanity had formed itself into a society whose social process was interstate relations.216

Nevertheless, there appears to be a discernible lack of focus specifically on public international law’s failures in the ocean’s transnational ecosystem. Instead, the vast majority of literature and research in marine governance focuses on Grotius’s doctrine of the “freedom of the seas” and the problems with relying upon flag states to regulate global concerns.


213 Barnes, supra note 102, at 321.


215 See generally Allot, supra note 207.

216 Allot, supra note 207, at 775–76.
IV. THE SCAPEGOATS OF OUR FAILING OCEAN LAW

A. Mare Liberum: Res Communis or Res Nullius?

After four centuries of compressing Grotius’s 1609 *Mare Liberum* into a dogma that propagates the interests of powerful flag states, this utopian vision for ocean governance has become frequently misinterpreted. Grotius’s now-famous monograph was originally written as a legal brief supporting the Dutch East India Company’s capture and sale of the Portuguese vessel, *Santa Catarina*, in the Malacca Strait. Its adept integration of wider literature from Ancient Rome and Greece, and from the views of natural and canonical law, along with Grotius’s posthumous fame as an international legal theorist, have caused it to have considerable influence—in name, at least—on our view of ocean law in the centuries that followed. Most commentators on the law of the sea today still hold considerable disdain for the supposed concept it introduced: the so-called and now infamous “Freedom of the Seas.” This concept allegedly sought to view the ocean’s resources as *res nullius* and open to appropriation and exploitation. Unfortunately, however, Grotius stands falsely accused. His original “Free Sea” did not promote the systemic free-for-all that we can still detect in global marine lore today but, as will be shown, was actually much more progressive.

Primarily, this loss in translation has been caused by the convenient simplification of ocean management into a binary system of competing strategies—that between the *Mare Liberum* and the *Mare Clausum*. On the one hand, there is the argument for increased territorialization, which would enable coastal states to more successfully control and effectively regulate large swaths of the

---


220 See generally supra note 219.
This argument is said to have arisen from John Selden’s equally famous rejoinder to Grotius in 1635, titled Mare Clausum, or “Closed Sea.” On the other hand, there is the argument that the ocean should not be territorially appropriated by anybody but should be open for free and unrestricted access by all—supposedly exemplified by the Mare Liberum.

These two doctrines have been contorted over the years to try and represent two binary Westphalian approaches to solving ocean governance. The Mare Clausum represents the view that coastal and port states should have more power, given that flag states are often too disconnected from distant ocean activities. Meanwhile, the Mare Liberum has been inappropriately reinterpreted by the most powerful flag states of the day—such as the Netherlands, France, the U.K., and the U.S.—to support the view that the ocean should be “free” from regulation as far as possible and that, instead, states should primarily regulate their own fleets through flag state law. In reality, as argued in Section IV.B, both are systems prone to failure because both entail states acting in exclusive national interest and, hence, externalization of losses to the international community.

Much of the misplaced blame upon the Mare Liberum seems to arise from the word “freedom.” Grotius did indeed argue that the ocean’s vastness, fluidity, and unsuitability to physical dominion should render it free from territorialisaton. Crucially, however, he never posited that the ocean and its resources should be “free” property to be appropriated or exploited. Rather, quite oppositely, he argued that the ocean’s resources cannot be appropriated without prior permission from the international community to which they belong. In other words, he saw its resources as res communis—rather than res nullius—which was to be shared among all humankind with “a common right [of usage] over which no other right could be asserted.” It was, therefore, a rejection of our contemporary understanding of private dominion over ocean property, in that ...

222 JOHN SELDEN, MARE CLAUSUM, SEU DE DOMINIO LIBRI DUO (1635).
such ownership means the ocean is “incapable of belonging to someone else as well.”\textsuperscript{227} This common ownership model was premised on the fact that “there [is] nothing to prevent a number of persons from being joint owners . . . of one and the same thing.”\textsuperscript{228} He argued, thus, that the ocean was a gift from God and “not to this or that individual, but to the human race,”\textsuperscript{229} and “that all that which has been so constituted by nature that although serving some one person it still suffices for the common use of all other persons, is today and ought in perpetuity to remain in the same condition as when it was first created by nature.”\textsuperscript{230} As a result, only those activities which can be done “without loss to anyone else” are tacitly permitted by the global community.\textsuperscript{231}

On this basis, Grotius was able to argue that certain “freedoms” should be automatically available because they are not prima facie causing loss to humankind. For example, his primary argument was actually that maritime navigation can only be free and unrestricted because—at his time at least—ships leave “no more than a ‘track in the sea,’”\textsuperscript{232} and the ocean “is not exhausted by that use.”\textsuperscript{233} This unrestricted freedom of access was based on the tacit permission given by humankind—as the ocean’s communal owners—given that (pollution-free) navigation benefits wider community interests.\textsuperscript{234} In another example, he posited that coastal development is permitted because “it is lawful to build upon the shore if it may be without the hurt of the rest.”\textsuperscript{235} Therefore, he did not promote a general “freedom” of high seas fishing under some kind of liberal system permitting the private accumulation of unclaimed property, but rather that fishing might be tacitly accepted by the international community as a free activity, given that fish stocks were—at that time, at least—seen as essentially inexhaustible.\textsuperscript{236} Naturally, in the face of industrial-scale fishing and the decimation of global fish

\textsuperscript{227} Grotius, supra note 226, at xiv–xv (“[Laws have been set down so] that all surely might use common things without the damage of all and, for the rest, every man contented with his portion should abstain from another’s.”); Hugo Grotius, Mare Liberum 1609-2009 55 (Robert Feenstra ed., 2009) (1609); Rossi, supra note 225, at 19.

\textsuperscript{228} Grotius, supra note 227, at 53.

\textsuperscript{229} Id.

\textsuperscript{230} Hugo Grotius, Freedom of the Seas or the Right Which Belongs to the Dutch to Take Part in the East Indian Trade 27 (James Brown ed., Ralph V.D. Magoffin trans., 1916) (1609).

\textsuperscript{231} Id.

\textsuperscript{232} Schrijver, supra note 96, at 1254.

\textsuperscript{233} Grotius, supra note 230, at 26–27 (“[I]f any of those things by nature may be occupied, that may so far forth become the occupant’s as by such occupation the common use be not hindered.”).

\textsuperscript{234} Id. at 33 (“If any should forbid another to take fire from his fire . . . and light from his light, by the law of human society I would accuse and sue him to condemnation.”).

\textsuperscript{235} Id. at 27.

\textsuperscript{236} Schrijver & Prislan, supra note 226, at 176; Grotius, supra note 230, at 38 (“If a man were to enjoin other people from fishing, he would not escape the reproach of monstrous greed.”).
stocks four hundred years later, this tacit approval of unrestricted appropriation of fisheries would patently no longer exist.

Recent reinterpretations of the *Mare Liberum* therefore suggest that pro-flag state theorists have, over the years, considerably skewed the concept of the freedom of the seas. As Vid Prislan and Nico Schrijver put it, Grotius’s arguments “later digressed into ‘first come, first served’ advantages for industrialized nations.” It is, therefore, not surprising that the two dominant-maritime nations, that have held hegemonic regulatory power over the oceans since 1805, have promoted a distinctly Anglo-American model of liberal regulation. This model relies on minimal state intervention and the promotion of anarchic self-governance, thus permitting U.K. and U.S. interests to profit considerably from this “free” maritime domination. This is the unfortunate entrenchment of the “free-for-all” interpretation of ocean law, which will continue to suit the dominant maritime powers in each regional ocean zone. It has also been suggested that the misinterpreted “freedom of the seas” did not actually come to dominate until the Industrial Revolution in the mid-19th century. Before this, it was actually John Selden’s argument in favor of territorial appropriation and intensive national regulation that ruled international thinking. Unsurprisingly, this was throughout the era that witnessed the ascendency and indoctrination of Westphalianism and the centering of all regulatory power in the hands of nation-states. Wherein, as a necessary corollary, all law without national territory was viewed as *lex nullius*.

U.S. Supreme Court Justice Story exemplified this liberal model of law in 1826, when he adjudged that “[u]pon the ocean, in time of peace, all possess an entire equality. It is the common highway of all, appropriated to the use of all, and

---


238 Schrijver & Prislan, supra note 226, at 206. For example, in 1950, Lauterpacht defined it as ensuring “freedom of navigation, unhampered by exclusive claims of individual states, and freedom of utilization of the resources of the sea to a degree to which they can be equitably utilized by all.” Hersch Lauterpacht, *Sovereignty over Submarine Areas*, 27 Brit. Y.B. Int’l L. 376, 407 (1950).

239 R.P. Anand, *Origin and Development of the Law of the Sea*, 152–53 (1983). Anand notes that the “hallmark of this law, ever since its acceptance in the early nineteenth century, has been, in accordance with the spirit of the time, ‘freedom’ meaning essentially non-regulation and *laissez faire*.” Id. at 152.

240 Id. at 151.

241 Id. at 228–29. Anand notes how Selden “won this protracted ‘battle’ not by the brilliance of his arguments, but by the ‘louder voice’ of the powerful British Navy.” Id. at 229. In fact, the idea and concern that the ocean can be enclosed by dominant maritime powers is one which has existed for millennia. See e.g., Rossi, supra note 225, at 39–41.
no one can vindicate to himself a superior or exclusive prerogative there.” And, a century later, the Permanent Court of Justice, in the Lotus Case, still understood that:

[v]essels on the high seas are subject to no authority except that of the State whose flag they fly. In virtue of the principle of the freedom of the seas, that is to say, the absence of any territorial sovereignty upon the high seas, no State may exercise any kind of jurisdiction over foreign vessels upon them.243

In other words, if there is no territorial or non-flag state that is permitted to impose unilateral regulation in open ocean waters, then the only alternative under an international system must be the state that is embodied in the vessel itself. This understanding that either territorial states or flag states govern the seas continues to the present day.244 As R.P. Anand states, over the past two centuries, it became “accepted as an indisputable, almost sacred, dogma which was supposed to be in the interest of all mankind and which nobody would dare challenge.”245 Its total transmogrification can, for example, be witnessed in Bo Johnson Theutenberg’s words, who understood it as meaning that the ocean is “open and free for the use of all nations,” where “[n]o nation could prevent another from carrying on traditional activities at sea”246—a far cry from the system of common society and pooled interests promoted by Grotius.

These accounts of the freedom of the seas became corrupted for one very simple reason: if only nation-states have exclusive and unencumbered regulatory power, but coastal states do not have territorial dominion, then only flag states are left as the option for national legal control of the high seas. By its very design, Westphalianism is innately antagonistic to any possible suggestions of “shared” governance over a territorial space. Furthermore, given that states are “equals” and free to act in self-interest, they cannot be compelled to accept shared ownership of the ocean’s resources. Therefore, under the Westphalian “inter-state” regulatory system, the only two choices for the “sovereign-free” high seas were either lawlessness or regulation by distant flag states.

In sum, by promoting a system that sees the ocean as a gift to the global community and shared by all, Grotius’s argument in favor of turning the ocean into a global commons was in fact highly analogous to Arvid Pardo’s famous 1967 appeal to turn much of the seas into the “Common Heritage of Mankind” and to foreclose the ocean from ever-increasing national territorialization and

---

242 See generally The Marianna Flora., 24 U.S. 1 (1826).
243 S.S. Lotus (Fr. v. Turk.), Judgment, 1927 P.C.I.J. (Ser. A) No. 10, ¶ 64 (Sept. 7) (emphasis added).
244 U.N. Secretariat, supra note 237, at 2–3 (“The idea of the freedom of the high seas is, paradoxically, a survival of the idea . . . that the high seas are subject to dominion and sovereignty, just like any territorial dominion.”).
245 ANAND, supra note 239, at 255.
246 Johnson Theutenberg, Mare Clausum et Mare Liberum, 37 ARCTIC 481, 492 (1984).
unilateralism.\textsuperscript{247} This engaging vision aimed to bring the entire community of humankind within the frame of ocean governance and, out of necessity, sought to redefine and mollify the territoraility of ocean space and the politically exclusive nature of all the ocean’s interconnected zones, including the territorial sea.\textsuperscript{248} Such a new and holistic global commons management model has thus been covered, with some social scientists concurring that the failed understanding of Grotius seems to come in failing to distinguish between “common access” and “common property.”\textsuperscript{249} As a result, the nation-state should no longer be seen as the only representative agent of humankind’s shared ocean and as the only source of transnational legal power.

B. The Fault of “Flag” States or Flag “States”?

The political tug-of-war between states to resolve the exclusive sovereign rights and freedoms inside hard borders in ocean space is primarily built on the economic ambitions of sovereign states and often plays out through diplomacy, political posturing, and displays of military power.\textsuperscript{250} As Daniel Cheever puts it, “[n]owhere is the indissoluble relationship between politics and law demonstrated more cogently than in the law of the sea.”\textsuperscript{251} The outmoded debate between the two original competing systems of governance—between flag state regulation or territorial state encroachment—is often couched in the language of a normative debate over the more suitable, effective, and customary system for managing the seas. A closer look at the contested negotiations, however, makes clear that the debate has never been about which mode of governance is more effective, but about which ultimately delivers the propounding state the greatest amount of wealth, opportunity, and power. Indeed, the growing tension between the U.S. and China in the South China Sea is really a defense of the right of states to navigate trade,


conduct military operations, and prevent coastal exploitation of a resource-rich region, rather than any justified defense of the quality of flag state regulation as the most effective system of marine governance.\textsuperscript{252}

The same can be seen in reverse, where arguments promoting the rightful placement of coastal state “jurisdiction” are really a masked promotion of coastal state power and ownership of distant resources, regardless of the environmental protection challenges this entails. For example, negotiations over the UNESCO 2001 Convention on the Protection of the Underwater Cultural Heritage were hopelessly distracted by this same contest between the flag states and the coastal states.\textsuperscript{253} The negotiators hardly touched the subject of “which is more effective” in terms of protecting the cultural heritage of the ocean. Instead, most of the time and political energy was focused on suspicions about the underhanded motivations and power aspirations of the “opposing” side.\textsuperscript{254} As a result, despite increased coastal state jurisdiction for protecting underwater cultural heritage in the EEZ being better on practical grounds, it was ultimately blocked by flag states who were distrustful of the motivations of coastal states, referring to their proposals disparagingly as “taking another bite at the jurisdictional apple.”\textsuperscript{255}

In many senses, flag state regulation—built around the aforementioned notion of liberalism and laissez-faire regulation of broad ocean space—has been the preferred ocean governance model for the states who stand to gain the most. Those powerful Western maritime nations—who accumulated considerable wealth and opportunity through colonial expansion and domination of the world’s transcontinental trade networks—are the same powerful nations who have conveniently designed today’s liberal flag state-based model of ocean management. The result is that, when levelling criticisms against the failing law of the sea, the vast majority of international law scholars have, up to now, aimed their cannons upon the over-reliance on “flag state jurisdiction” or upon the areas where this flag state regulation needs tweaking. Consequently, in recent decades, many theorists and policymakers have promoted a system with far greater


\textsuperscript{254} See generally Ariel W. González, The Shades of Harmony: Some Thoughts on the Different Contexts that Coastal States Face as Regards the 2001 Underwater Cultural Heritage Convention, in ART AND CULTURAL HERITAGE: LAW, POLICY AND PRACTICE 308 (Barbara T. Hoffman ed., 2006).

\textsuperscript{255} Id., Blumberg, supra note 253, at 499.
prescriptive and enforcement jurisdiction for coastal and port states in order to overcome the weaknesses of distant flag state regulation.²⁵⁶

This approach—which increases coastal and port state regulation in the international law of the sea framework—has certainly provided a number of important improvements. For example, it has put much greater pressure upon the dominant flag states or upon flags-of-convenience to raise their standards in order to maintain access for their fleets to certain ports.²⁵⁷ The difficulty, however, remains that coastal and port states also suffer from the same symptoms of Westphalian sovereignty—such as inter-state competition, races to the bottom, and the exclusive freedom to externalize global responsibility.²⁵⁸ In other words, the risk of further territorialization and encroachment by coastal states is just as undesirable as a laissez-faire model of distant and disinterested regulation by disconnected flag states.

Indeed, the emphasis on internal interests rather than the collective interests of humankind means that public laws governing port and coastguard authorities usually emphasize criminal activities of national interest, such as inward smuggling, market distortion, and migrant trafficking, above distant environmental concerns.²⁵⁹ As the Task Force on Illegal, Unreported, and Unregulated Fishing on the High Seas reported, not only are domestic enforcement agencies unaware of the true global costs, but any public expenditure by a state on criminal prosecution for activities having an impact beyond national jurisdiction is not recouped.²⁶⁰ Other practical factors also make non-flag states equally awkward as regulators, such as the physical challenges of interdiction and


²⁵⁹ See High Seas Task Force, supra note 85, at 24, 30; König, supra note 256, at 7.

²⁶⁰ See High Seas Task Force, supra note 85, at 33.
inspection, the difficulty with surveillance, and the pressure to avoid additional delays and disruption in port.  

For the very same reasons, market and transit states also make uncomfortable and ineffective regulators. What is more, particularly on account of processing supply chain and transshipment practices, it is difficult to isolate imports that breach environmental standards. They also have considerable difficulty regulating grey or black markets and are at constant risk of falling foul-of-trade protectionism rules. Home states, that have jurisdiction over their own nationals when in home-state territory, are also too distanced from the specific activities and are unsuitable prosecutorial enforcers or receptors of incriminating evidence. Indeed, in recent unverified reports of looting from the Battle of Jutland wrecks by converted trawlers in the North Sea, it appears that numerous port, coastal, and flag states are implicated, but none have shown a willingness to fully invest in the security of wrecks more strongly valued by Britain and Germany. Even in their position as a home state, the British authorities did not appear to have prosecuted a U.K. national reported to have been involved in the

---


illicit activities,\textsuperscript{267} despite this state having recognized the need to proactively protect such underwater cultural heritage.\textsuperscript{268}

Critically, therefore, none of these states—whether flag, port, coastal, market, transit, or national—can avoid the same weaknesses of national sovereignty. They can commonly refuse consent to any international law which does not provide them—and only them—with a political gain on balance. In other words, they require considerable political or economic incentives before they invest the political and economic resources in the punishment of offenders or in restricting their own economic activity in deference to external interests.\textsuperscript{269} Furthermore, even if these states did accept maximal jurisdictional responsibility, they would still stand to gain from competing in the provision of lax regulation and apathetic enforcement.

Understood from this perspective, the sovereign freedom of deficient flag state systems to profit from lax regulation, rather than the use of flag state systems, lies at the heart of the law of the sea’s underlying weakness. Certainly, increased regulation across all the governance nodes would significantly notch-up standards in specific areas and force flag states to join collective standards’ arrangements. The underlying fault, however, lies in the inability to coerce states, flag or otherwise, to modify their behavior in a manner against their own interests, as well as prohibiting states from externalizing global “losses” and not in the use of flag state regulation per se.

V. A Transnational Law of the Sea

The focus of this article has been to pinpoint the causes of the legal system’s failure to manage the oceans, as opposed to introducing every element of a new paradigm of governance. As such, it is beyond its purview to give a comprehensive account of a needed transnational law of the sea. While a detailed global vision of transnational ocean governance would be too large for the current discussion, this Section provides an overview of some important elements within such a system, exploring some of their advantages when productively eclipsing or interoperating with existing inter-state legal systems. Subsection A highlights recent academic efforts to explicate a model of Integrated Ocean Management (IOM), which many have regard as merely a search for an ecosystems-oriented model of governance. It should, in fact, be largely understood, more simply, as a search for a post-

\textsuperscript{267} See Brockman, supra note 266 and accompanying text; see also Dutch Salvors Accused of Looting Warship Wreck, THE MARITIME EXECUTIVE (Sept. 22, 2016), https://perma.cc/RG6Q-EEL8.

\textsuperscript{268} See Josh B. Martin & Toby Gane, Weaknesses in the Law Protecting the United Kingdom’s Remarkable Underwater Cultural Heritage: The Need for Modernisation and Reform, 15 J. MARINE ARCHAEOLOGY 69, 70 (2019).

\textsuperscript{269} See generally Shaffer, supra note 98; GLOBAL PUBLIC GOODS: INTERNATIONAL COOPERATION IN THE 21ST CENTURY, supra note 101 and accompanying text.
Westphalian model of ocean management. Then, building on previous sections, Subsection B suggests that multi-stakeholder inclusivity and multi-level governance are two primary features of such a transnational law of the sea.

A. Understanding Integrated Ocean Management as a Demand for a Transnational Law of the Sea

While the root causes are rarely examined, it in fact appears well-accepted that the problem with ocean governance lies with the system itself, rather than with the legal rules contained within. For example, during detailed studies between 2010 and 2015 by the Ad Hoc Open-ended Informal Working Group to Study Issues Relating to the Conservation and Sustainable Use of Marine Biological Diversity Beyond Areas of National Jurisdiction (BBNJ Working Group), there was consistent pessimism about the inevitably poor implementation or the likely dilution of any resulting commitments between states, rather than the content of any agreed rules. As a result, the BBNJ Working Group urged the need for better cooperation and coordination between “all sectors and all levels,” conceding that “a global universal governance structure remained the best way to promote sustainable marine biodiversity beyond areas of national jurisdiction.”

This gloomy allusion to the present system of law—highlighting concern for compliance, implementation, and enforcement of international agreements, or with pervasive regulatory, normative, or geographical gaps—has also become an increasingly vocalized issue among academics more broadly. For example, David Vousden says that “many leading experts on high seas and ocean governance are now convinced that . . . there is now an urgent need for a transformation to a more suitable legal regime which is more cross-sectoral and integrated in its management approaches and strategies.” As Freestone summarizes, therefore, “virtually all are in agreement . . . that we need far more effective means of enforcing compliance with the norms and structures than we have.”


271 Id. ¶ 12.


it appears to be almost unanimous that there is a need for better institutional structures, systems, and processes in the governance model itself.\textsuperscript{274}

In response to these widely reported systemic issues in the present model, there has been increasingly widespread calls for an approach built around far greater “integration” or, in other words, for IOM.\textsuperscript{275} It is not just global institutions and accords that are relaying these calls for integration across the legal landscape but also academics.\textsuperscript{276} For example, in 2008, Yoshifumi Tanaka wrote how “international documents tend to stress the importance of a holistic approach by referring to [IOM].”\textsuperscript{277} Many others have been equally clear on the need for a new paradigm within IOM.\textsuperscript{278} As Wright and his colleagues recently reported, “[m]any States, scientific experts and civil society groups have . . . repeatedly highlighted the need for integrated ocean governance.”\textsuperscript{279} Yet, despite its undisputed central placement on the global agenda, the actual meaning, content, and envisioned structure of this new integrated approach to ocean management remains remarkably undefined and under-examined.\textsuperscript{280}

It certainly appears that IOM is at least concerned with the holistic management of the ocean as one large interdependent—and, indeed, transnational—ecosystem.\textsuperscript{281} By understanding the placement of humans as an integral aspect of this interconnected ocean network, many see IOM as almost a


\textsuperscript{275} Although various different phrases have been used over the years, often with different emphases, the term “Integrated Ocean Management” has perhaps become most widely adopted. See, e.g., Jakobsen, supra note 201, at 297–98.


\textsuperscript{277} See TANAKA, supra note 15, at 16.


\textsuperscript{279} Wright et al., supra note 208, at 9.

\textsuperscript{280} See TANAKA, supra note 15, at 17; Scott, supra note 4, at 466; Jakobsen, supra note 201, at 297.

This, in essence, demands that all governance be founded upon a highly relational (interdependent), eco-centric, and cumulative understanding of long-term impacts of marine activities upon all stakeholders. 283 Most also concur that IOM’s meaning is rarely fixed, but is actually highly context-dependent.284 It is, therefore, better understood as an amassing collection of principles and processes to be called upon depending on the specific context, challenge, or circumstance in focus.285 Karen Scott, for example, draws upon various environmental principles and marine planning processes being used in various regions and nations across the world.286 Similarly, Elizabeth Kirk argues that an approach is needed which “combines both principle and process.”287 A brief review of the literature, however, suggests that IOM is likely to include many of the general environmental principles, such as sustainable development, intergenerational and intragenerational equity, ecosystem services, common but differentiated responsibilities, access and benefit sharing, precautionary management, and the polluter pays principle—as well as containing other more general principles such as public participation, transparency, and accountability. In addition to these principles, numerous governance processes can be utilized in order to achieve better integration, such as marine spatial planning, marine protected areas, environmental impact assessments, integrated coastal zone management, co-management, and various collaborative models of governance.

Many authors also understand IOM’s purpose as ultimately bridging the specific “gaps” in ocean governance which continuously appear. Barnes’s approach, for example, classified IOM as providing normative, spatial, sectoral, temporal, disciplinary, and user integration.288 Similarly, the BBNJ Working Group viewed IOM as bridging regulation, implementation, governance, coordination, and information sharing gaps.289 Further, although Tanaka broke IOM down into ecological, normative, and implementation vectors, his definition narrowly focused on institutional or agency coordination and, thus, largely excluded the

282 See Jakobsen, supra note 201, at 293–96; see also Scott, supra note 4, at 465; TANAKA, supra note 15, at 18. See generally Kirk, supra note 274; Stephen Jay, Marine Space: Maneuvering Towards a Relational Understanding, 14 J. ENV’T POL’Y & PLAN. 81 (2012); cf. Oanta, supra note 16, at 226 (suggesting that IOM necessitates end-to-end regulation, in order to improve cohesion and harmony across the regulatory framework).

283 See supra note 15 and accompanying text.

284 See Scott, supra note 4, at 466–67.

285 See generally id.; Kirk, supra note 274.

286 See generally Scott, supra note 4.

287 Kirk, supra note 274, at 33.

288 See Barnes, supra note 278, at 860–62.

various gaps that arise when looking more closely at the vital stakeholder level.\textsuperscript{290} As demonstrated in Section III.D, however, most ocean governance gaps can, in fact, be clearly linked back to the international system of law. Moreover, studies into transnational law and governance are often seen as a means to improve upon these same weaknesses of the international legal system itself.\textsuperscript{291} Ergo, in order to address the gaps in ocean governance, a focus on achieving effective transnational law and governance is needed.

Nevertheless, despite some commentators casually noting the transnational characteristics of the oceans\textsuperscript{292} and calling for solutions that capture transboundary activity from end-to-end,\textsuperscript{293} there has been a dearth of literature in the law of the sea field making the crucial connection between transnational law, as a broad discipline and field of study, with the coveted principles and processes of IOM.\textsuperscript{294} This is surprising and regretful given that the same weaknesses of the international law of the sea that are universally lamented—such as the need for an expanded role of supranational, private, and non-state laws, actors, and systems operating more fluidly at transnational scales—are almost identical to those same weaknesses that the transnational legal discipline seeks to address. More so, given that, in its normative guise, transnational law seeks to achieve a multi-level, holistic, and inclusive system of law in response to global regulatory challenges. As a result, any proposed model of IOM should incorporate the theories, processes, and approaches of transnational law and governance.

B. A Transnational Law of the Sea: Multi-Stakeholder Inclusivity and Multi-Level Governance

At the very heart of the coveted “integrated” model of ocean governance is the consistent desire for both greater stakeholder inclusivity and for increased multi-level regime-building. In other words, the desire to truly bring the global community of persons beyond states—the unfortunately titled “non-state actors”—into positions of governance authority, as well as to release some of the grip on legal authority by national governments and displace it to governance networks found at local, national, regional, and global scales. In fact, this ideal of post-national

\textsuperscript{290} See TANAKA, supra note 15, at 18–21.

\textsuperscript{291} See supra Section II.C.


\textsuperscript{293} See Oanta, supra note 16, at 221, 226; Betrand Le Gallic, \textit{The Use of Trade Measures Against Illicit Fishing: Economic and Legal Considerations}, 64 ECOLOGICAL ECON. 858, 858 (2008). See generally High Seas Task Force, supra note 85.

\textsuperscript{294} For example, a standard and unfiltered Google search by the author in May 2020 for “transnational law of the sea” returns zero results; whereas “transnational human rights” returns 98,200 results, “transnational environmental law” returns 42,600 results, “transnational criminal law” returns 95,000 results, and even “transnational family law” returns 26,500 results.
governance accords neatly with Elisabeth Mann Borgese’s visionary 1998 monograph, *The Oceanic Circle: Governing the Seas as a Global Resource*, which promoted a new global approach to ocean governance.\(^{295}\) The same recognition could perhaps also be attributed to Allott’s thesis on the *Mare Nostrum* in 1992, which touched on the case for a human-centered and post-Westphalian approach to ocean governance—indeed, as well as Grotius’s *Mare Liberum* over four centuries ago.

1. Multi-stakeholder inclusivity

In every case, one can locate *stakeholder participation*—and the capacity of ocean users at all levels both within and without the state to effectively and productively communicate, cooperate, coordinate, and collaborate—at the heart of the IOM approach. For example, even under those headings that are discernibly regulatory in focus, such as the need for normative frameworks which cross political boundaries, the same underlying motive is to address the conflicting and overlapping interests of fragmented and dissociated communities.\(^{297}\) Barnes, therefore, rightly summarizes that a “truly integrated approach” would empower stakeholders to be engaged in the regulation process.\(^{298}\) It is possible to detect these same common themes throughout social scientific research that seeks to improve upon the presently beleaguered model of Westphalian law.\(^{299}\) Research and associated policies dedicated to expanding public participation have, therefore, proliferated in recent years, with the principle of civil-society inclusivity becoming something of a revolutionary movement.\(^{300}\)

Environmental public participation can range from small-scale local resource management to the representation of global communities and interest groups at

---

296 See Freestone et al., supra note 211, at 13. See generally Allott, supra note 207.
297 See supra note 211 and accompanying text.
298 See Barnes, supra note 278, at 862.
international fora.\footnote{See C. Aloni et al., The Importance of Stakeholders Involvement in Environmental Impact Assessment, 5 RES. 
& ENV’T 146, 148–49 (2015); Tamara Tschentscher, Promoting Sustainable Development Through More Effective Civil Society Participation in Environmental Governance: A Selection of Country Case Studies from the EU-NGOs Project 8 (2016).} It can also come in many forms: from a simple requirement to improve public information and transparency of decisionmakers to consultations with affected individuals and communities in the development of new legislation, to full-scale regulatory management of public goods by private actors.\footnote{See supra notes 299–301 and accompanying text.} There are also many representative forms of the global demos, such as: non-governmental organizations (NGOs), including charities, advocacy networks, and interest groups all operating at the national, regional, and global levels; corporations and other enterprises, also operating at or across the local, national, regional, and global scales; epistemic (or “expert”) communities and standards bodies, across all levels; and all communities and individual humans themselves.\footnote{See generally supra note 14 and accompanying text; Robert Pomeroy & Fanny Douvere, The Engagement of Stakeholders in the Marine Spatial Planning Process, 32 MARINE POL’Y 816 (2008); Jutta Brunnee & Ellen Hey, Transparency and International Environmental Institutions, in TRANSPARENCY IN INTERNATIONAL LAW 23 (Andrea Bianchi & Anne Peters eds., 2013).}

In all cases, there is clear evidence that an expanded demos has been effective in enhancing the protection of the ocean. There are numerous examples where the progress toward post-national representation in global governance is at last providing solutions to the ocean’s paradigmatic transnational context. Such examples include: the increased stakeholder participation in marine planning or environmental protection systems;\footnote{See generally Maarten Bavinck & Svein Jentoft, Subsidiarity as a Guiding Principle for Small-Scale Fisheries, in WORLD SMALL-SCALE FISHERIES: CONTEMPORARY VISIONS 311 (Ratana Chuenpagdee ed., 2011); Berkes, supra note 195.} enhanced subsidiarity of biodiversity management and protection;\footnote{See generally Torrens, supra note 214.} expansion of international legal standing to private actors and NGOs;\footnote{See generally David Lewis and Nazneen Kanji, Non-Governmental Organizations and Development (2009); Marloes Knaan et al., How to Dance? The Tango of Stakeholder Involvement in Marine Governance Research, 50 MARINE POL’Y 347 (2014); Robert E. Kelly, From International Relations to Global Governance Theory: Conceptualizing NGOs After the Rio Breakthrough of 1992, 3 J. CIV. SOC’Y 81 (2007).} integration of stakeholder groups into governance networks; public-private partnerships or advisory bodies;\footnote{See generally Josh B. Martin, Harnessing Local and Transnational Communities in the Global Protection of Underwater Cultural Heritage, in TRANSNATIONAL ENVIRONMENTAL LAW (forthcoming 2020); Steve Charnovitz, Two Centuries of Participation: NGOs and International Governance, 18 MICH. J. INT’L L. 183} utilization of private communities and NGOs within the governance mix;\footnote{See generally supra note 309 and accompanying text.} utilization of maritime
communities in improving security measures;\textsuperscript{309} and the facilitation of coordination and information exchange across epistemic communities\textsuperscript{310} or private and subnational actor networks.\textsuperscript{311}

2. Multi-level governance

It similarly becomes essential, therefore, that multi-level governance (MLG)—governance that is intentionally arranged as interlinking and overlapping regimes at different spatial levels or scales at the global, regional, national, and local community level—should provide both the descriptive and normative frame for understanding the coveted stakeholder-inclusive model. Not only does MLG provide an appealing vision for a simple response to the complex challenges of transnational governance and provide avenues toward greater stakeholder inclusivity, but it also helps clarify the roles and functions of the state and of non-state actors within the governance framework.\textsuperscript{312} In fact, a multi-stakeholder approach naturally calls for a multi-level approach, given that most actors beyond the state are characteristically found above or below the national level.\textsuperscript{313} It appears clear that such a multi-level approach is sought in the attainment of integrated governance. Indeed, the International Union on the Conservation of Nature Draft

\begin{flushleft}

\textsuperscript{310} Peter M. Haas, Epistemic Communities, Constructivism, and International Environmental Politics (2015).


\end{flushleft}
International Covenant,\textsuperscript{314} Rio Declaration,\textsuperscript{315} and Agenda 21\textsuperscript{316} each call unequivocally for action at all levels—international, regional, national, and local—in order to effectively achieve sustainable development. MLG thus becomes a necessary byproduct of the worldwide search for governance beyond government.

Indeed, the interdependence of all stakeholders in the shared ocean commons leads to a widespread tradition of externalities and spillover effects, resulting in underproduction by nation-states in work toward global objectives. MLG can therefore resolve these collective action failures by integrating states into regimes or processes that delegate decision making authority and accountability to more suitable levels, such as through local, regional, or global regimes, where international law is failing to produce global public goods.\textsuperscript{317} It also permits states to pass technical management and administrative responsibility for global goods onto more appropriate external agents and non-state actors,\textsuperscript{318} or even to pass on the blame for stringent social and environmental policies.\textsuperscript{319} Furthermore, the centralization of responsibility can actually enhance, rather than diminish, the capacity of civil society to press for changes to policies, by creating a clearer and more familiar target for political lobbying.\textsuperscript{320}

The above can be evidenced in the ocean governance context. An illustrative example is the sustainable management of global fish stocks. It is true that certain epistemic communities (for example, the International Council for the Exploration of the Sea or the International Union on the Conservation of

\textsuperscript{314} See generally IUCN ENVIRONMENTAL LAW PROGRAMME, DRAFT INTERNATIONAL COVENANT ON ENVIRONMENT AND DEVELOPMENT (5th ed., 2017).


\textsuperscript{319} See Wälti, supra note 317, at 414.

\textsuperscript{320} See id. at 418–19; Jeffrey L. Dunoff, Levels of Environmental Governance, in THE OXFORD HANDBOOK OF INTERNATIONAL ENVIRONMENTAL LAW 85, 97, supra note 166; Nele Matz-Lück & Johannes Fuchs, The Impact of OSPAR on Protected Area Management Beyond National Jurisdiction: Effective Regional Cooperation or a Network of Paper Parks?, 49 MARINE POL’Y 155, 163 (2014).
Nature), standards bodies (for example, the Marine Stewardship Council or the U.N. Food and Agricultural Organization), or advocacy NGOs (for example, Oceana, Greenpeace or the World Wildlife Fund) are more efficiently organized by providing data, technical rules, standards, schemes, and political pressure at a global level. However, it is also necessary to leave sufficient latitude to local communities to develop their own rules and systems for preventing or recycling bycatch or allocating fishing zones between local stakeholders with the efficient use of the guidance, tools, enforcement architecture, and resources which are provided from higher levels. Thus, higher levels can be used to overcome the accepted limitations of national enforcement, by encouraging or sidestepping state-level implementation.

In addition to having the global governance level bolstered by strategic public-private partnerships expanded participative democracy, it is particularly the regional level that has continuously proven itself to be indispensable within effective ocean governance models. Regional-level governance carries many advantages, by lowering the critical mass needed before significant commitments can be made in cross-border negotiations, as well as providing for a more harmonized and coherent system of management for a large body of water. As such, there are countless success stories in the protection of the marine environment driven by regional networks and organizations, such as through the Barcelona Convention, Antarctic Treaty System, Baltic Marine Environment

326 See supra notes 301–11 and accompanying text.
Protection Commission (or Helsinki Commission), the OSPAR Commission, and numerous other regimes, such as those under the U.N. Regional Seas Programme. There is also the enhanced level of protection and stock management derived by regional fisheries management organizations and global improvements in general flag state compliance as a result of closed regional systems of port state measures, such as the Paris Memorandum of Understanding. Furthermore, based on the analyses above, the more that states are willing to acquiesce national sovereignty across regional contexts, the more effective and powerful the resulting regimes can become. For example, the highly supranational European Union has made several noteworthy achievements in the marine governance context and is regarded as a world-leader in this field, including extensive improvements in pollution and waste management, fisheries, transboundary spatial planning, maritime security, and biodiversity protection. This makes sense when one considers that the creation of an overarching system of accountability above states will have the capacity to drive higher levels of state compliance and a wider servitude to regional objectives.

For all these reasons, it is understandable that the present negotiations over an international legally-binding instrument to protect biodiversity beyond national jurisdiction (BBNJ) have included discussions on a “hybrid” global-regional approach. An efficient system of MLG would therefore handle the complexity and polycentricity of different governance polities, with numerous public, private, and hybrid actors operating at different levels and playing different roles.

331 See OSPAR COMMISSION (May 6, 2020), https://perma.cc/FN6W-SWWC.
depending on the precise issue in question. To be effective, MLG must also use an appropriate mix of centralization and decentralization, to permit autonomous and meaningful community self-governance at the right levels.

MLG is thus understood as the vital provision of polycentric coordination within Transnational Environmental Governance, given this field’s concern “with the migration and impact of legal norms, rules and models across borders.” As Campbell and her colleagues suggested in 2016, an explanation for failed oceans governance is scalar mismatch; that is, a governance intervention . . . not well matched to the ecological scale of the feature or process being governed . . . . The concern for scalar mismatch fuels support for governance at global or regional scales and coordination among scales.

However, rather than viewing MLG as merely the realignment of scalar mismatch between eco-systems and governance systems, we must be clear that the marine environment continues to suffer severe scalar mismatch across normative, regulatory, jurisdictional, collaborative, and compliance systems as well.

VI. CONCLUSION: LOOKING BEYOND THE HORIZONTALISM

Ascending up the rigging to the crow’s nest atop our new globalized world order and peering through the spyglass, it is clear that coming into view beyond the horizon is a new paradigm of ocean governance. While sovereignty and territoriality may continue to rule the waves for some time, any suggestion that this system of global legal accountability is working for the protection of our shared global ocean must at last be jettisoned. Instead, it is time to consciously expand beyond the horizontalism and decentralism of the present inter-national law of the sea, by focusing on building more forceful systems of legal accountability above, below, within, and without the nation-state. The failure to govern the oceans has less to do with Grotius’s res communis or flag state regulation per se, as it does with our griped obsession with legal positivism and the eternal truism of sovereign exclusivity, equality, and territoriality. All three of these features of national sovereignty—which remain resolute and ever-present in our global legal order for the seas—have manifested a list of critical weaknesses in the management of an incredibly complex, multi-leveled, prototypically transnational, 


See Dunoff, supra note 320, at 88–100.


intergenerational, fluid, interdependent, and reflexive global commons, that itself remains accessible and open to rivalry from all directions.

There has been a distinct lack of criticism of our present international legal system for governing the seas, despite the huge expanse of research which has flourished in the fields of transnational law and governance, global legal pluralism, global governance, and multi-level governance. Fortunately, the time is ripe to adapt the principles, findings, and theories across these subject fields for use within the marine environment. The time for reflecting on our presently Westphalian “constitution for the oceans”—as epitomized by the 1982 LOSC—therefore appears to have been upon us for some time. While some law of the sea experts have dispassionately inferred that this may require a “new paradigm beyond positivism,” there has been a lack of conviction and concerted research towards this end. Perhaps the best effort thus far has been Tanaka’s 2008 monograph on a regionally oriented understanding of “integrated” management for the oceans that, although appearing to have slipped past the radar of wider academic discussions, did at last highlight the weaknesses of a sole reliance on the LOSC.

Yet, it is clear that ocean jurisprudence continues to proceed along the same tired inter-national lines. For example, the current negotiations over an internationally binding legal instrument on the protection of BBNJ appear to carry the same weaknesses of inter-state consent-based bargaining and lowest common denominator weaknesses which were lamented throughout this article. More investment is needed in the mobilization and advancement of private and hybrid actors, such as NGOs, transnational corporations, subnational actors, and standards bodies, as well as effective supranational and regional institutions, which are able to force meaningful consent and compliance across ocean space.

Far more can be discussed on the expansion of pluralistic, multi-stakeholder, and multi-level systems of maritime law and governance, as well as the means of achieving such a global system of regulation. Nevertheless, this article has focused an early critical point in these discussions, by arguing that the worldwide search among academics and experts over the past three decades for new “integrated” modes of ocean governance has, in fact, been merely a calling for a new post-Westphalian, multi-level, and stakeholder-inclusive governance regime. In other words, the coveted search for an integrated ocean governance is a search for a new transnational law of the sea. The principal feature of this new transnational governance model is an advanced level of stakeholder participation in the

342 See, e.g., Freestone et al., supra note 211.
343 See generally Allott, supra note 207; Barnes, supra note 102; Tanaka, supra note 15; Borgese, supra note 295.
344 See generally Tanaka, supra note 15.
governance framework, as well as the migration of normative regimes from a stubbornly horizontal system toward a multi-layered and polycentric network of legal rule-makers and takers.

The integration and coordination between the traditional inter-national order of the oceans, as epitomized and constitutionalized through the LOSC, with the emerging seaward migration of a new transnational understanding of global governance, is a ripe area for future research. The first port of call would be to finally identify our outmoded devotion to Westphalianism as being at the heart of our failing system up to now. If we are ever to govern this blue planet effectively, a transnational law of the sea is not just desirable—it is indispensable.
The Legal Man in the Moon: Exploring Environmental Personhood for Celestial Bodies

William B. Altabef*

Abstract

The rise of the commercial space industry endangers the preservation of environments, such as the lunar surface and other celestial bodies, with the threat of contamination and resource exploitation. In the coming decades, flights to space will become commonplace—but at present, there is no way to hold outer space polluters accountable. The existing international legal regime is weak, with the United Nations’ space treaties offering limited enforcement mechanisms against offenders. The increasingly popular concept of environmental personhood offers a solution by rethinking the meaning of a juridical person within the text of the United Nations Outer Space, Space Liability, and Moon treaties. Utilizing the International Court of Justice, outer space environmentalists can seek to recognize celestial bodies as juridical persons and gain third-party standing to protect the rights of the Moon and seek damages for environmental degradation. Through the exploration of contentious and advisory avenues within the International Court of Justice, this Comment advances a new way of thinking to save extraterrestrial environments.

* The University of Chicago Law School, J.D. Candidate, 2021. I would like to extend my gratitude to Professor Aziz Huq for his guidance and to the CJIL Board and Staff for their support. I would also like to thank my family, Veronica Patterson, and Conor McGannon for their encouragement in pursuing this topic.
# Table of Contents

I. Introduction......................................................................................................................... 478

II. The Problem ...................................................................................................................... 479
   A. The Practical Problems ................................................................................................. 480
      1. Exploitation .............................................................................................................. 480
      2. Contamination ......................................................................................................... 481
   B. The Legal Problems ..................................................................................................... 484
      1. Lacking Regulatory Regime ..................................................................................... 484
      2. Unarticulated Basis for International Litigation ....................................................... 487

III. Legal Background .......................................................................................................... 487
   A. Existing International Law for Celestial Bodies ......................................................... 487
      1. The Outer Space Treaty .......................................................................................... 488
      2. The Space Liability Convention ............................................................................. 489
      3. The Moon Treaty ..................................................................................................... 491
      4. Interpreting the U.N. Treaties ................................................................................ 493
   B. Environmental Personhood .......................................................................................... 494
      1. Introducing the Concept .......................................................................................... 494
      2. Subsequent International Adoption ........................................................................ 495

IV. Recognizing Environmental Personhood for Celestial Bodies as a Solution ............. 498
   A. Jurisdiction of the International Court of Justice ...................................................... 499
   B. Bringing a Contentious Claim ..................................................................................... 500
      1. Bringing the Claim .................................................................................................... 501
      2. Deciding a Case Explicitly Under Environmental Personhood .............................. 501
      3. Deciding a Case Using a Common Law Approach ................................................. 502
      4. Prospect of Success ................................................................................................. 504
   C. Seeking an Advisory Opinion ...................................................................................... 505
      1. Requesting an Advisory Opinion ........................................................................... 506
      2. Advantages and Relative Value of an Advisory Opinion ...................................... 507
   D. Shortcomings ................................................................................................................ 507
      1. National Opposition .................................................................................................. 508
      2. Institutional Capacity ............................................................................................... 508
      3. Interpreting the Treaties’ Gaps as Enabling Mining ............................................... 509
   E. Alternative Solutions ..................................................................................................... 511
      1. Establishing a Moon Authority ................................................................................ 511
      2. Permanent Court of Arbitration ............................................................................. 511

V. Conclusion ....................................................................................................................... 512
I. INTRODUCTION

Humanity is on the eve of a new era of industry, during which natural resources, manufacturing, and research will be conducted beyond terrestrial boundaries. This will include the mining of asteroids and the transformation or colonization of untouched celestial bodies for economic gain by human actors.\(^1\) Presently, there is no effective way to enforce environmental regulation on celestial bodies, leaving them vulnerable to contamination and exploitation without repercussion.

This is an important issue as the vast mineral riches of outer space could fundamentally alter Earth’s economy and existing geopolitical rivalries, extending them to outer space like during the Cold War. While it is unknown when and where the first human settlement on another celestial body will occur, the rising number of national and commercial actors interested in achieving such a goal make it seem probable.\(^2\) The success of the multinational effort to maintain a human population on the International Space Station for nearly two decades demonstrates the feasibility of this vision.\(^3\) Space entrepreneur Jeff Bezos sees the utilization of space as critical to the resolution of global issues including hunger, poverty, and pollution.\(^4\) These hopes will all be moot if contamination or rapid exploitation of celestial bodies renders them unable to be used for a greater good.

From a pessimistic perspective, a failure to introduce a regime for regulating the usage of celestial resources could lead to international conflict in space and on Earth.\(^5\)

In order to understand the scope of the problem, Parts A and B of Section II of this Comment examine the threat posed by the contamination and exploitation of outer space resources. Part C of Section II discusses how the current international regime is lacking solutions and is unlikely to create a legislative solution given the national incentives to develop space industries. Part


D of Section II frames the Comment’s focus: solving the problem of standing for international litigation over extraterrestrial environmental damage. Section III Part A provides a more substantive legal background, first by detailing the relevant U.N. treaties: the Outer Space Treaty, the Space Liability Convention, and the Moon Treaty. Section III Part B introduces the legal concept of environmental personhood, an idea quickly gaining traction as national and local governments seek to preserve natural resources on Earth and protect against climate change. Environmental personhood bestows juridical personhood upon natural features, enabling them to have standing so that other entities or persons can bring claims on their behalf. Finally, Section IV hypothesizes the application of environmental personhood in the realm of outer space and how the existing legal framework can provide a system of regulation and justice for celestial natural resources. The proposed means of incorporating environmental personhood into the international law of space would be a judgment by the International Court of Justice (ICJ), either through an advisory opinion or a contentious case. While there are potential problems and alternative solutions, implementing environmental personhood through a judicial decision represents a rapid solution requiring limited consensus to a problem that could quickly grow beyond control.

II. THE PROBLEM

The introduction of new national space programs and commercial ventures has ensured that space will become even more crowded in the coming years. Activity will not be limited to mere scientific exploration. Already, commercial actors are planning to mine celestial bodies for profit. While the near-term future for spacefaring consumers appears to be tourism, the goal of national space programs and space entrepreneurs is resource extraction, lunar colonization, and beyond. The Moon is proposed to be an abundant source of Helium-3, a few

---

7 See Higginbotham, supra note 1.
hundred tons of which could meet Earth’s energy needs for an entire year.\textsuperscript{11} There is an incentive in the space mining and colonization industries to build a track record of success to lure investment in a capital-intensive field, making the nearby Moon’s energy reserves an appealing target for early commercial missions. The private space race will potentially create the first trillionaire, leaving little room for the thought of the environmental effects on faraway places.\textsuperscript{12}

First, this Section discusses the two main environmental threats associated with new commercial enterprises: exploitation and contamination. Then, this Section explains how the existing legal framework compounds the practical problems, leaving a gap for the articulation and measurement of the environmental harm in the commons of outer space.

A. The Practical Problems

Without actions causing direct harms, there is no need for a legal framework to regulate activity and hold bad actors accountable. In order to understand the problem that the legal solution of extraterrestrial environmental personhood is attempting to solve, first this Comment will introduce the practical problems of exploitation and contamination.

1. Exploitation

The risk of exploitation is exponentially rising as the prospect of harnessing resources in outer space becomes commercially viable. Exploitation in this context can be defined as the extraction and consumption of extraterrestrial resources for non-scientific purposes, potentially without research on the long-term impact of such activities. The global space economy currently produces revenues of $350 billion, a number conservatively expected to rise to $1 trillion by 2040.\textsuperscript{13} Traditional aerospace companies such as Boeing and Airbus continue to focus primarily on designing rockets for national program usage.\textsuperscript{14} Newer entrants such as Planetary Resources are explicitly focusing on the private exploitation of


\textsuperscript{12} Tiffany Terrell, \textit{Physicist Says Asteroid Mining Ventures Will Spawn First Trillionaire}, GLOBAL NEWswire (Jan. 30, 2018), https://perma.cc/J75M-87NA.

\textsuperscript{13} Foust, \textit{ supra note 8}; \textit{Space: Investing in the Final Frontier}, MORGAN STANLEY (July 2, 2019), https://perma.cc/9FUJ-432V.

asteroids through space mining, while others, such as SpaceX and Blue Origin, have taken an ‘all of the above’ approach with broad goals to commercialize space and support the eventual colonization of outer space.

Commercial actors may be less likely to concern themselves with implementing procedures to mitigate or prevent pollution, and they could perform launch operations from jurisdictions with minimal requirements in order to maximize profitability. 

Previous missions to space were conducted for scientific purposes by governments, with American and Soviet space programs implementing planetary protection precautions for landers to prevent forward contamination.

The growing commercial industry raises the potential for a much more crowded outer space with less commitment to protocol. Already, commercial space ventures are launching thousands of satellites, endangering the low Earth orbit ecosystem, with little planning for the safe decommission of these satellites. While sustainable usage and extraction of extraterrestrial resources would be ideal, the risky nature of the space industry coupled with limited current enforcement is likely to produce a tragedy of the commons.

2. Contamination

Human exploration has a long history of contamination, and space is no exception. For the purposes of this Comment, contamination specifically refers to the introduction of foreign substances and lifeforms into extraterrestrial

---

17 See Caroline Delber, SpaceX Says There Are No Laws on Mars, So Maybe Elon Musk Will Be President, POPULAR MECHANICS (Oct. 30, 2020), https://perma.cc/338Z-TY6K (discussing SpaceX’s terms of service which claim Mars is a free planet and that no Earth-based government has authority over Mars).
20 See Jonathan O’Callaghan, The FCC’s Approval of SpaceX’s Starlink Mega Constellation May Have Been Unlawful, SCIENTIFIC AMERICAN (Jan. 16, 2020), https://perma.cc/GVZ7-J877 (detailing the launch of SpaceX’s Starlink satellite system around the Earth has already begun to brighten the sky, inhibiting terrestrial stargazing).
environments. Recent events further evidence the risks of contamination in an unregulated field. In February 2019, a private rocket carrying a lunar lander crashed on the surface of the Moon.\textsuperscript{22} The rocket was funded by the American nonprofit Arch Mission Foundation and was launched from Florida by the Israeli corporation SpaceIL.\textsuperscript{23} Unbeknownst to international regulators at the time of launch, the rocket was carrying thousands of tardigrades, a terrestrial creature known for its ability to survive nearly anywhere.\textsuperscript{24} Nova Spivack, the cofounder of the Arch Mission Foundation, admitted to placing the tardigrades on the SpaceIL lander at the last minute without disclosing the nature of the addition to SpaceIL.\textsuperscript{25} Although there is no definitive analysis of the consequences of introducing the tardigrades into the lunar environment, they are the only creature known to survive the vacuum of space.\textsuperscript{26} There is limited immediate threat posed by the tardigrades as they exist in a state of cryptobiosis in space, unable to reproduce with their metabolism held to a minimum, but the concern is that the next species sent to the Moon might not be as harmless.\textsuperscript{27}

Regardless, the actions of the Arch Mission Foundation violate existing planetary protection guidelines, practices set out by the international Committee on Space Research (COSPAR) and national space agencies to prevent cross contamination between planetary bodies.\textsuperscript{28} There has been no reported sanction for the irresponsible private actors involved in the tardigrade launch, signaling a low risk to subsequent commercial actors and increasing the chance of future contamination or exploitation. National systems are disincentivized from


\textsuperscript{23} \textit{Id.}


\textsuperscript{25} Chris Taylor, \textit{I'm the First Space Pirate!} How Tardigrades Were Secretly Smuggled to the Moon, \textit{Mashable} (Aug. 8, 2019), https://perma.cc/98W5-JTVY (noting that Spivack considers himself to be the first space pirate after his smuggling of the tardigrades).


\textsuperscript{27} Ari Shapiro, \textit{Thousands of Tardigrades are Stranded on the Moon After a Failed Lunar Mission}, NPR (Aug. 8, 2019), https://perma.cc/L4GB-J4CD.

\textsuperscript{28} COSPAR, \textit{The COSPAR Panel on Planetary Protection Role, Structure and Activities}, \textit{205 Space Rsch. Today} 14 (Aug. 2019) (providing an overview of the planetary protection framework and examples of procedures, including the requirement that missions to other planetary bodies “adhere to stringent planetary protection measures to abide the first rationale for planetary protection to not interfere with ‘scientific investigations of possible extraterrestrial life forms, precursors, and remnants’ and not to impose terrestrial biological contamination to these objects of high astrobiological interest”).
regulating harshly, as launches can happen across the globe and fledgling space companies might take their businesses to friendlier jurisdictions.

While the Earth’s atmosphere has proven to be relatively durable in the face of carbon emissions and other pollutants, the atmospheres of our neighbors are far more fragile. The emissions of twenty Apollo mission landings would have effectively doubled the lunar atmosphere; the Martian atmosphere is similarly tenuous. When imagining the scale of a lunar mining operation or colony, it is easy to predict the potential human-caused climate change on the lunar surface. Introducing enough biological or chemical contaminants could produce carbon emissions that start a dangerous process.

Various actors have proposed larger environmental offenses, particularly terraforming, the process by which an Earthlike ecosystem is created on another planet. Already, small steps have been taken to test our ability to bring Earth to other surfaces, such as when China attempted to grow cotton on the Moon. There is concern that an effort to terraform will damage the natural ecosystem of the targeted planet. Elon Musk, CEO of SpaceX, has suggested terraforming Mars and potentially nuking ice deposits on its surface. Musk’s plan to bring life to Mars through terraforming has numerous critics and doubters in the scientific community, particularly astrobiologists. It could be that nuking Mars leads to the degradation of natural features that humanity might wish to preserve for future generations. Just as Americans have protected Yosemite and the Grand Canyon, perhaps future Martian settlers will wish that Olympus Mons had been protected.

Some amount of extraterrestrial resource usage is permissible to satisfy human needs for research and scientific gains, but there should be a contamination threshold beyond which there is some sort of legal ramification. The contamination threshold could be determined by considering the value of the contaminated body and the severity of the contamination. The value of the extraterrestrial body can be governed by the categories of planetary protection priority already established by COSPAR based upon the probability that those bodies have life on them. For example, the contamination threshold might be

---


33 See COSPAR, supra note 28.
higher for a barren asteroid and lower for Europa, which is believed to have a subsurface ocean. There would be variation by the type of contaminant and the amount of exposure. Leaving sealed bags of human waste on the Moon from the Apollo missions might not be as harmful as releasing a school of fish into an extraterrestrial ocean.

B. The Legal Problems

The potential damage from exploitation and contamination could be mitigated if there was a sufficient regulatory regime or enforcement mechanism to curb the activities of future polluters. As outlined below, the current national and international regulatory regimes fail to control independent actors seeking a profit. Although it may be possible to craft regulations capable of protecting extraterrestrial environments through permitting and planetary protection systems, political capital requirements and national incentives make regulation a less likely solution. Instead, a few environmentally conscious actors can seek judgments through international litigation, lowering the potential political costs and providing a more immediate solution than the drafting of regulations. The issue that this Comment focuses on is how to cure potential defects in standing and causation in potential international litigation over environmental damage to an extraterrestrial environment, such as the Moon.

1. Lacking Regulatory Regime

Despite decades of increasing usage and dependency on outer space as a resource, the international regime governing outer space is weak. The United Nations Committee on the Peaceful Uses of Outer Spaces (COPUOS) oversees the United Nations Office for Outer Space Affairs (UNOOSA) and created the five current U.N. treaties covering outer space.\textsuperscript{34} While the first space treaty from 1967 has 109 states parties,\textsuperscript{35} subsequent treaties offering more specific regulation of space received much less support, with the Moon Treaty having only 18 states parties.\textsuperscript{36} The five main space treaties were all introduced between 1967 and 1979, with no substantive development of an international regulatory regime in the subsequent years. This gap in regulation has increased along with the possibilities and realities of human use of space.

Many nations have supplemented the international agreements with their own space regulations. In the United States, the Federal Aviation Administration

\textsuperscript{34} Roles and Responsibilities, U.N. Office for Outer Space Affairs, https://perma.cc/536N-7FQU.


\textsuperscript{36} Agreement Governing the Activities of States on the Moon and Other Celestial Bodies, Dec. 5, 1979, 1363 U.N.T.S. 3 [hereinafter Moon Treaty].
oversees the space launch licensing process, and maritime jurisdiction extends to spaceships. In the European Union, Norway, Sweden, Belgium, the Netherlands, and France have all passed laws to regulate private space enterprises. These laws, among many other examples at the national level, often require private actors to secure permission or comply with a national registry before launching objects into space. While these national regulations impose some restrictions on commercial actors, the international regime never reached a sufficient level of development to do so.

The development of independent national laws is not necessarily beneficial to the protection of celestial environments. Given the vast amount of resources and money at stake, it may be more likely that national legislation leads to a race to the bottom to enable domestic space corporations to engage in riskier but more profitable activities than their international competitors. Increasing values of asteroid minerals in combination with lower barriers to entry as space technology improves will encourage more commercial players to enter the industry. Without an international regulatory regime in place, commercial actors will be incentivized to lobby against regulations as revenues increase and they gain more influence within national governments. National governments will also have limited incentives to regulate their own space industries if it will hurt their competitiveness in the broader market. Even if national governments were willing to create and enforce a working environmental system, the lack of uniformity between national standards still calls for an international approach.

Seabed mining provides a terrestrial example of this problem. After initial proposals to collect minerals from the sea floor developed in the 1960s, the U.N.

---

40 Lov om oppskyting av gjenstander fra norsk territorium m.m. ut i verdensrommet, 13 juni 1969 nr. 38 (Nor.).
41 2 § LAG OM RYMDVERKSAMHET, (Svensk f-författningssamling [SFS] 1982:963) (Swed.).
42 Loi relative aux activités de lancement, d’opération de vol ou de guidage d’objets spatiaux of Sept.17, 2005, MONITEUR BELGE [M.B.] [Official Gazette of Belgium], Nov.4, 2008 (Belg.).
43 Wet rimtevaartactiviteiten, 24 januari 2007, Stb. 2007, 80 (Neth.).
45 See Tronchetti, supra note 11, at 810.
enacted the United Nations Convention on the Law of the Sea (UNCLOS) in 1982.\textsuperscript{47} UNCLOS established an international framework for several international maritime legal issues, including the establishment of the International Seabed Authority (ISA) under Part XI of the convention. Article 136 declares, “The [seabed] Area and its resources are the common heritage of mankind.”\textsuperscript{48} The ‘common heritage of mankind’ language mirrors language that was used to describe the Moon in the Moon Treaty, which was being drafted contemporaneously.\textsuperscript{49} The result of UNCLOS and the ISA has been far from ideal. The ISA has operated with limited transparency,\textsuperscript{50} giving out twenty-seven contracts for the mining of 1.4 million square kilometers with limited assurances of the environmental controls desired by some conservationists.\textsuperscript{51} A small number of contracts for mining in international waters come from closed sessions of the ISA, while many larger contracts are given by national governments to mine the seabed within their exclusive economic zones with limited research on the ultimate environmental impact.\textsuperscript{52} The U.S. notably objected to parts of UNCLOS, undermining its effectiveness and leading to competing national regulatory systems.\textsuperscript{53} It is not unreasonable to imagine the same dual licensing system taking hold in space, whereby some corporations are licensed by national launch authorities and others by an international body. The danger is much the same, that corporations will seek the nations willing to give early licenses in order to get ahead in space.

The interaction of actors incentivized to be the first to make a large profit in a risky industry with the lack of real regulation may lead to the contamination or exploitation of celestial resources with potentially irreversible consequences. A historical analogy would be that of the older oil wells in Texas, drilled without long-term concern for environmental impact, which are now leaking contaminants across the state.\textsuperscript{54} When starting a risky natural resources venture,

\textsuperscript{48} Id. art. 136.
\textsuperscript{50} Kirsten F. Thompson et al., Seabed Mining and Approaches to Governance of the Deep Seabed, Frontiers Marine Science (Dec. 11, 2018), \url{https://perma.cc/8BGH-C78E},
\textsuperscript{52} Id. (detailing contracts given to mine off the coast of Africa and Oceania).
\textsuperscript{54} Jim Malewitz, Abandoned Texas Oil Wells Seen as “Ticking Time Bombs” of Contamination, Texas Tribune (Dec. 21, 2016), \url{https://perma.cc/6E2W-UGQM}.
the first goal is to make a substantial return, and the fear of environmental liability is often an afterthought. If a wildcatter does not find oil or cannot get it out of the ground, they will be just as insolvent as if they were hit with large amounts of environmental liability. Additionally, there is a short-term bias to conducting many ventures as the environmental impact and full consequences often are not calculable until decades later, potentially long after the initial mining venture has concluded. History cautions against the lack of regulation. A legal regime is needed to hold offending actors accountable in a time horizon short enough to create an incentivizing impact.

2. Unarticulated Basis for International Litigation

In the absence of either a working international regulatory framework or comparable national systems, pursuing environmental damages claims under the existing U.N. space treaties in the ICJ presents a viable path to create accountability and promote extraterrestrial conservation. The language of the treaties, detailed in the following Section, enables a nation to seek monetary damages when space debris from a second nation strikes the territory or property of the first nation. In the extraterrestrial context, the difficulty arises when the damage occurs to another planetary body, which no nation has a territorial claim to protect. There is a gap in the current practice and scholarship on international law to show how a litigant could have standing to sue for damage to extraterrestrial environments, such as that of the Moon. This Comment will focus on answering this problem by using the legal concept of environmental personhood to articulate what is damaged when the Moon is polluted and how third-party standing will hold bad actors accountable.

III. LEGAL BACKGROUND

A. Existing International Law for Celestial Bodies

The U.N. is the primary governing authority on international laws and regulations pertaining to outer space. Through the U.N. Office for Outer Space Affairs, the U.N. tracks satellites orbiting the Earth and works to implement the five adopted space treaties. The two treaties most relevant to this Comment are the Outer Space Treaty and the Moon Treaty. The Outer Space Treaty laid out an initial framework for international ambitions to regulate outer space activity, but it left gaps and ambiguities for subsequent treaties and regulations to fill in or refine. The Outer Space Treaty’s general spirit provides a lodestar for subsequent laws to follow. The Space Liability Convention, for example, built upon the Outer Space Treaty to create a mechanism for nations to seek damages when debris fall into their sovereign territory, an important building block for nations seeking to

55 Id.
protect extraterrestrial environments from contamination or exploitation. Finally, the Moon Treaty introduced additional protections for the Moon and represented the most progressive attempt to prevent exploitative usage of extraterrestrial resources.

While the Moon Treaty was less widely adopted, the Outer Space Treaty and Space Liability Convention can work in concert to provide a path for a case to be heard in the ICJ if the court were to adopt an environmental personhood reading of certain provisions of the treaties. The Moon Treaty still provides a persuasive example of where the international legal community might have gone had competitive intentions been removed.

1. The Outer Space Treaty

The Treaty on Principles Governing Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies (Outer Space Treaty) was the first attempt to establish an international regime for outer space, ratified just 10 years after Sputnik and two years before the Apollo 11 landing.\(^{56}\) It has been ratified by 110 parties, including all major spacefaring nations, and serves as the most widely adopted source of international space law. Certain sections were later clarified by subsequent agreements such as the Agreement on the Rescue of Astronauts, the Return of Astronauts and the Return of Objects Launched into Outer Space.\(^{57}\)

The preamble of the Outer Space Treaty recognizes “the common interest of all mankind in the progress of the exploration and use of outer space for peaceful purposes.”\(^{58}\) This language demonstrates the reliance on historic conceptions of nature serving human needs, an idea that will be countered by the concept of environmental personhood and subsequent U.N. treaty language. In the absence of any foundational law for outer space, it may have seemed natural to transplant the legal regime that governed the property of Earth to outer space.

Article I of the Outer Space Treaty establishes the broad jurisdiction of the treaty as “[o]uter space, including the Moon and other celestial bodies.”\(^{59}\) Article II declares “[o]uter space, including the [M]oon and other celestial bodies, is not subject to national appropriation by claim of sovereignty, by means of use or occupation, or by any other means.”\(^{60}\) However, some commentators have

\(^{56}\) Outer Space Treaty, supra note 35.


\(^{58}\) Outer Space Treaty, supra note 35.

\(^{59}\) Id. art. I.

\(^{60}\) Id. art. II.
speculated whether resources are similarly unclaimable once they have been extracted.61

Under Article IX, “States Parties to the Treaty shall pursue studies of outer space, including the [M]oon and other celestial bodies, and conduct exploration of them so as to avoid their harmful contamination . . .”62 The introduction of tardigrades is clearly a form of contamination, but the harm caused by their introduction is still unknown. Previous ventures to the Moon during the Apollo missions left bags of human waste, which some scientists have hypothesized could have introduced microbial life to the previously sterile lunar surface.63 Terraforming a planet would almost certainly be a harmful contamination as it would purposefully change the ecosystem of a planet. The harm would be to the natural planet itself, rather than to the human interests that are traditionally at the core of human-centric Anglo-American property law. The environmental personhood concept requires that the protection from harm given to the celestial body be the same protection the law would give any human body. Even using the most stringent and narrow definitions of harm, requiring potentially irreversible decimation of any native lifeforms and the introduction of invasive species from Earth as proposed by terraforming advocates would meet these criteria. Yet, the current international framework has limited means to punish violations of this treaty.

2. The Space Liability Convention

Following the Outer Space Treaty, the U.N. enacted the Convention on International Liability for Damage Caused by Space Objects (Space Liability Convention) in 1972.64 The Space Liability Convention clarified language regarding the consequences of damages derived from space travel. While the focus of the treaty appears to be on compensation for damages stemming from an object falling from space, much of the treaty is still operable for “damage being caused elsewhere than on the surface of the earth.”65 Under Article I, damages can occur to “persons, natural or juridical . . .”66 The inclusion of juridical persons does not appear to be limited to persons who are a citizen of a state party. Rather the injury of any juridical person could be sufficient for a claim should the litigant have standing, enabling monetary damages for harm to juridical persons under the

---

62 Outer Space Treaty, supra note 35, art. IX.
63 Andrew C. Schuerger, John E. Moores, David J. Smith & Günther Reitz, A Lunar Microbial Survival Model for Predicting the Forward Contamination of the Moon, 19 ASTROBIOLOGY 730, 752 (2019).
65 Id. art. II.
66 Id. art. I.
environmental personhood model. Liability extends to “a State from whose territory or facility a space object is launched.” Thus, had the SpacEL rocket caused damage, the U.S. would be the liable nation as the rocket was initially launched from Cape Canaveral, Florida.

The Space Liability Convention suggests that diplomatic negotiations should be the primary means of settling damages claims. After stating a claim, a party “may also present its claim to the Secretary-General of the United Nations” under Article IX. Articles XIV–XX outline the use of an international Commission to resolve disputes between the parties when negotiations fail. The Commission is listed as the tertiary option, less preferable to negotiation or the assistance of the Secretary-General. There has only been one resolution of a claim thus far under the Space Liability Convention. In 1978, a Soviet nuclear satellite scattered radioactive debris over Canada, leading Canada to claim several million dollars in damages under the Space Liability Convention. The claim was then settled through diplomatic means before the point at which the Secretary-General would have made a recommendation. While this system worked for a state-sponsored satellite, it has tremendous shortcomings in dealing with the coming private space industry. In particular, the Space Liability Convention ostensibly requires the claims of private parties to be sponsored by a state party to the treaty.

Although not discussed in the text of the Space Liability Convention, the consultation of the ICJ appears to be a possible avenue under a recommendation by the Secretary-General. With no precedent demonstrating the mechanics of the Space Liability Convention at a more contentious or substantive procedural point, the Secretary-General would likely look to proven dispute resolution mechanisms. The ICJ would be at the top of the list given its proven record of equitable arbitration, statutory status under the U.N., and reputation as the premier international court. Demonstrating its institutional capacity, the ICJ has decided claims over other extraterritorial areas, including a claim over whaling in Antarctic...
water under the International Convention for the Regulation of Whaling.\textsuperscript{75} Hearing a case under the Space Liability Convention could become a routine practice for the ICJ in the coming decades as space traffic increases. An environmental personhood claim might eventually be heard as the ICJ produces substantially more precedent proving institutional expertise in extraterrestrial legal matters.

3. The Moon Treaty

The Agreement Governing the Activities of States on the Moon and Other Celestial Bodies (Moon Treaty)\textsuperscript{76} was the final of the five U.N. space treaties. Since its introduction in 1979, the Moon Treaty has been significantly less adopted than the Outer Space Treaty, with only 18 states parties. While the Moon Treaty is the most progressive on environmental issues, the lack of widespread adoption “renders the instrument practically meaningless.”\textsuperscript{77} India is the only Moon Treaty signatory to also have a significant national space program thus far.\textsuperscript{78} France is an original signatory to the Moon Treaty and is also a member state of the European Space Agency, potentially tying one of the largest space agencies to the Moon Treaty through a key member. Other nations continue to slowly join the Moon Treaty with accessions by Turkey and Saudi Arabia in 2012, Venezuela in 2016, and Armenia in 2018. The slow pace of adoption has left a “vacuum” of international law over the Moon,\textsuperscript{79} but the Moon Treaty still serves an important role as the best expression of the international objectives for a legal framework to govern the Moon.

The Moon Treaty’s status as the only treaty explicitly about the Moon should still guide behavior and inform any future discussions over the law governing the Moon. Most of the opposition to the Moon Treaty by the spacefaring powers at the time, the U.S. and the Union of Soviet Socialist Republics, was based upon reaction to the potential requirement to share extracted mineral wealth with other nations.\textsuperscript{80} This Comment does not address the legality of lunar mining itself, instead focusing on the legal ramifications for the environmental impact of such activities. The Outer Space Treaty set principles broadly for outer space activity;

\begin{flushleft}
\textsuperscript{76} Moon Treaty, supra note 36.
\textsuperscript{77} Lotta Viikari, \textit{Environmental Aspects of Space Activities, in Handbook of Space Law}, 717, 726 (Frans von der Dunk and Fabio Tronchetti eds., 2015).
\textsuperscript{78} Chandrayaan-2 Days Away from Moon’s Orbit. What Next, supra note 6.
\end{flushleft}
whereas subsequent treaties, such as the Moon Treaty, sought to fill in the gaps.\textsuperscript{81} The Space Liability Convention created the deeper legal framework initially called for in Article VII of the Outer Space Treaty covering liability for space activities. The Agreement on the Rescue of Astronauts, the Return of Astronauts and the Return of Objects Launched into Outer Space\textsuperscript{82} built out the vision set forth in Article V of the Outer Space Treaty. So, while the Moon Treaty received fewer initial ratifications, it can serve as an explanatory document for the framework that the signatories of the Outer Space Treaty had envisioned.

Article 1 of the Moon Treaty states that the treaty applies to the Moon and also “to other celestial bodies within the solar system other than the earth.”\textsuperscript{83} The Moon Treaty recognizes that outer space law is not separate from the international law framework but rather a subset of it, as “[a]ll activities on the moon, including its exploration and use, shall be carried out in accordance with international law, in particular the Charter of the United Nations . . . .”\textsuperscript{84} Under this conception, the broader international legal mechanisms, such as the ICJ and the Security Council, still govern the operations of space actors.

Later sections of the Moon Treaty deal more directly with the importance of environmental preservation. Article 7 § 1 states that “[i]n exploring and using the [M]oon, States Parties shall take measures to prevent the disruption of the existing balance of its environment, whether by introducing adverse changes in that environment, by its harmful contamination through the introduction of extra-environmental matter or otherwise.”\textsuperscript{85} This represents a more specific and stronger version of the language seen in Article IX of the Outer Space Treaty of 1967. The examples of terraforming and the introduction of tardigrades would likely be considered violations of this language. Under Article 11 § 1, “[t]he moon and its natural resources are the common heritage of mankind . . . .”\textsuperscript{86} Unlike the “province of all mankind” language from the Outer Space Treaty, the phrase “common heritage of mankind” is stronger and points to the Moon’s unique

\textsuperscript{81} Each of the preambles to the various space treaties refers to the prior agreements and the spirit of law the new treaty seeks to build upon. See Moon Treaty, supra note 35, ¶ 15 (“Recalling the Treaty on the Principles Governing the Activities in the Exploration and Use of Outer Spaces including the Moon and Other Celestial Bodies, the Agreement on the Rescue of Astronauts, the Return of Astronauts and the Return of Objects Launched into Outer Space, the Convention on International Liability for Damage Caused by Space Objects, and the Convention on Registration of Objects Launched into Outer Space . . . .”).

\textsuperscript{82} Rescue Agreement, supra note 57.

\textsuperscript{83} Moon Treaty, supra note 36, art. 1.

\textsuperscript{84} Id. art. 2.

\textsuperscript{85} Id. art 7.

\textsuperscript{86} Id. art. 11 § 1.
freedom from ownership. This language reinforces the human-centric attitude toward natural resources, but also indicates a common interest in the regulation or sustainable usage of the Moon. By recognizing a common heritage value of the Moon to all humans, the Moon Treaty might create an avenue for a plaintiff to bring a claim for the degradation of the lunar environment. The common heritage aspect of the Moon can enable third parties to serve as guardian ad litem for the environmental person that is the Moon.

Article 11 § 3 states that “[n]either the surface nor the subsurface of the [M]oon, nor any part thereof or natural resources in place, shall become property of any State . . . or non-governmental entity or of any natural person.” Based on this language, it seems that there should be no ownership or territorial claims to the Moon, creating a situation analogous to Antarctica. Additionally, it would be hard to legally extract and sell material from the Moon if it cannot be owned. Prohibiting ownership may prevent transfer or encourage actors to adopt a parallel system more supportive of their commercial needs.

Article 11 § 5 requires that “States Parties to this Agreement hereby undertake to establish an international regime, including appropriate procedures, to govern the exploitation of the natural resources of the Moon as such exploitation is about to become feasible.” Article 11 § 7 states that “[t]he main purposes of the international regime to be established shall include: (a) The orderly and safe development of the natural resources of the Moon; (b) The rational management of those resources; (c) The expansion of opportunities in the use of those resources . . .” It is conceivable that these responsibilities could be delegated to UNOOSA or COPUOS, but no regime or delegation of authority exists yet to carry out the Moon Treaty’s vision. Article 15 §§ 2–3 set out methods for resolving disputes, including consultation with the offending nation, the assistance of the Secretary-General, or “other peaceful means of [the states parties’] choice appropriate to the circumstances and the nature of the dispute.”

4. Interpreting the U.N. Treaties

While the U.N. outer space treaties provide a legal framework, judicial interpretation of the treaties’ terms can enable greater reach and regulation. In the

---

88 Moon Treaty, supra note 36, art. 11 § 3.
89 Juan Francisco Salazar, Antarctica and Outer Space: Relational Trajectories, 7 POLAR J. 259, 261 (2017) (detailing how the Antarctic Treaty System and Outer Space Treaties both created “extraterritorial zones” in which no nation could claim sovereign territory).
90 Moon Treaty, supra note 36, art. 11 § 5.
91 Id. art. 11 § 7.
92 Id. art. 15 § 2–3.
common law tradition, judges have the power to fill in the gaps of statutes and choose the rules that fulfill the enactor’s intent.\(^93\) While the U.N. treaties do not define “juridical person,” the growing adoption of environmental personhood can fill a gap in the law by providing a legal mechanism for the environmental regulation envisioned by the treaties. Currently, commercial and national actors engage in environmentally hazardous behavior with limited fear of international sanction or economic penalty. Recognition of celestial bodies as juridical persons will create a more sustainable and just future for humanity in space.

B. Environmental Personhood

1. Introducing the Concept

Protecting nature from exploitation is not a new legal problem: numerous international agreements have sought to promote the conservation of the Earth’s resources,\(^94\) atmosphere,\(^95\) and environments.\(^96\) Treaties are often drafted with exceptions that later become problematic or offer nations the opportunity to reject specific provisions. Without standing, nature and the organizations seeking to protect it have no basis to seek remedy from those profiting from contamination and exploitation. The concept of “environmental personhood” presents a legal means for the preservation and regulation of natural resources and can be extended to the outer space context.

Environmental personhood was first introduced by Professor Christopher D. Stone in a law review article advocating a reconsideration of humanity’s relationship with nature.\(^97\) Stone’s idea gained initial prominence when it was cited by U.S. Supreme Court Justice William O. Douglas in his dissent in *Sierra Club v. Morton*.\(^98\) In *Morton*, the plaintiffs could not seek judicial review. They lacked standing because the injury identified was to the natural environment and not to the individuals themselves. Justice Douglas’s dissent drew analogies to other juridical personhoods such as ships and corporations to illustrate that

---


environmental personhood could create a better sense of justice on behalf of a harmed ecosystem.  

Stone subsequently expanded upon his initial article in a book of the same name, detailing the inspiration for his idea and the challenges he foresaw. Stone noted that “throughout legal history, each successive extension of rights to some new entity has been, theretofore, a bit unthinkable,” citing the extension of the franchise and other protections to women and African Americans. In order to ensure the protection of an environmental person, a court would need to appoint a guardian ad litem or a representative as courts do in cases involving incompetent parties. The expansion of corporate rights is in many ways a precursor to the current growth of the environmental personhood movement, and subsequent proponents of expanding rights to non-natural persons have followed this model.

2. Subsequent International Adoption

In the years since Stone introduced his theory, there has been significant discussion of the environmental personhood concept within academic circles, with over 1500 articles citing his original journal article alone. Beyond academic momentum, various national and local governments have adopted forms of environmental personhood, either through legislative or judicial avenues. Around the world, there is a growing movement to preserve and secure the rights of nature.

Ecuador and Bolivia are the strongest adopters, implementing national environmental safeguards through juridical personhood. Ecuador adopted a constitutional amendment in 2008 to give nature the right to “exist, persist, maintain, and regenerate its vital cycles, structure, functions, and its processes in evolution.” Further, everyone in Ecuador has the right to sue on behalf of the

99 Id. at 741–43.
100 See generally CHRISTOPHER D. STONE, SHOULD TREES HAVE STANDING? LAW, MORALITY, AND THE ENVIRONMENT (3rd ed. 2010).
101 Id. at 2.
102 Id. at 8.
105 As with many legal concepts, there are both proponents of and detractors from environmental personhood. The merits are not fully discussed in this Comment; the focus is rather on the application of the concept.
environment. The principles of environmental personhood enshrined in their constitution have been upheld in Ecuadorian court, with the Provisional Court of Loja granting an injunction against the construction of a road and ordering the remediation of the Vilcabamba River in a 2011 decision. The initial claim was brought by two natural persons on behalf of the environment, demonstrating the feasibility of environmental personhood claims in court.

Along similar lines, Bolivia introduced legislation granting “Mother Earth” rights equal to those of natural persons in 2010 and 2012. Bolivia authorized the creation of governmental agencies to litigate on behalf of the earth and oversee climate change related policies. These laws, like Ecuador’s, promote the right-to-life for natural ecosystems and enable litigation on behalf of nature.

The U.S. and New Zealand have used a more limited approach, granting juridical personhood to specific environmental features, rather than to the entire environment. Within the U.S., the most notable effort to introduce environmental personhood was the Lake Erie Bill of Rights passed by a ballot measure in the City of Toledo, Ohio. The ballot measure gave Lake Erie the right to “exist, flourish, and evolve naturally” and empowered citizens to sue on behalf of the lake to enforce those rights against polluters. However, this measure was later nullified by the Ohio General Assembly in a provision added to an annual budget. The law stated that “[n]ature or any ecosystem does not have standing to participate in or bring an action in any court of common pleas” and “[n]o person, on behalf of or representing nature or an ecosystem, shall bring an action in any court of common pleas.” Additionally, a U.S. district court later found the Lake Erie Bill of Rights to be unconstitutionally vague and exceeding the

---


112 Jason Daley, Toledo, Ohio, Just Granted Lake Erie the Same Legal Rights as People, SMITHSONIAN MAGAZINE (Mar. 1, 2019), https://perma.cc/BF9D-EBEA.

113 OHIo REV. CODE ANN. § 2305.011 (West 2019).

power of a municipal government in Ohio. Although this effort to introduce environmental personhood was thwarted, the fact that it passed democratically shows the growing appetite for such legal solutions in the U.S. A number of local ordinance proposals across the U.S. indicate interest among local environmental activists.

New Zealand passed legislation granting environmental personhood to the Urewera Forest and the Whangangui River. The protection of the river is grounded in recognizing it as an ancestor of the Māori people. A representative of the Crown and a representative of the Whanganui iwi act as protectors of the river and its rights, giving them standing much like Justice Douglas envisioned.

Just days after the New Zealand parliament granted rights to the Whangangui River, a judicial ruling in India’s Uttarakand High Court extended similar protections to the Ganges River. The court appointed two state officials to serve as guardians of the river and its rights. This represents an interpretation closer to the aim of this Comment: rather than granting new rights, the Indian high court recognized the existing importance of nature and, thus, expanded legal standing.

In 2016, the Constitutional Court of Colombia reached a result similar to that of the Indian court, declaring that the “Atrato River basin possesses rights to ‘protection, conservation, maintenance, and restoration.’” This decision created a joint guardianship between a representative of the government and a member the indigenous peoples living in the river basin. Following the Constitutional Court’s decision, the Supreme Court of Justice of Colombia recognized the Colombian portion of the Amazon river as a “subject of rights.”

---


117 Te Urewera Act 2014 (N.Z.).


decisions represent a model for a high court adjudicating a claim involving environmental personhood.

In 2019, the Supreme Court of Bangladesh issued an even broader decision, granting legal rights to all of Bangladesh’s rivers to protect them from pollution and dredging. The Bangladeshi court appointed the National River Conservation Commission, a government agency, to serve as the legal guardian of the rivers and thus bring suits on the river’s behalf.

As can be expected from their extremely strong stances on granting the environment juridical standing within their own borders, Ecuador and Bolivia lead the way in advocating for international adoption of the environmental personhood model. During a 2010 conference in Bolivia, the Universal Declaration of the Rights of Nature was drafted with the intention of creating a new international treaty similar to the laws found locally in Bolivia. Similarly, Ecuador has supported the concept of an International Rights of Nature tribunal. Neither of these ideas have gained substantial traction within the international governmental community, but there has been support from environmentalist organizations.

The legal background for this Comment is bifurcated. Currently, the international treaty regime fails to substantially address what appears to be an imminent problem, creating a grim outlook for extraterrestrial environments. In contrast, the potential solution for extraterrestrial environmental protection appears to be taking off on Earth. The next section hypothesizes the fusion of these two realities.

IV. RECOGNIZING ENVIRONMENTAL PERSONHOOD FOR CELESTIAL BODIES AS A SOLUTION

Judicial application of the environmental personhood concept to the outer space context can create an effective regulatory regime by utilizing the existing treaty framework. Enacting large scale international legislation instituting environmental personhood is unlikely as support for even the Moon Treaty has been limited and corporate interests would likely oppose potential liability.

122 Rina Chandran, 
Fears of Evictions as Bangladesh Gives Rivers Legal Rights, Reuters (July 4, 2019), https://perma.cc/4SRX-WSUM.

123 See Andres Schipani, 
Grassroots Summit Calls for International Climate Court, The Guardian (Apr. 23, 2010), https://perma.cc/LSL3-PCQZ; Brandon Keim, 

124 Cormac Cullinan, 

125 See Linda Sheehan, 
Instead, environmentalists should pursue the judge-led expansion of rights for celestial bodies through legal action. Although the application to outer space might be new, the use of courts and judgments to enumerate and enforce rights is a proven tactic.

Utilizing the Outer Space Treaty and the Space Liability Convention, to which all major spacefaring nations are parties, can create a swift and clear result in favor of the environmental rights of celestial bodies. Through the texts of these two treaties, a case can be made that the ability to bring claims on behalf of celestial bodies already exists—it merely needs to be articulated by a judge.

Section A below explains the jurisdiction of the ICJ over the current treaties and environmental claims. Section B explores how a hypothetical plaintiff could have third-party standing in a contentious claim, with the ICJ either explicitly interpreting the term “juridical person” to include environmental person or using a common law approach to apply environmental personhood. Section C discusses the potential appeal for the ICJ to avoid a contentious decision that could be rejected by a defendant and instead issue an advisory opinion at the request of a specialized U.N. agency. The contentious claim is more binding but has less viability when compared to an advisory opinion. Subsections B and C will discuss their possibility of success weighed against their relative strengths. Section D explores the shortcomings of the ICJ adopting environmental personhood, and Section E discusses alternative solutions and their appeal.

Once environmental personhood for outer space bodies is implemented, there will be a deterrent effect of liability for actors that fail to prevent contamination through sufficient planetary protection protocols. A commercial space venture might implement better precautions if it believes precautions are a good investment to prevent or reduce potential liability. As a secondary benefit, the ICJ could order remediation efforts or funds as well, but once a contamination occurs it is hard to stop as we have seen with invasive species on Earth. Similarly, the value gained by exploiting space resources and diminishing the common heritage of mankind must be weighed against a potential judgment.

A. Jurisdiction of the International Court of Justice

Under the U.N. Charter, the ICJ has jurisdiction to issue advisory opinions and hear contentious cases. Members of the U.N. are automatically subject to the authority of the ICJ, so it cannot be avoided the same way that so many nations have simply failed to ratify the Moon Treaty. Presently, 74 countries have accepted compulsory jurisdiction of the ICJ, including Japan, Canada, and many members

---

126 U.N. Charter arts. 92, 94–96.
of the European Space Agency. In recent years, the ICJ has heard contentious cases involving environmental claims, including a 2008 claim by Ecuador against Colombia for the aerial spraying of herbicides and a claim by Argentina against Uruguay for the contamination of the Uruguay River. Although both of those claims were eventually withdrawn, they demonstrate the initial willingness of the ICJ to adjudicate and potentially assess the compensatory damages for environmental claims. In February 2018, the ICJ decided its first case involving environmental injury, holding that Nicaragua had to compensate Costa Rica for damages to the ecology along the border of the two nations. What is new to the court is not the idea of environmental claims, but rather the setting and idea of third-party standing on behalf of the environment.

B. Bringing a Contentious Claim

There is a growing belief that the ICJ can hear and settle disputes involving space law. The logic under the current regime is as follows: if disputes arising under the Outer Space Treaty fail to be resolved through diplomatic channels, then resolution must come from “international law, including the Charter of the U.N., in the interest of maintaining international peace and security and promoting international co-operation and understanding.” Article 33 of the U.N. Charter directs disputes between nations to be referred to the ICJ, granting the court jurisdiction over unsettled space claims. The Space Liability Convention also holds nations responsible for the actions of private parties launching from within their borders. The International Law Association drafted a proposed “Convention on the Settlement of Space Law Disputes” in 1984 before revising the language and formally adopting the text in 1998. The drafted Convention states a preference for a proposed International Tribunal of Space Law but provides for adjudication by the ICJ as the next alternative means of dispute resolution. This demonstrates a broader sentiment in the international

131 See Frans G. von der Dunk, Space for Dispute Settlement Mechanisms – Dispute Resolution Mechanisms for Space? A Few Legal Considerations, UNIV. OF NEB. COLLEGE OF LAW, SPACE, CYBER, AND TELECOMMUNICATIONS LAW PROGRAM FACULTY PUBLICATIONS (2001); see also VI IKARI, supra note 18, at 287 (noting the availability of the Permanent Court of Arbitration for use by private parties).
132 Outer Space Treaty, supra note 35, art. III, with similar language in the Space Liability Convention.
133 U.N. Charter art. 33; see also VI IKARI, supra note 18, at 289–90.
134 See VI IKARI, supra note 18, at 307.
legal profession that the ICJ has the institutional competence to handle such a case.

The most important part of the case would not necessarily be the merits of the claim, but the underpinning of the plaintiff’s standing. Examination of these processes shows the path toward environmental personhood for celestial bodies. The grounds for a contentious case could be based upon violations of the Outer Space Treaty, particularly the aforementioned language in Article IX directing parties to conduct their exploration and studies of celestial bodies while avoiding harmful contamination. Introducing tardigrades to the surface of the Moon could have negative consequences, as could plans to mine the lunar ice; a plaintiff would need to sue an actor causing some substantial effect of environmental degradation. If a private actor such as SpaceIL or the Arch Mission Foundation did contaminate the lunar environment and substantial environmental damage were subsequently proven, the U.S. would be the nation liable under the Space Liability Convention.

1. Bringing the Claim

A plaintiff-nation such as Ecuador could bring a contentious claim against the U.S. or Israel for their negligence in regulating their space industries and allowing the contamination of the lunar surface with tardigrades. The plaintiff-nation would need to sue the home country of any private actor rather than that actor themselves since the ICJ would not have jurisdiction over non-state parties. The defendant-nation could then seek to collect judgment from the private party responsible for the environmental damage. Ecuador could claim that it has standing based on the violation of the Outer Space Treaty through environmental personhood of the Moon and seek judicial interpretation by the ICJ of the relevant space treaties.

The Outer Space Treaty’s terminology prohibiting “adverse changes in the environment” and “harmful contamination” are largely undefined. The ICJ could conclude that exploitation or contamination of the lunar environment is a violation of the Outer Space Treaty. This presents a relatively straightforward way for a nation like Ecuador to bring a claim against offending nations. In order to seek damages though, it would be more useful to receive a ruling granting third-party standing.

2. Deciding a Case Explicitly Under Environmental Personhood

The Statute of the International Court of Justice includes sections determining the competence of the court and guiding the decision-making process. Under Article 38, the ICJ is to apply law based on international

---

135 See Viikari, supra note 77, at 729–30.
136 ICJ Statute, supra note 74.
convention, international custom, “general principles of law recognized by civilized nations,” and “judicial decisions and the teachings of the most highly qualified publicists of the various nations.” Utilizing this jurisprudential guidance, the ICJ could find that the international momentum is shifting towards recognizing environmental personhood as an increasingly accepted legal principle. Article III of the Outer Space Treaty makes all public international law applicable to space activities, so an international customary law of environmental personhood could be thought of as operating in the background of the Outer Space Treaty. The court would not have to reach far to see the many recent examples of high courts adopting the principle of environmental personhood and implement it in the international field of space law. Under this logic, the ICJ could recognize the Moon as a juridical person under the text of the Outer Space Treaty and enable other parties, such as Ecuador, to bring a claim on its behalf when the treaty has been violated.

Beyond the Outer Space Treaty, the ICJ could look to the Moon Treaty as the starting point for discussions regarding the laws governing the Moon. Despite its limited adoption, the Moon Treaty serves as the best articulation of international law for the Moon and could be used as an explanatory companion of the Outer Space Treaty as discussed previously. The language of the Moon Treaty was the product of an extensive drafting process by the U.N., including the American and Soviet edits that were implemented. While those nations ultimately declined to ratify the Moon Treaty as an objection to specific language, it still serves as the best demonstration of international intention in this area. In the same way that New Zealand explained its implementation of environmental personhood as recognition of the ancestry of the Māori people, the ICJ could utilize the language of the Moon Treaty, which states that the Moon is part of the “common heritage of all mankind.” Enabling claims by third parties brought on behalf of the Moon validates the language of the Moon Treaty, with the New Zealand implementation as precedent.

3. Deciding a Case Using a Common Law Approach

Beyond the argument that environmental personhood is becoming an accepted principle, the judges could also be motivated by the even broader idea that the creation of common law principles can be a form of regulation. Professor

---

137 Id. art. 38.
140 See Te Urewera Act, supra note 117; Te Awa Tupua Act, supra note 118.
141 Moon Treaty, supra note 36, art. 11.
Fabio Tronchetti, a leading space law expert, has called for the development of a legal regime that can protect the environment through reviewing and even interrupting activity that threatens the environment. Tronchetti has theorized that this regime could be an instrument attached to the Moon Treaty, Outer Space Treaty, or an independent legal instrument. Creating a liability system for environmental protection of celestial bodies would also be consistent with the Moon Treaty’s call for the establishment of an international regime governing the Moon. The Moon Treaty does not state that the regime must be administrative in nature. A judicially constructed liability regime may qualify under the treaty as a sort of delegation of authority by the States Parties under Article 11 § 5 of the Moon Treaty. The regime would need to address the Moon Treaty’s stated goals of safe and rational management of lunar resources and equitable benefits to all States Parties. Enabling standing for claims of harm to celestial bodies would allow states to sue when another state or private actor has acted in a manner that is not safe for the juridical person of the Moon. The use of a more common law approach through the ICJ would be a unique solution but still has the potential to satisfy many of Tronchetti’s criteria. Under the common law framework, judges will alter and improve rules in order to create a regulatory system of liability.

Indeed, the pollution of outer space brings in several familiar situations that may be addressable based on common law tort and property theories. When a river is being polluted upstream, there is an expectation in tort law that someone will be able to show an ex post injury downstream, providing a regulating effect through the plaintiff’s claim. Alternatively, a governmental authority can step in before an injury occurs and create an ex ante regulatory system. But both avenues to liability and thus regulation are absent in outer space. There is no governmental authority with the power to regulate or levy fines. The ex post deterrent is also weaker as it will be hard for any individual or government to demonstrate their present harm from extraterrestrial pollution or even the certainty that potential future harm will impact them specifically. A plaintiff such as Ecuador would struggle to show causation for an environmental damage occurring on the other side of the Earth, and it would be even more difficult when that distance is multiplied nearly twenty-fold to the surface of the Moon. Through the lens of property law, the resources on the Moon and other celestial bodies are common resource pools and thus threatened by the tragedy of the commons.

---

142 Tronchetti, supra note 11, at 811.
143 Id. at 812.
144 See generally Susan Rose-Ackerman, Regulation and the Law of Torts, 81 AM. ECON. REV. 54 (1991) (discussing the relationship between “private” tort law and “public” statutes in regulatory frameworks).
Without some rule protecting their usage, these resources will be exploited. But again, judges would be hard pressed to find who could claim the resources in order to preserve them, and thus the common law approach at first appears to fall short.

These common law principles begin to work again when environmental personhood is introduced. Recognizing the juridical personhood of the Moon would allow individual or organizational custodians to sue on behalf of the damaged ecosystem. While it would be necessary for another nation to bring the claim, such as Ecuador, the ICJ could grant a custodianship to an organization or select group of individuals. This is consistent with the models pioneered in India, New Zealand, and Colombia. The U.N. Committee on the Peaceful Uses of Outer Space or the U.N. Office for Outer Space Affairs could serve this role.

Grounding a decision in the tradition of the common law might be more appealing to ICJ justices. Rather than be accused of implementing a relatively new legal concept without international proof of concept and giving substantial power to single state-plaintiffs, the common law method is a smaller leap forward in judicial reasoning. The ICJ would be much closer to the customary international law sources of precedent in national courts and could draw on deeper wells of international precedent in the environmental and tort areas of law, demonstrated by Costa Rica v. Nicaragua and the Russian-Canadian settlement.

4. Prospect of Success

The potential of the contentious case strategy can be evaluated on two factors: the probability of securing a favorable judgment in the law and the ability to secure the desired remedy. Monetary damages are likely the best remedy for a claim of lunar environmental degradation. This form of penalization gives polluters an economic choice between adopting precautions and paying for remediation. Alternatively, an injunction would be too hard for the ICJ to enforce given that a defendant might choose to ignore the decision, knowing the ICJ lacks serious enforcement power. The contentious case strategy is weaker in its likelihood of success but stronger in its ability to provide a substantive remedy.

The probability of securing a judgment is undermined by the fact that it requires a plaintiff nation who is willing to finance the litigation, risk the diplomatic consequences of suing a powerful spacefaring nation, and lose future space industry revenue as a result of an anti-space industry reputation. Although Ecuador or Bolivia might be willing to take this step given their constitutional dedication to the environment and present lack of a space industry, the ICJ might

---

146 See Chandran, supra note 122.
147 See Warne, supra note 118.
148 See Bryner, supra note 121.
be concerned about the legitimacy costs of adopting such a new concept. But, grounding the decision as common law progress could mitigate this concern.

Given the history of damages awarded in recent international environmental cases, the contentious case strategy provides a useful avenue to secure the desired remedy. The compensation paid to Canada by Russia evidences that monetary damages can be appropriate compensation for damage in space.\(^{149}\) Instead of payment by the offending nation to the nation bringing the claim on behalf of a celestial body, the damages could be redirected toward lunar conservation or the funding of environmental impact research through COPUOS. Rather than require the direct sharing of economic benefits, which many industrialized countries found objectionable with the Moon Treaty, the contribution to lunar conservation could present a more acceptable tribute to the common heritage of all mankind. The relevant U.N. organs could manage an environmental conservation trust on behalf of the Moon. This idea combines the only precedent under the Space Liability Convention, the Russia-Canada settlement,\(^{150}\) with the recent precedent of environmental rulings at the ICJ, including the Costa Rica-Nicaragua dispute.\(^{151}\) Directing the use of funds towards the appropriate destinations is consistent with the Costa Rica-Nicaragua ruling. If the ICJ continues to move in the direction of awarding damages for environmental claims, the prospect of using a contentious case for resolution will grow.

During contentious cases, States Parties not represented in the composition of the ICJ’s bench have the opportunity to appoint an ad hoc judge pursuant to Article 31 of the Statute of the ICJ. Even if the court were to dismiss a case on the merits and not discuss the element of standing through environmental personhood, a judge appointed by the environmentalist nation bringing the claim could issue a dissenting or concurring opinion. An opinion approvingly citing environmental personhood would create kindling for future claims in the same manner as Justice Douglas in \textit{Morton}.

C. Seeking an Advisory Opinion

Alternatively, the ICJ could reach a similar conclusion without the use of a contentious case through its capacity to issue advisory opinions. The advisory opinion offers several benefits: it would not require a single nation to initiate the process, it would avoid creating an immediate loser, and it would offer an opportunity to create a legal regime without limiting the ICJ to a presented set of facts. Advisory opinions are not binding, but the requesting agency or organ can

\(^{149}\) Claims Protocol, \textit{supra} note 70.

\(^{150}\) \textit{Id}.

adopt the opinion to make it international law. Additionally, advisory opinions carry the weight of the international court and can influence subsequent behavior for actors wishing to avoid a detrimental contentious case.

1. Requesting an Advisory Opinion

Article 96 of the U.N. Charter says that “[t]he General Assembly or the Security Council may request the [ICJ] to give an advisory opinion on any legal question.” Further, “[o]ther organs of the United Nations and specialized agencies which may at any time be so authorized by the General Assembly, may also request advisory opinions of the Court on legal questions arising within the scope of their activities.” Previous U.N. agencies that have received advisory opinions include the World Health Organization and the Inter-Governmental Maritime Consultative Organization, now the International Maritime Organization.

The most relevant part of the U.N., COPUOS, would likely be an entity capable of asking for an advisory opinion as it operates as a subcommittee of the General Assembly. In order to bring a claim, members of the committee would need to pass a resolution asking the ICJ to clarify whether the term “juridical person” from Article I of the Space Liability Convention extends to environmental persons. Alternatively, an agency could request an advisory opinion on the same question. There are currently 17 U.N. specialized agencies, with the International Civil Aviation Organization and the U.N. Industrial Development Organization being the specialized agencies best positioned to request an advisory opinion given their normal areas of expertise. The most germane U.N. organ to request an advisory opinion would be UNOOSA, because of its outer space expertise, but it is below the specialized agency status and therefore lacks standing to request an advisory opinion. Generally, agencies might be the more likely to act, since committees can be paralyzed by protesting nations profiting from space exploitation.

The opinion would ideally come from an agency first requesting clarification as to whether the term “juridical person” in the various treaties could be interpreted as including environmental persons. The most important treaty for

---

153 U.N. Charter art. 96.
154 Id.
155 Legality of the Use by a State of Nuclear Weapons in Armed Conflict, Advisory Opinion, 1996 I.C.J. 93 (July 8).
157 Space Liability Convention, supra note 64, art. 1.
this part of the advisory opinion might be the Space Liability Treaty as it would then give rise to claims for damages. The next part of a request for an advisory opinion would seek to understand who can bring suit on behalf of juridical persons within the ICJ. While the ordinary answer might be based on the nationality of the juridical person, the space treaties would already prohibit the celestial bodies as being considered part of any country. The ICJ may recognize that any state party to the Outer Space or Space Liability treaties would have standing to bring a claim against a violator. It is possible that the ICJ would recognize the ability of natural persons to also bring suits on behalf of environmental persons under the idea of “common heritage” similar to the New Zealand example discussed previously, but this seems unlikely as the ICJ would probably caution against a deluge of claims from individuals. Beyond states parties, the next best plaintiffs would likely be the U.N. agencies themselves.

2. Advantages and Relative Value of an Advisory Opinion

In the context of an advisory opinion, the ICJ might be more willing to take a bigger leap in protecting the environmental futures of celestial bodies. The ICJ would be less afraid of losing legitimacy or seeing the immediate withdrawal of nations from its jurisdiction. Furthermore, issuing an advisory opinion is an inherently prospective exercise; it does not require adjudication between states parties and therefore avoids the potentially undesirable optics of creating an immediate loser in an area of previous legal uncertainty.

In comparison to a favorable contentious case opinion, securing a favorable advisory opinion is less valuable. The advisory opinion is not binding on specific parties and fails to deliver the precedent of monetary damages for extraterrestrial environmental damage. Partially redeeming the value of the advisory opinion is the limitation of the risk for the ICJ, with fewer political consequences for generating a potentially controversial ruling. Weighing the comparative benefits and risks, the contentious case likely offers the better strategy for environmentalists to attract significant international attention to the problem and potentially secure a remedy.

D. Shortcomings

Implementing environmental personhood for celestial bodies would be a substantial step forward for the jurisdiction of the ICJ. Ordering substantial damages or administrative action would likely stretch the boundaries of the court’s power. A stronger ICJ may be necessary as the world becomes more connected and some authority over space becomes essential to avoiding international conflict. In the absence of the international committee envisioned by the Moon Treaty, the ICJ may not be a perfect solution, but it might be one of the only available solutions at present. This Section will discuss several of the difficulties
associated with the proposed solution: national opposition, limits on institutional capacity, and alternative interpretation of the outer space treaties.

1. National Opposition

Nations are clearly interested in promoting the development of their space industries to grow their economies and acquire early dominance. The U.S. has repeatedly introduced legislation to economically incentivize the development of its space industry\(^{159}\) and has recognized the potential growth limiting factors of cooperating with the international regime.\(^{160}\) A critical flaw of the plan is the potential refusal or reluctance of the U.S. to comply with decisions of the ICJ.

After the ICJ ruled in favor of Nicaragua in *Nicaragua v. United States*,\(^{161}\) the U.S. refused to pay damages. The U.S. protested the court’s jurisdiction despite decades of previous compliance.\(^{162}\) While the U.N. Security Council has the power to enforce judgments of the court, the U.S. is a permanent member with veto power, a status it used repeatedly against attempts to collect reparations. The undetermined level of compliance by the U.S., particularly given its withdrawal from the Paris Agreement,\(^{163}\) may undercut the effectiveness of an ICJ judgment in favor of protecting celestial bodies. One commenter noted that an international regime regulating space would “be meaningless unless the U.S., the 800-pound gorilla in space, agrees to go along with the results.”\(^{164}\) Major commercial actors based in the U.S., such as SpaceX and Blue Origin, would be more shielded from a judgment, although the overall industry may be more global in nature. However, the language of the Outer Space Treaty, to which the U.S. is a party, does obligate it to cooperate in the international law of outer space. To the extent that the ICJ can dictate what the law is through advisory opinions and contentious holdings, the U.S. would be bound to it, at least in principle.

2. Institutional Capacity

Currently, there are a limited number of spacecraft and the focus of commercial spaceflight is on low Earth orbit, so it is conceivable to employ the

---


160 See Rosanna Sattler, *Transporting a Legal System for Property Rights: From the Earth to the Stars*, 6 CHI. J. INT’L L. 23 (2005) (discussing how the current space regime does not provide for property rights, a major impediment for space exploration, exploitation, and development).


164 Tannenwald, *supra* note 87, at 421.
ICJ and state plaintiffs to establish an early common law system for governing outer space environmental liability. More space traffic is undoubtedly coming as outlined in prior sections, which may call into question the efficacy of using a litigation-based regime. Hoping for environmentally virtuous nations to bring claims requires that they track the movements of potentially thousands of spaceflights and mining activities and then fund their claims at the ICJ. Even asking for a U.N. body to perform a regulatory capacity of this magnitude could quickly strain the resources of the U.N.

One possible remedy is to implement a regime of plaintiff’s attorney fees for bringing a successful claim on behalf of the environment to incentivize nations or organizations with standing to bring good claims. Attorney fees are common in international commercial arbitration, although importing standards from commercial arbitration could seem initially disconcerting within the context of the ICJ. Such a proposal might seem unappealing in more controversial contexts such as war claims but could be acceptable in a specialized tribunal for space claims.

Turning from international arbitration to the custom in domestic legal systems, the English rule of loser pays is nearly universal outside the U.S. The ICJ or another adjudicating entity could adopt loser-pays fees as customary international law, but this might also disincentivize plaintiffs fearful of footing a legal bill for a lost claim in an uncertain field of law.

There would be significant complications for a “polluter pays” liability regime as it is hard to track debris and other contaminants to the particular spacecraft depositing them. Further, the cost of pollution is difficult to calculate and the resulting damages might exceed any economic benefit from space activity, chilling adoption since an absolute prohibition is incompatible with human need. Determining the cost of pollution might require a determination about what has been taken from the commons entitled to all mankind or what cultural diminishment comes from altering a faraway surface.

3. Interpreting the Treaties’ Gaps as Enabling Mining

A final issue would be a defensive claim that mining the Moon is legal under international law. The language of the Outer Space Treaty and Moon Treaties can be construed as allowing the extraction of resources from celestial bodies. While lunar resources cannot be claimed as property while still in the Moon, it could be

167 Viikari, *supra* note 77, at 764.
168 Id.
argued that extracted resources are claimable.\textsuperscript{169} A ruling by the ICJ in favor of a defending extractor might occur under this interpretation, but it would not foreclose arguments based on contamination.

While it is conceivable that the text of the treaties could be used to enable private property, it requires creative and generous interpretation of the U.N. treaties to find sufficient loopholes. Article II of the Outer Space Treaty states that “[o]uter space, including the [M]oon and other celestial bodies, is not subject to national appropriation by claim of sovereignty, by means of use or occupation, or by any other means.”\textsuperscript{170} This language, echoed in subsequent treaties, strongly places the Moon in an extraterritorial category and potentially precludes property rights over natural objects in outer space. The Outer Space Treaty also requires states parties to ensure their citizens comply with the treaty, prohibiting private property by extension. Arguments that the treaty only bans sovereign territorial claims and not individual property claims fail on this point.\textsuperscript{171} Attempts by private citizens to prospectively claim asteroids have been rejected by the U.S. Court of Appeals for the 9th Circuit, which held that an individual could not charge NASA for parking fees for landing on an asteroid he claimed to own.\textsuperscript{172} The district court held that the fact that the U.S. was not a party to the more stringent Moon Treaty was not important; the plaintiff “failed to demonstrate that either statement establishes legal basis for his claim of a private property right on an asteroid.”\textsuperscript{173}

In practice, the U.S. and other spacefaring nations have taken possession of samples from outer space bodies, representing a form of property for national governments. Unlike the proposed large-scale mining operations, the samples taken by national space agencies thus far are relatively small amounts of material taken for scientific purposes with strict cross-contamination prevention protocols. Extraterrestrial sample collecting runs parallel with the scientific presence in Antarctica, pursuing scientific goals to benefit all with the resources belonging to the “common heritage of mankind.” Private exploitation and profit-seeking colonization does not fit within the spirit or straightforward reading of the U.N. space treaties.

Even if the ICJ or another body were to clearly state that the current outer space regime enables private property rights in what is meant to be an extraterritorial zone, the environmental personhood argument remains valid. Rather than deny the existence of private property rights in outer space, the

\textsuperscript{169} See Bilder, supra note 138, at 268.

\textsuperscript{170} Outer Space Treaty, supra note 35, art. II.


\textsuperscript{172} Nemitz v. NASA, 126 F. App’x 343 (9th Cir. 2005).

juridical personhood of the Moon would enforce a common property right to lunar resources. Bringing claims on behalf of the Moon would vindicate these rights and collect damages for restoration or conservation efforts equal to what the individual infringer has taken from the commons. Private actors would retain rights to technology developed on the Moon, but any contamination caused by research would be grounds for a judgment for restoration of the natural state of the lunar environment.

E. Alternative Solutions

1. Establishing a Moon Authority

Many of the issues discussed in this paper, as well as other outer space legal problems, could be solved by the establishment of a lunar authority.\textsuperscript{174} Such an authority could oversee permitting, conservation, and contamination prevention protocols. This authority is authorized by the Article 11 of the Moon Treaty, but the Moon Treaty’s limited adoption is perhaps prohibitive of the creation of the organization. Any effort to organize an International Moon Authority would need to address the concerns that prevented the adoption of the Moon Treaty, namely the tension between nations with the capability to exploit celestial resources and those nations still developing space technology.\textsuperscript{175} Utilizing the ICJ and environmental personhood may be preferable to establishing an administrative agency charged with the protection of celestial bodies. When the potential gains are trillions of dollars, the threat of regulatory capture of any administrative organization is grave.\textsuperscript{176} It is possible that the administrative body could become used not to conserve the Moon but to conserve the opportunities for exploitation of celestial bodies for those actors with the necessary resources and influence.

2. Permanent Court of Arbitration

An often-discussed alternative to ICJ dispute resolution is the Permanent Court of Arbitration (PCA). Unlike the ICJ, the PCA can hear claims by private parties, eliminating the need for a state to agree to bring a claim in the ICJ.\textsuperscript{177} If a claim were first brought in the PCA, it is less likely that the PCA would seek to progress international law and generate standing. The PCA also has fewer member states and lacks any form of compulsory jurisdiction. As a result, it would be better


\textsuperscript{175} Id. at 1404.

\textsuperscript{176} This refers to the idea that regulatory agencies can be “captured” by the influence of lobbyists from the industries they seek to regulate in order to coopt them to meet the industry’s goals.

\textsuperscript{177} Dispute Resolution Services, PERMANENT COURT OF ARBITRATION, https://perma.cc/8VHG-MSVG.
for a claim to first come in the ICJ, establish standing, and then allow parties to look to the PCA as a future alternative mechanism.

V. CONCLUSION

Even if a requested advisory opinion was not sufficiently clear on the issue of standing or a contentious claim failed on the merits, the act of bringing such a claim could lead to action within the international community. Faced with the possibility of liability for their extraterrestrial exploits, nations and private actors may begin to independently craft their own legal regime. While law enacted with the defendants might not be as appealing as a judicially crafted form of environmental justice, the discussion around a regulatory or legislative framework would at least bring attention to the risks already present. The U.S. and other nations may become more comfortable with the Moon Treaty if the alternative is a less predictable form of ICJ decision.

More broadly, outer space presents a *carte blanche* to explore a new way of considering humanity’s relationship with nature. On Earth, we are tethered to tradition and fear the costs of moving away from the known principles of standing and torts. Yet, just as a pair of daring nations reached for the stars, now a few bold nations are reconsidering what it means to exist with nature. Embracing environmental personhood offers humanity an opportunity to test the legal concept as a way to preserve the environment and imagine a new way of coexisting with nature, rather than destroying it.

While the dividends of protecting celestial bodies may not be appreciated by extraterrestrial human inhabitants for generations, there is a potential collateral benefit to the adoption of environmental personhood of the Moon. It is possible that a ruling by the ICJ in favor for environmental personhood spawns a reflective impact on the international community of Earth. Nations may view environmental personhood as a new international norm to which they should conform. Alternatively, the success of an outer space regime for environmental regulation could provide a roadmap for the establishment of systems to combat climate change on Earth. The Paris Agreement sought to take an untested leap forward, but perhaps an experiment in space will provide an example of international environmental cooperation that can be replicated at home.
The Application of International Tax Treaties to Digital Services Taxes
Katherine E. Karnosh*

Abstract

As digital services and electronic commerce have become more prevalent aspects of the global economy, there have been concerns over how tax systems will adapt to this change. International tax treaties in particular seem to be outdated and unprepared for the digital economy. Many international tax treaties provide that businesses are to be taxed on their income only in jurisdictions where they have a sufficient physical presence. By establishing their European headquarters and digital servers in countries with low corporate income tax rates (such as Ireland) and then using those headquarters to provide digital services to the rest of Europe, large, sophisticated, multinational digital businesses have been able to generate much revenue from most European countries without paying significant taxes in those countries.

The issues surrounding digital tax laws made news headlines in the summer of 2019 when France passed a law that imposed a 3 percent tax on revenue earned from digital services in France. Scholars have suggested that this tax may violate existing tax treaties, arguing that it provides for the taxation of the income of businesses without a significant enough physical presence in the country imposing the tax. This Comment analyzes this potential violation with regards to the French digital services tax and the U.S.—France Treaty for the Prevention of Double Taxation. This Comment concludes that the French tax does not violate the Treaty because the tax is a consumption tax that falls outside the scope of the Treaty.

* J.D. Candidate, 2021, The University of Chicago Law School. The author wishes to thank her family for their unending love, support, and encouragement, her advisor, Professor Adam Chilton, for his helpful suggestions, and the entire board and staff of the Chicago Journal of International Law for their valuable comments and thorough review.
Table of Contents

I. Introduction....................................................................................................................... 516
II. Overview of Principles of Taxation and International Taxation .................. 520
   A. Income Tax vs. Consumption Tax ................................................................. 520
   B. Double Taxation ............................................................................................... 522
   C. Base Erosion and Profit Shifting ................................................................... 524
III. International Tax Treaties ...................................................................................... 526
   A. Purpose ............................................................................................................... 526
   B. U.S.–France Tax Treaty ..................................................................................... 529
      1. The primary purpose of this Treaty is the prevention of double taxation. .................. 529
      2. The scope of the U.S.–France Treaty includes income taxes and wealth taxes but does not mention consumption taxes. ........................................ 530
      3. A country is allowed to levy a tax on business profits only when the business has a “permanent establishment” in the country. ............................. 531
   C. Other Relevant Tax Treaties ........................................................................... 532
      1. The Ireland-France Treaty is substantially similar to the U.S.–France Treaty. ................................................................................................................. 532
      2. The Netherlands-France Treaty is substantially similar to the U.S.–France Treaty. ................................................................................................................. 534
IV. Application—Do Digital Service Taxes Violate International Tax Treaties? ............................... 535
   A. Defining Digital Services and Digital Business ............................................. 535
      1. The three most common features of digital businesses are cross-jurisdictional scale without mass, the heavy reliance on intangible assets, and the importance of data, user participation, and network effects. ...... 535
      2. The three main business models for digital businesses are the value network, the value shop, and the value chain. .................................................. 536
      3. A definition of digital business for tax purposes must consider these features and business models and the problems they create for taxing authorities. ................................................................................................................. 537
      4. The EC’s definition of digital services accounts for the wide variance in the characteristics and business models of digital businesses. ............ 539
   B. Description of Enacted and Proposed Digital Services Taxes...................... 540
   C. Analysis ................................................................................................................. 541
      1. Under the typical understanding of the term “profits,” the French tax is not a tax on profits. ................................................................................................. 542
      2. Revenue was likely not intended to be considered an “element of income” under the Treaty. ................................................................................................. 543
3. The Treaty would likely be interpreted to give effect to the parties’ intents, which would mean that the French tax is not within the scope of the Treaty. ........................................................................................................................................... 544

4. The French tax does not come within the scope of the U.S.–France Treaty and therefore does not violate the Treaty. ............................................................................... 546

V. Conclusion........................................................................................................................................... 547
I. INTRODUCTION

As digital services and electronic commerce have become more prevalent aspects of the global economy, there have been concerns over how tax systems will adapt to this change. International tax treaties in particular seem to be outdated and ill-equipped to adapt to the digital economy. Historically, these treaties have provided that businesses should be taxed in the jurisdiction where they have a physical presence. For example, the U.S.–France Treaty for the Prevention of Double Taxation (U.S.–France Treaty) allows business profits to be taxed only if the business has a “permanent establishment” in the country. Permanent establishments include places such as offices or factories but do not include facilities used solely for storage and delivery or subsidiaries. The relevant provisions of the U.S.–France Treaty are typical across international tax treaties between other countries as well.

It has been relatively easy for firms participating in electronic commerce to avoid creating a permanent establishment in a country while still doing much business in and generating significant revenue from the country. Large, sophisticated, multinational companies have often established their European headquarters in countries with low corporate income tax rates (such as Ireland). This practice has contributed to the Base Erosion and Profit Shifting (BEPS) scheme, whereby companies “exploit gaps and mismatches in tax rules to artificially shift profits to locations with no/low tax rates and no/little economic activity.” Google is just one example of a business that has taken advantage of this phenomenon. For many years, Google issued its contracts for advertisements

---

2 Id. art. 7.
3 Id. art. 5.
in Europe from its European headquarters in Dublin, Ireland. This allowed the profits from those contracts to be taxed in Ireland, rather than the countries in which the advertisements took place. The Organisation for Economic Co-operation and Development (OECD) reports that governments lose between 100 and 240 billion USD annually as a result of BEPS tactics by multinational companies. Furthermore, because it is particularly easy for digital service providers to avoid having a permanent establishment in a country while still generating much revenue there, some scholars argue that BEPS has contributed to the discrepancy in the effective tax rate of digital businesses and more traditional businesses—digital businesses pay an effective tax rate of 9.5 percent, while traditional businesses pay 23.2 percent.

European countries (notably France) have called for reform in order to address the fact that large, multinational corporations are able to generate large amounts of revenue from their countries without being subject to the countries’ taxes. The OECD and the European Commission (EC) have both issued reports on digital taxes that outline the issue and propose possible solutions. The EC has proposed a permanent solution that involves allowing countries to tax businesses that have a “digital presence” in the country even if they do not have a permanent establishment there. A business would have a digital presence in a country if it met one of the following criteria: annual revenues of 7 million EUR in that country, more than 100,000 users in that country annually, or over 3,000 business contracts for digital services created annually. The EC has also proposed an interim policy to be used until a permanent solution is implemented. The OECD has “committed to continue working toward a consensus-based long-

---

8 Id.
10 Wamsley, supra note 5.
14 Id. at 8.
15 See id.
term solution” and has established a Program of Work for coming to a long-term solution.\(^\text{17}\)

The issues surrounding digital tax laws made news headlines in the summer of 2019 when France passed a law imposing a 3 percent tax on revenue earned from digital services in France.\(^\text{18}\) This tax applies to companies that earn more than 25 million EUR in French revenue and more than 750 million EUR in global revenue.\(^\text{19}\) This law largely reflects the interim policy proposed by the EC.\(^\text{20}\) The passage of this law generated strong reactions around the world, most notably with the U.S. challenging the law as discriminatory against American companies and threatening to impose tariffs on France in response.\(^\text{21}\) The U.S. and France later reached a deal of sorts about the law, under which France promised to repeal the French tax, adopt the OECD solution, and refund any difference in tax amount once the OECD finalized its plans for a unified, international framework for overcoming the problems presented by BEPS.\(^\text{22}\)

There are potential legal problems with both the French tax and the long-term solution proposed by the EC. Scholars have suggested that both likely violate existing tax treaties because they provide for the taxation of businesses without a “permanent establishment.”\(^\text{23}\) This Comment analyzes this potential violation with regards to the French digital services tax and the U.S.–France Treaty and concludes that the tax does not violate the Treaty. The U.S.–France Treaty’s permanent establishment requirement applies only to income taxes on “profits of

\(^{16}\) International Community Agrees on a Road Map for Resolving the Tax Challenges Arising from Digitalisation of the Economy, OECD (May 31, 2019), http://perma.cc/NK79-XUHU [hereinafter OECD, International Community Agrees]. The original deadline for reaching an agreement was the end of 2020; however, the COVID-19 pandemic has caused delays in negotiations, and the OECD now hopes to reach an agreement by mid-2021. International Community Renews Commitment to Address Tax Challenges from Digitalisation of the Economy (2020), http://perma.cc/H3BT-WQEB.

\(^{17}\) OECD/G20 INCLUSIVE FRAMEWORK ON BEPS, OECD, PROGRAMME OF WORK TO DEVELOP A CONSSENSUS SOLUTION TO THE TAX CHALLENGES ARISING FROM THE DIGITALISATION OF THE ECONOMY (2019), http://perma.cc/9HZ3-3URD [hereinafter OECD, PROGRAMME OF WORK].

\(^{18}\) Wamsley, supra note 5. This Comment refers to this tax as the French digital services tax, the French digital tax, and the French tax interchangeably.

\(^{19}\) Alderman, supra note 11.

\(^{20}\) See Fair Taxation of the Digital Economy, EUR. COMM'N (Mar. 21, 2018), http://perma.cc/V87H-GXNK.

\(^{21}\) Wamsley, supra note 5; Alderman, supra note 11.


an enterprise of a Contracting State.24 Because the French tax is more aptly characterized as a consumption tax on revenues rather than an income tax on profits, the French tax is likely outside the Treaty’s scope and is therefore not subject to the permanent establishment requirements. When appropriately characterized as a consumption tax, the French tax does not violate the U.S.–France Treaty.

This Comment proceeds as follows: Section II provides an overview of international taxation principles, focusing on the historical desire to avoid double taxation and the relatively recent rise of BEPS. This Section also explores the differences between income taxes and consumption taxes. Section III provides an overview of international tax treaties, focusing on how they came to be as well as their primary purposes. This Section highlights the U.S.–France Treaty and its provisions that apply to the digital services tax analysis. Finally, Section IV applies principles of treaty law to the digital services tax and concludes that the French digital service tax does not violate the U.S.–France Treaty.

This Comment is timely because of France’s recent action as well as discussions that other European countries will follow France’s lead and impose digital taxes of their own.25 As digital transactions become a more important part of the global economy, base erosion continues to impact governments’ ability to raise revenue.26 Additionally, the economic downturn caused by the COVID-19 pandemic is expected to drastically reduce most countries’ tax revenues.27 However, as many businesses and countries experience fiscal distress related to COVID-19, revenues and profits from many companies providing digital services have increased.28 These circumstances could make taxes on digital services an appealing option for countries searching for a stable tax base and striving to combat base erosion. As countries enact and consider digital taxes, it is important for them to understand how they can change their tax laws and the effect those

24 U.S.–France Treaty, supra note 1, art. 7.
25 As of May 2020, in addition to France, Austria, Hungary, India, Italy, Turkey, and the United Kingdom have implemented their own versions of digital services taxes. The Czech Republic, Latvia, Norway, Poland, Slovakia, Slovenia, and Spain have published proposals for or announced intentions to implement digital services taxes of their own. DANIEL BUNN, ELKE ASEN & CRISTINA ENACHE, DIGITAL TAXATION AROUND THE WORLD 18–19 (2020), http://perma.cc/PVJ3-9JJ8.
changes may have on treaty compliance. It is also important for countries to understand what types of changes they may need to make to their treaties in order to achieve their desired tax policy. Additionally, as the OECD works toward finalizing a proposed multilateral long-term solution, it will be helpful to have scholarly input and analysis.

While many scholars have noted the potential problems with digital services taxes and suggested solutions for dealing with these problems, there is need for greater study into the legality of these taxes under international tax treaties. This Comment’s analysis of the legality of these taxes will provide a basis for evaluating proposed solutions under the current treaty requirements.

II. OVERVIEW OF PRINCIPLES OF TAXATION AND INTERNATIONAL TAXATION

A. Income Tax vs. Consumption Tax

This Subsection will explain the background principles necessary for understanding income taxes and consumption taxes and how they relate to broader tax goals. Understanding the underlying theory behind these taxes and some of the important differences between them allows for evaluation of proposed and enacted taxes and the impact they are expected to have. Additionally, because international tax treaties often refer explicitly to “income taxes,” understanding what makes a tax an income tax allows for evaluating whether a tax is covered by a treaty or not.

According to the Haig-Simons definition of income, “[p]ersonal income may be defined as the algebraic sum of (1) the market value of rights exercised in consumption and (2) the change in the value of the store of property rights between the beginning and the end of the period in question.” This definition of income (as consumption plus change in net wealth) helps define the differences


30 See Gianni, supra note 29.

31 The most prominent types of taxes are income taxes, consumption taxes, and wealth taxes. Some combination of these taxes is used by nearly all countries around the world. See Joseph Bankman, Daniel N. Shaviro, Kirk J. Stark & Edward D. Kleinbard, Federal Income Taxation 51 (Wolters Kluwer, 18th ed. 2019). Wealth taxes are not implicated in the analysis of digital services taxes and international tax treaties, so this Subsection will focus on income taxes and consumption taxes.


33 Bankman et al., supra note 31, at 52 (quoting Henry Simons, Personal Income Taxation 50 (1938)).
between an income tax, a consumption tax, and a wealth tax. In broad strokes, an income tax is levied on everything that a taxpayer earns during the taxation period, regardless of whether the taxpayer spends what she earns or saves what she earns. Conversely, a consumption tax is levied only on what a taxpayer spends (consumes) during the taxation period. Income taxes differ from consumption taxes primarily in their treatment of savings (savings are taxed under an income tax but not under a consumption tax). Wealth taxes (such as property taxes or estate taxes) are typically levied based on a taxpayer’s total wealth holdings. The remainder of this Subsection will focus on the income tax and the consumption tax, as those are most relevant for the discussion of digital services taxes.

The primary types of income taxes include personal income taxes, which are levied on individuals’ wages and investment income, and corporate income taxes, which are levied on the profits of corporations.

In contrast to income taxes, which are levied against both individuals and businesses, consumption taxes tend to affect only individuals. “[O]nly sales to ultimate consumers are appropriately included in the base of a consumption tax.” There is an assumption that “businesses, as such, are never consumers” because a business purchases only that which is required for the business to prepare goods or services for the ultimate consumers (individuals). Although consumption taxes are often remitted to the taxing authority by businesses rather than individuals, the ultimate burden of the tax is almost always paid by the individual consumers in the form of higher retail prices.

The most common types of consumption taxes are sales taxes and value added taxes. Sales taxes are “simply a tax on final sales by businesses to consumers.” Sales taxes can be levied on either the sale of goods or services but are more commonly levied on the sale of goods. The value added tax is a tax on the seller’s “contribution to the value of the product.”

---

36 Id.
37 Id.
39 Aucutt, supra note 35, ¶ 401.1.
40 Id. ¶ 401.2.
41 Id. ¶ 401.3.
B. Double Taxation

The phenomenon of double taxation is not new and has been analyzed by scholars for centuries.\(^42\) Double taxation, or "the taxation of the same person or the same thing twice over,"\(^43\) stems from the increased complexity of trade, income earning, property holding, and other taxable transactions.\(^44\) In the international context, cross-border transactions can give rise to taxes from two or more countries on the same underlying transaction.\(^45\) The relevant countries compete with each other for taxing jurisdiction and both (or all) claim jurisdiction to levy their taxes.\(^46\)

This phenomenon is further complicated by the fact that the different jurisdictions have different principles underlying their taxing claims.\(^47\) The differing principles include taxing based on citizenship, temporary residence, permanent residence, location of property, and economic interest.\(^48\) Without consensus on the principles outlining where and by which jurisdiction taxation should occur, there will continue to be double taxation. This concern is especially prevalent in international contexts because countries worry about protecting their tax sovereignty.\(^49\)

Tax scholars and economists have criticized double taxation for centuries. The OECD has gone so far as to say that

[the] harmful effects [of double taxation] on the exchange of goods and services and movements of capital, technology and persons are so well known that it is scarcely necessary to stress the importance of removing the obstacles that double taxation presents to the development of economic relations between countries.\(^50\)

There are generally two lines of argument against double taxation. The first draws on normative notions of fairness. The second relies on empirical and economic arguments, including advocating for capitalistic free movement and economic prosperity.

---


\(^{43}\) Id. at 98.

\(^{44}\) Id.

\(^{45}\) Id. at 99.

\(^{46}\) Id.

\(^{47}\) Id. at 111.

\(^{48}\) Id. at 111–13.


First, double taxation is often criticized for being unjust. 51 For example, in cases where people live in one country and work in a different country, it is seen as an unfair burden that they might be required to pay tax on their income in two countries, given that people who live in the same country in which they work would only be required to pay tax on their income in one country. 52

Second, double taxation is said to “hinder the free movement of capital, technology and persons” around the world. 53 The taxation of the same underlying transaction by two different taxing authorities can create a disincentive to participate in that transaction. Tax policy is often used to intentionally shape economic behavior; for example, many countries have implemented tax laws designed to incentivize investment in environmentally-friendly technology. 54 However, uncoordinated tax policies that have not been carefully and intentionally designed have the potential to distort economic behavior in unintended and undesired ways. This risk is especially prevalent when separate and sovereign taxing authorities independently tax the same transaction. Uncoordinated double taxation could cause businesses to avoid cross-border transactions that would otherwise have positive impacts on the economy. Furthermore, “[d]ouble taxation is often cited as a major obstacle to unfettered economic progress.” 55 Scholars argue that “double taxation represents an unfair burden on existing investment and an arbitrary barrier to the free flow of international capital, goods and persons.” 56 Therefore, eliminating double taxation “facilitate[s] international trade by minimizing tax barriers in the exchange of goods and services across national boundaries.” 57

Depending on the taxing authorities’ policy goals and desired results of taxation, double taxation is not necessarily always a problem. 58 However, the type of double taxation by two or more different sovereign countries that occurs in the context of international transactions and multinational companies is typically not a type of double taxation that is the result of a specific policy goal but rather is the result of a lack of coordination between different countries. Therefore, it makes

51 SELIGMAN, supra note 42, at 114.
52 Jogarajan, supra note 49, at 690.
53 Id. at 680.
58 Dagan, supra note 55, at 942 n.6.
sense to criticize double taxation in the international context. It also makes sense for countries to take steps to reduce and eliminate any international double taxation of the same underlying transaction, as doing so will encourage more international transactions.

C. Base Erosion and Profit Shifting

The rapid development of digital services and electronic commerce creates difficulty for countries as they attempt to maintain fair levels and methods of taxation. The OECD reports that governments lose between 100 and 240 billion USD annually as a result of BEPS tactics by multinational companies. BEPS occurs when companies “exploit gaps and mismatches in tax rules to artificially shift profits to locations with no/low tax rates and no/little economic activity.” Such tactics are generally not illegal, but they are widely considered to be unfair.

An example of how a company is able to shift profits will be illustrative for understanding the problem posed by BEPS. Google is just one example of a business that has taken advantage of profit shifting. For many years, Google issued its contracts for advertisements in Europe from its Irish subsidiary located in Dublin, Ireland. The revenues from selling advertisements in other European countries were collected by Google’s Irish subsidiary, but the Irish subsidiary had no permanent establishments in these other European countries. This allowed the profits from those advertisements to be taxed in Ireland rather than the country in which the advertisement took place. Because Ireland has one of the lowest corporate income tax rates in Europe at 12.5 percent (for comparison, France has a corporate income tax rate of 34.4 percent), having profits taxed in Ireland rather than the other European countries allowed Google to reduce its overall tax liability. Overall tax liability is reduced even further by transferring the revenue from the first Irish subsidiary to a second subsidiary located in Bermuda but incorporated in Ireland. This revenue transfer is accomplished by selling Google’s technology and IP to the second subsidiary located in Bermuda and having that subsidiary license the technology to the first Irish subsidiary.

59 Hadzhieva, supra note 26, at 16.
60 OECD/G20 Inclusive Framework, supra note 6.
61 Id.
62 Dillet, supra note 7.
64 Dillet, supra note 7.
66 Kahn, supra note 63.
first Irish subsidiary then pays royalties to the subsidiary located in Bermuda, which greatly reduces the profits of the first Irish subsidiary. This results in the Irish subsidiary having little or no profits and the subsidiary located in Bermuda having nearly all of the revenue generated from the sale of the European advertisements.68 Due to the operation of the tax laws in Ireland and Bermuda—the Bermuda taxing authority treats the second subsidiary as located in Ireland because it is incorporated in Ireland, and the Irish taxing authority treats it as located in Bermuda because its management and control are located in Bermuda—the revenue now located in the second subsidiary in Bermuda is not subject to tax in either Ireland or Bermuda.69 This tactic, known as a “Double Irish Tax Sandwich” or simply a “Double Irish,” is not technically illegal—in 2017, the Paris administrative court ruled that Google did not act illegally by using this technique.70

The exact financial effect of such profit shifting tactics is difficult to calculate because of differences in tax deductions allowed and tax credits provided in different countries, but it is not difficult to imagine how shifting profits to countries with lower tax rates would generate tax savings for a business. For one potential measure of the success of BEPS techniques, France has previously alleged that Google owed French taxing authorities 1.12 billion EUR in back taxes for the period between 2005–2010.71 Google eventually settled disputes with French authorities for taxes from 2005 to 2018 by paying a 500 million EUR fine as well as 465 million EUR in back taxes.72

While BEPS does not apply only to digital service businesses, the ease with which digital businesses are able to use these tactics has received much attention. Some argue that this practice has contributed to the discrepancy in the effective tax rate of digital businesses and more traditional businesses—digital businesses pay an effective tax rate of 9.5 percent, while traditional businesses pay 23.2 percent.73

68 Id.
69 Id.
70 Gaspard Sebag, Google Spared $1.3 Billion Tax Bill with Victory in French Court, BLOOMBERG TECH. (July 12, 2017), http://perma.cc/32KN-Z26G.
71 Id.
72 Dillet, supra note 7.
73 Hadzhieva, supra note 26. But see, e.g., MATTHIAS BAUER, DIGITAL COMPANIES AND THEIR FAIR SHARE OF TAXES: MYTHS AND MISCONCEPTIONS 7 (ECIPE Occasional Paper No. 3, 2018), http://perma.cc/842F-U7N3 (“[The European Commission’s] selective use of firms and [its] obscure way of estimating effective corporate tax rates may . . . be convenient for the suggestion that traditional businesses pay higher rates than digital businesses. Yet it is also a misleading way to frame problems of international taxation. In fact, it is difficult to find the supporting evidence that digital firms (irrespective of their definition) and their profits are a big reserve of potential tax
III. INTERNATIONAL TAX TREATIES

Bilateral tax treaties are the most popular mechanism for addressing international double taxation.\textsuperscript{74} As of 2011, there were over 3,000 different bilateral tax treaties in force.\textsuperscript{75} The first bilateral tax treaties were largely based on a theory of reciprocity.\textsuperscript{76} Reciprocity in an international economic context is a way of ensuring that both countries grant the same privileges to citizens of the other country and thereby place citizens of both countries “upon an equal basis in certain branches of commerce.”\textsuperscript{77} The early bilateral tax treaties also alleviated double taxation by “allocating taxing rights for different types of income between jurisdictions.”\textsuperscript{78} Taxing rights were allocated “based on the source of income or the residence of the taxpayer.”\textsuperscript{79} These essential principles found in the very first international tax treaties (reciprocity and allocation based on source and residence) continue to be the principles that underlie modern tax treaties.\textsuperscript{80}

Subsection A of this Section will provide a historical overview of bilateral tax treaties from the first recorded treaty to modern day treaties. It will focus on the purpose of these treaties and the extent to which the purpose of these treaties has changed or remained the same. Subsection B will then turn toward the U.S.–France Treaty for the Prevention of Double Taxation, describing the purpose and scope of the Treaty and the requirements for taxation under the Treaty. Finally, Subsection C will briefly describe the tax treaty between Ireland and France and the treaty between the Netherlands and France. Because many multinational companies based in the U.S. have operating subsidiaries to service Europe in Ireland or the Netherlands, the relevant tax treaties when analyzing the French digital services tax may often be those between Ireland and France or the Netherlands and France.

A. Purpose

Beginning with the very first recorded international treaty on the topic of taxation (an 1843 agreement between France and Belgium), a primary factor

\textsuperscript{74} Jogarajan, \textit{supra} note 49, at 680.
\textsuperscript{75} \textit{Id.}
\textsuperscript{76} \textit{Id.} at 683.
\textsuperscript{77} \textit{Id.} at 682 (quoting JAMES LAURENCE LAUGHLIN \& HENRY PARKER WILLIS, RECIPROCITY 1 (Baker \& Taylor Company 1903)).
\textsuperscript{78} \textit{Id.} at 683.
\textsuperscript{79} \textit{Id.} at 683–84.
\textsuperscript{80} \textit{Id.}
motivating the agreement was the promotion and maintenance of significant economic benefit.\textsuperscript{81} Allowing Belgium access to markets throughout France benefited the economies of both Belgium and France, and the 1843 agreement facilitated this access by regulating the transfer of information between the two governments to “assist the effective and regular collection of taxes imposed by the laws of the two countries.”\textsuperscript{82} Modern tax treaties continue to include information-sharing clauses.

The first recorded international treaty entered to prevent double taxation was an 1899 agreement between the Austro-Hungarian Empire and Prussia.\textsuperscript{83} This agreement was designed to prevent double taxation in cases of dual nationality or residence in border areas, which was perceived to be unfair to border residents.\textsuperscript{84} This agreement allocated taxing rights by requiring the countries to levy personal taxes only on individuals with permanent residence in that country and business taxes only on businesses with permanent establishments in that country. This treaty was made in part with the goal of establishing strong political and economic relations between the Austro-Hungarian Empire and Prussia, as evidenced by other trade agreements between the countries that implemented favorable terms and attempted to strengthen the alliance.\textsuperscript{85} Prussia, the economically dominant country, wanted to firmly establish its economic control by creating an alliance, and the Austro-Hungarian Empire was eager to have greater access to the Prussian economy.\textsuperscript{86} As this first treaty demonstrates,

\begin{quote}
[r]he political context of the early treaties highlights the fact that, although countries may be concerned with the burden of double taxation on their citizens and double non-taxation, action to address these issues does not take place in isolation and is likely to be part of broader political and economic relationships.\textsuperscript{87}
\end{quote}

As countries shifted away from the view that the primary purpose of the income tax was government financing (especially of wars) to the view that its primary purpose was justice and wealth redistribution, they became more concerned with double taxation and worked to eliminate it.\textsuperscript{88} Because double taxation is often seen as unjust, a desire to design a just taxing system could lead to a desire to avoid double taxation, especially the double taxation of individuals. Eliminating double taxation...
taxation of individuals is largely what the earliest treaties aimed to do by reducing the tax burdens on border residents.

Tax treaties today play largely the same role that the earliest treaties played—that is, “facilitating economic integration” by “enabling the movement of people between states without the burden of double taxation.” This historical analysis is relevant because tax treaties today are motivated by economic and political factors just as they were in 1899. “Tax treaties do not exist in a vacuum and consideration of economic factors and political motivation will always be relevant.” Like the earliest treaties, the primary purposes of modern tax treaties are the prevention of double taxation and the facilitation of information sharing. Modern tax treaties also aim to combat tax evasion and fraud.

According to the introduction to the OECD’s Model Tax Convention on Income and on Capital (the OECD’s model convention), modern tax treaties have two primary purposes. First, tax treaties aim to “remov[e] the obstacles that double taxation presents to the development of economic relations between countries.” Second, tax treaties are intended to “improve administrative cooperation in tax matters, notably through exchange of information and assistance in collection of taxes, for the purpose of preventing tax evasion and avoidance.”

To further these purposes of removing double taxation and preventing tax evasion, there have been movements toward standardization of bilateral and multilateral tax treaties. This has led to the development of model treaties that are intended to be used as templates for countries drafting bilateral and multilateral treaties. Because the problems of double taxation and tax evasion are often identical or extremely similar even across jurisdictions, countries can use common solutions to combat these problems. The first model bilateral convention was established in 1928 by members of the League of Nations; it was followed by the Model Convention of Mexico in 1943 and the Model Convention of London in 1946, but none of these models was fully accepted. In 1956 the OECD began

---

89 Id. at 706.
90 Id. at 706–07.
91 Id. at 707.
94 OECD, MODEL TAX CONVENTION 2017, supra note 4, at 9.
95 Id.
96 Id.
97 Id.
98 Id. at 9–10.
working on a model convention that would be accepted by all member countries, which resulted in the 1963 Draft Double Taxation Convention on Income and Capital.\textsuperscript{99} This model convention has since been updated a number of times, with the most recent amendments in 2017.\textsuperscript{100}

The OECD’s model convention has had a significant impact. Since 1963, “OECD member countries have largely conformed to the Model Convention when concluding or revising bilateral conventions.”\textsuperscript{101} The model convention has influenced non-member countries as well, serving as the basis for negotiations of bilateral treaties between member countries and non-member countries as well as between two non-member countries.\textsuperscript{102} The model convention also served as the basis for the United Nation’s Model Double Taxation Convention between Developed and Developing Countries.\textsuperscript{103} Finally, the commentaries to the model convention have been generally accepted as a guide to the interpretation and enforcement of bilateral treaties.\textsuperscript{104}

The widespread use of the OECD’s model convention has furthered the historic purposes of avoiding double taxation and facilitating information sharing as well as the more modern purposes of avoiding tax evasion, tax avoidance, and tax fraud.

B. U.S.–France Tax Treaty

The starting point for an analysis of whether the French digital tax violates the U.S.–France Treaty for the Prevention of Double Taxation is the text of the Treaty itself. This Subsection begins by describing the purpose of the U.S.–France Treaty. It then turns to the scope of the Treaty as described in Article 2. It concludes by describing the Treaty requirements for taxation of commercial profits as established in Article 7, including the definition of a permanent establishment as defined in Article 5.

1. The primary purpose of this Treaty is the prevention of double taxation.

The stated purpose of the U.S.–France Treaty is the “avoidance of double taxation and the prevention of fiscal evasion with respect to taxes on income and capital.”\textsuperscript{105} Furthermore, the letter of submittal from the U.S. State Department

\textsuperscript{99} Id. at 10.
\textsuperscript{100} Id. at 11.
\textsuperscript{101} Id. at 12.
\textsuperscript{102} Id.
\textsuperscript{103} Id.
\textsuperscript{104} Id.
\textsuperscript{105} U.S.–France Treaty, supra note 1.
for the Treaty notes that the Treaty “also provides for administrative cooperation between the tax authorities of the two countries in applying the [Treaty] and the taxes covered by the [Treaty].”\textsuperscript{106}

The U.S.–France Treaty was updated in 1996 to replace the 1967 income tax treaty between the U.S. and France.\textsuperscript{107} Periodic updates to international tax treaties such as this one are required to reflect changes in domestic taxation law and policies and address new problems in cross-border transactions that arise in the years between the tax treaties. The letter of transmittal from the U.S. President to the U.S. Senate also states that the 1996 Treaty “more accurately reflects current income tax treaty policies of the [U.S. and France].”\textsuperscript{108} The U.S. State Department noted that an important improvement to the 1996 Treaty was the strengthening of the compliance aspects of the Treaty.\textsuperscript{109} This strengthened compliance was achieved by updating the provisions concerning exchange of information and the provisions concerning associated enterprises.\textsuperscript{110}

The primary purpose of this Treaty, the avoidance of double taxation, is achieved in part through restrictions on when the U.S. or France is allowed to tax a certain entity (described in Section III.C.3 below) and in part through Article 24. Article 24 requires that the U.S. allow taxpayers in the U.S. to claim credits against their U.S. income tax for certain taxes paid to France and that France allow taxpayers in France to claim credits against their French income tax for certain taxes paid to the U.S.\textsuperscript{111} By agreeing to only tax entities in accordance with the Treaty and by allowing credits for foreign taxes paid, the U.S.–France Treaty prevents a significant number of cases of double taxation.

2. The scope of the U.S.–France Treaty includes income taxes and wealth taxes but does not mention consumption taxes.

Under Article 2, the U.S. taxes that are subject to the Treaty include federal income taxes (except social security taxes) and excise taxes on certain insurance premiums.\textsuperscript{112} The scope of French taxes covered by the Treaty is somewhat broader. Article 2 states that “all taxes imposed on behalf of the State [of France], irrespective of the manner in which they are levied, on total income, on total capital, or on elements of income or of capital, including taxes on gains from the alienation of movable or immovable property, as well as taxes on capital

\textsuperscript{106} \textit{Id.} at Department of State Letter of Submittal.
\textsuperscript{107} \textit{Id.}
\textsuperscript{108} \textit{Id.} at Presidential Letter of Transmittal.
\textsuperscript{109} \textit{Id.} at Department of State Letter of Submittal.
\textsuperscript{110} \textit{Id.}
\textsuperscript{111} \textit{Id.} art. 24.
\textsuperscript{112} \textit{Id.} art. 2.
appreciation” are subject to the Treaty. Article 2 specifically notes that France’s income tax, company tax, tax on salaries, and wealth tax are all included. Notably, the Treaty does not mention consumption taxes (such as France’s value added tax or the U.S.’s sales taxes). These taxes are therefore presumably not included in the scope of the Treaty and are unaffected by any of the Treaty requirements.

3. A country is allowed to levy a tax on business profits only when the business has a “permanent establishment” in the country.

Article 7 of the Treaty provides that

[the profits of an enterprise of a Contracting State shall be taxable only in that State unless the enterprise carries on business in the other Contracting State through a permanent establishment situated therein. If the enterprise carries on business as aforesaid, the profits of the enterprise may be taxed in the other State but only so much of them as is attributable to that permanent establishment.]

A permanent establishment is defined in Article 5 of the Treaty as “a fixed place of business through which the business of an enterprise is wholly or partly carried on.” In particular, a permanent establishment includes “(a) a place of management; (b) a branch; (c) an office; (d) a factory; (e) a workshop; and (f) a mine, an oil or gas well, a quarry, or any other place of extraction of natural resources.” The following do not constitute a permanent establishment:

(a) the use of facilities solely for the purpose of storage, display, or delivery of goods or merchandise belonging to the enterprise;
(b) the maintenance of a stock of goods or merchandise belonging to the enterprise solely for the purpose of storage, display, or delivery;
(c) the maintenance of a stock of goods or merchandise belonging to the enterprise solely for the purpose of processing by another enterprise;
(d) the maintenance of a fixed place of business solely for the purpose of purchasing goods or merchandise, or of collecting information, for the enterprise;
(e) the maintenance of a fixed place of business solely for the purpose of carrying on, for the enterprise, any other activity of a preparatory or auxiliary character;
(f) the maintenance of a fixed place of business solely for any combination of the activities mentioned in subparagraphs (a) to (e), provided that the overall

---

113 Id.
114 Id.
115 Id. art. 7.
116 Id. art. 5.
117 Id.
activity of the fixed place of business resulting from this combination is of a preparatory or auxiliary character.\footnote{118}{Id.}

Importantly, the permanent establishment requirement only applies when a country wishes to tax the “profits of an enterprise of a Contracting State.” The Treaty does not provide an explicit definition of profits, but it does note that “[i]n determining the profits of a permanent establishment, there shall be allowed as deductions expenses which are reasonably connected with such profits, including executive and general administrative expenses, whether incurred in the State in which the permanent establishment is situated or elsewhere.”\footnote{119}{Id.}

C. Other Relevant Tax Treaties

It is important to note that the relevant treaty is the treaty between the country where the operating entity or subsidiary is based (which is not necessarily the country where the parent corporation is located) and the country imposing the tax.\footnote{120}{Thompson & Grandjouan, supra note 23.} Because many multinational companies based in the U.S. have operating subsidiaries in Ireland or the Netherlands that service Europe, the relevant tax treaties may often be those between Ireland and France or the Netherlands and France.\footnote{121}{Id.}

1. The Ireland-France Treaty is substantially similar to the U.S.–France Treaty.

The current tax treaty in effect between Ireland and France is the 1970 Double Taxation Treaty between Ireland and France (the Ireland-France Treaty). This Ireland-France Treaty uses much of the same language that is found in the U.S.–France Treaty.\footnote{122}{Double Taxation Treaty between Ireland and France: Convention between Ireland and France for the Avoidance of Double Taxation and the Prevention of Fiscal Evasion with Respect to Taxes on Income, Fr.-Ire., Jul. 14, 1970.} The stated purpose of the Ireland-France Treaty is “to avoid, so far as possible, double taxation and to prevent fiscal evasion in the matter of taxes on income.”\footnote{123}{Id.} The most relevant provisions of the Ireland-France Treaty are Article 1 (describing the scope of the agreement), Article 2 (defining “permanent establishment”), and Article 4 (providing for the taxation of commercial profits).

The Ireland-France Treaty applies to “taxes on income,” which are broadly defined as “all taxes imposed on total income or on the elements of income.”\footnote{124}{Id. art. 1.}
Article 1 specifically lists the French individual income tax, complementary tax, and tax on companies as well as the Irish income tax and corporate profits tax.\textsuperscript{125} Much like the U.S.–France Treaty, the Ireland-France Treaty does not mention consumption taxes.

Article 2 of the Ireland-France Treaty uses the same language as the U.S.–France Treaty to define “permanent establishment” as “a fixed place of business in which the business of an enterprise is wholly or partly carried on.”\textsuperscript{126} The Ireland-France Treaty also uses the same examples of permanent establishments as the U.S.–France Treaty, with the addition of (g):

(a) a place of management;
(b) a branch;
(c) an office;
(d) a factory;
(e) a workshop;
(f) a mine, quarry or other place of extraction of natural resources;
(g) a building site or construction or assembly project which exists for more than twelve months.\textsuperscript{127}

The Ireland-France Treaty’s specific exclusions from the definition of permanent establishments are slightly different from the U.S.–France Treaty’s exclusions:

(a) the use of facilities solely for the purpose of storage, display or delivery of goods or merchandise belonging to the enterprise;
(b) the maintenance of a stock of goods or merchandise belonging to the enterprise solely for the purpose of storage, display or delivery;
(c) the maintenance of a stock of goods or merchandise belonging to the enterprise solely for the purpose of processing by another enterprise;
(d) the maintenance of a fixed place of business solely for the purpose of purchasing goods or merchandise, or for collecting information, for the enterprise;
(e) the maintenance of a fixed place of business solely for the purpose of advertising, for the supply of information, for scientific research or for similar activities which have a preparatory or auxiliary character, for the enterprise.\textsuperscript{128}

Article 4 of the Ireland-France Treaty provides that “[t]he industrial and commercial profits of an enterprise of a Contracting State shall be taxable only in that State unless the enterprise carries on business in the other Contracting State through a permanent establishment situated therein.”\textsuperscript{129} It also states that the tax

\textsuperscript{125} Id.
\textsuperscript{126} Id. art. 2.
\textsuperscript{127} Id.
\textsuperscript{128} Id.
\textsuperscript{129} Id. art. 4
shall be imposed only on the amount of profits “as is attributable to that permanent establishment” with deductions allowed for “expenses which are incurred for the purposes of the permanent establishment.” Much like the U.S.–France Treaty, the Ireland-France Treaty does not provide an explicit definition of “profits.”

2. The Netherlands-France Treaty is substantially similar to the U.S.–France Treaty.

The current tax treaty in effect between the Netherlands and France is the 1973 Tax Treaty, as amended by the 2004 Protocol (the Netherlands-France Treaty). The relevant provisions of the Netherlands-France Treaty are Article 2 (describing the taxes covered by the agreement), Article 5 (defining “permanent establishment”), and Article 7 (providing for the taxation of business profits).

The Netherlands-France Treaty applies to “taxes on income and on fortune,” which are defined as “all taxes imposed on total income, on total fortune or on elements of income or of fortune.” Article 2 specifically lists the Netherlands’s income tax, wages tax, corporation tax, dividends tax, and fortunes tax as well as the French income tax and company tax. Just like the U.S.–France Treaty and the Ireland-France Treaty, the Netherlands-France Treaty does not mention any consumption taxes.

“Permanent establishment” is defined as “a fixed place of business in which the business of the enterprise is wholly or partly carried on.” The Netherlands-France Treaty uses the same examples and same specific exclusions of permanent establishments as the Ireland-France Treaty.

The Netherlands-France Treaty also uses the same language as the Ireland-France Treaty to describe when “[t]he profits of an enterprise” are taxable.

The fact that the U.S.–France Treaty, Ireland-France Treaty, and Netherlands-France Treaty do not vary in any significant ways is not surprising given the widespread use of the OECD model convention as a basis for bilateral tax treaties. Because of the similarity of the treaties’ relevant text, this Comment’s analysis will proceed using the text of the U.S.–France Treaty.

130 Id.
132 Id. art. 2.
133 Id.
134 Id. art. 5.
135 Id.
136 Id. art. 7.
137 OECD, MODEL TAX CONVENTION 2017, supra note 4, at 12.
IV. APPLICATION—DO DIGITAL SERVICE TAXES VIOLATE INTERNATIONAL TAX TREATIES?

A. Defining Digital Services and Digital Business

Defining digital business is more complicated than it may first appear. The spread of technology and the digitalization of the economy have affected businesses in every sector and every market. In addition to the exponential growth of the digital economy, the entire economy is going digital at a rapid pace.\(^\text{138}\) The percentage of businesses with a web presence increased from 2009 to 2017 in nearly all OECD countries, and 95 percent of businesses in OECD countries benefit from a high-speed internet connection.\(^\text{139}\) The World Economic Forum estimates that the value of the digital transformation will reach 100 trillion USD by 2025.\(^\text{140}\) Digitalization has facilitated an increased number of businesses providing services rather than tangible goods.\(^\text{141}\) Technological innovations, such as machine learning, are facilitating the digitalization of traditional businesses such as retailers, manufacturers, publishers, and health care providers.\(^\text{142}\) Digitalization has also facilitated the globalization of businesses by making it easier for formerly domestic businesses to interact digitally with customers around the world.\(^\text{143}\) As digitalization becomes more and more ubiquitous, even small firms are becoming capable of reaching customers worldwide.\(^\text{144}\) In a broad sense, the digital economy can be defined as “all activities that use digitized data,” which would encompass, in essence, the entire modern economy.\(^\text{145}\) More narrowly defined, the digital economy is comprised of “online platforms, and activities that owe their existence to such platforms.”\(^\text{146}\)

1. The three most common features of digital businesses are cross-jurisdictional scale without mass, the heavy reliance on intangible assets, and the importance of data, user participation, and network effects.

According to the OECD’s 2018 Base Erosion and Profit Shifting Interim Report, the three most salient and common features of digitalized businesses are

\(^{138}\) Hadzhieva, supra note 26, at 14.

\(^{139}\) OECD, TAX CHALLENGES, supra note 12, at 14.

\(^{140}\) Hadzhieva, supra note 26, at 14.

\(^{141}\) OECD, TAX CHALLENGES, supra note 12, at 38.

\(^{142}\) Hadzhieva, supra note 26, at 14.

\(^{143}\) OECD, TAX CHALLENGES, supra note 12, at 52.

\(^{144}\) Id.

\(^{145}\) Hadzhieva, supra note 26, at 14.

\(^{146}\) Id.
cross-jurisdictional scale without mass, the heavy reliance on intangible assets, and the importance of data, user participation, and network effects.\textsuperscript{147} Cross-jurisdictional scale without mass refers to the fact that digitalization “allows some highly digitalized enterprises to be heavily involved in the economic life of a jurisdiction without any, or any significant, physical presence.”\textsuperscript{148} Heavy reliance on intangible assets means that investments in IP and “the intense use of IP assets such as software and algorithms supporting [] platforms, websites, and many other crucial functions” are important aspects of the business models of many digital businesses.\textsuperscript{149} Data, user-generated content, and network effects from increased participation are also central to the business models of highly digitalized businesses; in fact, some digital businesses, such as social media networks, would not exist without a high level of user participation.\textsuperscript{150} These three characteristics are not exclusive to digital businesses, but they are more prevalent among digital businesses and are helpful for defining what counts as a digital business.

2. The three main business models for digital businesses are the value network, the value shop, and the value chain.

A common way of characterizing traditional businesses is based on their business model, especially where the business generates value. Because existing tax systems are based on the principle of taxing profits where value is created,\textsuperscript{151} understanding how value creation in a digital business differs from that of a traditional business will be helpful for defining what counts as a digital business for tax purposes.

Digital businesses can take the general form of three different business models: the value network, the value shop, and the value chain. Highly digitalized businesses often take the business model form of value networks.\textsuperscript{152} In a value network, value is created by organizing and facilitating an exchange between users through the business’s mediating technology.\textsuperscript{153} This type of business model leads to revenue generation that often differs from non-digital businesses.

Revenue in value networks may be generated through subscription fees (e.g., LinkedIn Premium) or pay-as-you-go fees when services are consumed (e.g., Airbnb, BlaBlaCar). In other cases, such as Instagram, Facebook, Twitter and Weibo, the business may, in what may be perceived by some countries as a type of barter transaction, offer access to the platform without a demand for financial compensation upon the user providing some input valuable to the

\begin{footnotesize}
\begin{itemize}
\item\textsuperscript{147} OECD, TAX CHALLENGES, supra note 12, at 24.
\item\textsuperscript{148} Id.
\item\textsuperscript{149} Id.
\item\textsuperscript{150} Id.
\item\textsuperscript{151} Hadzhieva, supra note 26, at 19; Proposal for a Council Directive, supra note 12.
\item\textsuperscript{152} OECD, TAX CHALLENGES, supra note 12, at 38.
\item\textsuperscript{153} Id.
\end{itemize}
\end{footnotesize}
platform operator. Such input could be personal information about the user’s interests that can be employed to generate revenue from targeted advertising. It could also be content accessible by other users, which increases the platform’s utility and value.\textsuperscript{154}

The second type of business model that is common among digital businesses is the value shop.\textsuperscript{155} In a value shop, a business uses its technology to solve a specific customer demand or problem, such as specialized data analysis, software development, or cloud computing.\textsuperscript{156} The final type of business model, the value chain, is less common among digital businesses. In a value chain, a business designs, produces, markets, delivers, and supports its products.\textsuperscript{157} The value chain most often describes traditional businesses that produce tangible goods, but can also include businesses that operate as resellers of goods and businesses that produce intangible goods or services such as movies, games, music, or software.\textsuperscript{158} Additionally, as the business landscape and global economy increase in complexity, many digital businesses span several of these three general business models.\textsuperscript{159} “For example, Amazon’s retail business is considered a value chain . . . whereas Amazon Marketplace, which links buyers and sellers in order for them to trade, is considered a value network, and Amazon Web Services is considered a value shop.”\textsuperscript{160}

3. A definition of digital business for tax purposes must consider these features and business models and the problems they create for taxing authorities.

The most salient features of digital businesses, combined with the use of new business models such as value networks, create certain problems for taxing authorities as they attempt to tax these businesses. Cross-jurisdictional scale without mass allows “a growing number of businesses [to] have an economic presence in a jurisdiction without having a physical presence.”\textsuperscript{161} As noted in Section III of this Comment, a lack of a physical presence could lead to an inability to tax such business’s income under current bilateral tax treaties. Heavy reliance on intangible assets means that “[t]he location in which a business’ intangible assets are controlled/managed can [] have a material impact on where that business’ profits are subject to tax” regardless of where those IP rights are used.

\begin{footnotesize}
\begin{enumerate}
\item[154] Id. at 39.
\item[155] Id. at 40.
\item[156] Id.
\item[157] Id. at 38.
\item[158] Id. at 38–39.
\item[159] Id. at 42.
\item[160] Id.
\item[161] Id. at 51.
\end{enumerate}
\end{footnotesize}
to generate revenue.\textsuperscript{162} Heavy reliance on intangibles also makes it easier for “companies to structure themselves to minimize their tax liabilities and makes it more cumbersome for tax authorities to assess how income from such assets should be identified, valued and allocated amongst different parts of multinational groups.”\textsuperscript{163} The importance of data, user participation, and network effects, especially when used in a value network business model, can lead to barter transactions in which the digital business provides non-financial compensation (for example, data hosting, email services, or digital entertainment) to users in exchange for user-generated content and the ability to collect the users’ data.\textsuperscript{164} Today’s income tax systems rarely capture this type of barter exchange (in which no cash payment is made on either side of the transaction), but many believe that this type of transaction should be subject to taxation.\textsuperscript{165} By way of summary, the EC notes that

\begin{quote}
[\textbf{t}he application of the current corporate tax rules to the digital economy has led to a misalignment between the place where the profits are taxed and the place where value is created. In particular, the current rules no longer fit the present context where online trading across borders with no physical presence has been facilitated, where businesses largely rely on hard-to-value intangible assets, and where user generated content and data collection have become core activities for the value creation of digital businesses.]\textsuperscript{166}
\end{quote}

A definition of digital services must consider each of these characteristics of digital businesses and business models and the problems they create for taxing authorities. Additionally, the perfect definition would consider that the characteristics of digital businesses and business models are not exclusive to digital businesses nor comprehensive of digital businesses that pose problems for the existing tax systems. Not all digital businesses reflect the characteristics described above, and sometimes non-digital businesses can have those characteristics. For example, a traditional manufacturing business can have cross-jurisdictional scale without mass by selling products in countries where it does not have a manufacturing site.\textsuperscript{167} Additionally, non-digital businesses can use the value network and value shop business models, and digital businesses can use the value chain business model. For example, a traditional employment agency uses the value network business model by bringing together employers and job seekers.\textsuperscript{168}

\begin{flushright}
\textsuperscript{162} Id. at 52.
\textsuperscript{163} Hadzhieva, \textit{supra} note 26, at 17.
\textsuperscript{164} OECD, \textit{TAX CHALLENGES}, \textit{supra} note 12, at 58–59.
\textsuperscript{165} Id. at 59.
\textsuperscript{167} OECD, \textit{TAX CHALLENGES}, \textit{supra} note 12, at 38.
\textsuperscript{168} Id.
\end{flushright}
4. The EC’s definition of digital services accounts for the wide variance in the characteristics and business models of digital businesses.

The EC considered these problems in defining digital businesses in its proposed solution to the taxation challenges posed by digital businesses.\textsuperscript{169} The EC defines “digital services” as “services which are delivered over the internet or an electronic network and the nature of which renders their supply essentially automated and involving minimal human intervention, and impossible to ensure in the absence of information technology.”\textsuperscript{170} Its proposed definition explicitly includes:

(a) the supply of digitized products generally, including software and changes to or upgrades of software;

(b) services providing or supporting a business or personal presence on an electronic network such as a website or a webpage;

(c) services automatically generated from a computer via the internet or an electronic network, in response to specific data input by the recipient;

(d) the transfer for consideration of the right to put goods or services up for sale on an internet site operating as an online market on which potential buyers make their bids by an automated procedure and on which the parties are notified of a sale by electronic mail automatically generated from a computer;

(e) Internet Service Packages (ISP) of information in which the telecommunications component forms an ancillary and subordinate part, in other words packages going beyond mere internet access and including other elements such as content pages giving access to news, weather or travel reports, playgrounds, website hosting, access to online debates or any other similar elements . . . .\textsuperscript{171}

According to the EC’s guidance on this definition, “involving minimal human intervention” refers to the amount of human service provided by the business “without any regard to the level of human intervention on the side of the user.”\textsuperscript{172} A business that “initially sets up a system, regularly maintains the system, or repairs it in cases of problems linked with its functioning” will be considered to involve only a “minimal human intervention” and thus be providing a digital service.\textsuperscript{173} Additionally, “the sale of goods or other services which is facilitated by using the internet or an electronic network” is explicitly excluded from the definition of digital services.\textsuperscript{174} According to the EC’s guidance, “giving access [] to a digital


\textsuperscript{170} Id. at 14.

\textsuperscript{171} Id. at 14–15.

\textsuperscript{172} Id. at 7.

\textsuperscript{173} Id.

\textsuperscript{174} Id. at 15.
marketplace for buying and selling cars is a digital service, but the sale of a car itself via such a website is not.175

This definition of digital services appears to recognize and account for the difference between truly digital businesses and traditional businesses that have adopted some aspects of digitalization. By excluding the mere sale of goods or services over the internet and requiring only minimal human intervention, the definition excludes many businesses that might involve some aspects of digitalization or digital service (such as a manufacturer with an online webstore) but whose traditional business models of producing and selling goods do not pose the same problems to the tax system that truly digital businesses pose.

The EC’s proposed solution to BEPS uses the term “significant digital presence” to describe when a digital business has enough of a presence in a country to be subject to taxes in that country.176 The EC proposed three different ways for a business to meet the significant digital presence test. A business would have a significant digital presence in a country if 1) the business’s revenues from providing digital services to users in the country exceed 7 million EUR in the tax period, 2) the number of users of the business’s digital services in the country exceeds 100,000 in the tax period or 3) if the business’s number of business-to-business contracts for digital services exceeds 3,000 in the tax period.177

These three tests for significant digital presence were intended to “reflect the reliance of digital businesses on a large user base, user engagement and user’s contributions as well as the value created by users for these businesses.”178 The criteria therefore reflect the most salient characteristics of digital businesses. Additionally, the use of three different criteria reflects an understanding that digital businesses can assume various different business models179 and tries to capture as many of these different models as possible. Finally, the minimum thresholds were designed to leave out trivial cases where profits from a digital presence are not even expected to be large enough to cover the cost of tax compliance.180

B. Description of Enacted and Proposed Digital Services Taxes

The French digital tax law imposes a 3 percent tax on revenue earned from digital services in France on all companies that earn more than 25 million EUR in

---

175 Id. at 7.
176 Id. at 2.
177 Id. at 8.
178 Id. at 7–8.
179 Id.
180 Id.
French revenue and more than 750 million EUR in global revenue.\textsuperscript{181} This law would affect over 24 companies, many of which are not French-based.\textsuperscript{182} The passage of this law generated strong reactions around the world, most notably with the U.S. challenging the law as discriminatory against American companies and threatening to impose tariffs on France in response.\textsuperscript{183} The U.S. and France later reached a deal of sorts about the law, under which France promised to repeal the law and refund any difference in tax amount once the OECD finalized its plans for a unified, international framework for the taxing of digital companies.\textsuperscript{184}

As noted above, the EC’s proposed permanent solution would allow a country to tax a business without a permanent establishment as long as that business had annual revenues of at least 7 million EUR in that country, more than 100,000 users in that country annually, or over 3,000 business contracts for digital services created annually.\textsuperscript{185} This solution may require the amendment of most bilateral and multilateral tax treaties which still require a business to have a physical location in the country in order for that country to tax the business.

The OECD has “committed to continue working toward a consensus-based long-term solution”\textsuperscript{186} and has established a Program of Work for coming to a long-term solution.\textsuperscript{187} The Program of Work focuses on two major areas in which change and consensus is needed: 1) nexus requirements and profit allocation and 2) ensuring that a minimum global level of tax is paid.\textsuperscript{188}

C. Analysis

Several scholars have pointed out that these proposed digital tax laws may violate bilateral and multilateral tax treaties.\textsuperscript{189} It seems clear that companies falling under the proposed digital taxes may not meet the permanent establishment requirement under Articles 5 and 7 of the U.S.–France Treaty.

However, before the issue of permanent establishment arises, there is a more basic question of how to classify these digital taxes. The permanent establishment requirement applies only to taxes on “profits of an enterprise of a Contracting

\textsuperscript{181} Wamsley, supra note 5.
\textsuperscript{182} Alderman, supra note 11.
\textsuperscript{183} Wamsley, supra note 5; Alderman, supra note 11.
\textsuperscript{184} Gold, supra note 22.
\textsuperscript{186} OECD, International Community Agrees, supra note 16.
\textsuperscript{187} OECD, PROGRAMME OF WORK, supra note 17.
\textsuperscript{188} Id.
\textsuperscript{189} See Govindarajan et al., supra note 23; Gottlieb & Ali, supra note 23; Thompson & Grandjouan, supra note 23.
The French digital tax is a tax on revenue, rather than income. It is not immediately clear that such a tax would be categorized as a tax on “business profits.”

1. Under the typical understanding of the term “profits,” the French tax is not a tax on profits.

Article 7 does not include an explicit definition of “profits,” but the analysis must begin with whether the digital services tax on revenue qualifies as a tax on profits or not. Article 3 states that when the U.S. or France is applying the Treaty, any undefined terms will “have the meaning which [they have] under the taxation laws of that State.” The typical understanding of profit is revenue less expenses required to earn that revenue. Revenue refers to “the total amount of money the business receives . . . for its products and services.” Income, or profits, refers to what remains after expenses are subtracted from revenue. This understanding is reflected in the U.S.’s income tax law which allows deductions for “all the ordinary and necessary business expenses” of the business. This understanding is also reflected in France’s corporate tax law (l’impôt sur les sociétés) which taxes profits (les bénéfices) earned in France and allows deductions for normal business expenses. This understanding of profits is also reflected in Article 7 of the Treaty, which notes that the countries shall allow deductions for “expenses which are reasonably connected with such profits[.]” The French digital tax allows no such deduction of expenses. Given this understanding of the difference between revenue and profit, it seems unlikely that the French digital tax’s explicit tax on revenues without a deduction for any expenses would be considered a tax on business profits.

In further support of this argument that the French digital tax is not a tax on business profits, sales taxes are generally levied on the total sales prices, which is equal to the total revenues, from the goods or services being taxed. Sales taxes are classified as consumption taxes rather than income taxes (see Section II.A.

190 U.S.–France Treaty, supra note 1, art. 7.
191 See id.
192 U.S.–France Treaty, supra note 1, art. 3.
197 U.S.–France Treaty, supra note 1, art. 7.
198 Aucutt, supra note 35, ¶ 401.2.
above). The French tax on revenue seems more like a sales tax than an income tax. This characterization as a consumption tax seems especially apt considering that, shortly after France announced its digital services tax, Amazon announced that it would increase the Amazon Marketplace fees it charges its sellers by 3 percent (which Amazon then expects to be passed on to consumers in the form of higher prices).

This method of passing taxes through to the ultimate individual consumer is a hallmark of consumption taxes (rather than income taxes). These considerations would seem to imply that the French digital tax is not a tax on business profits under Article 7 of the Treaty.

2. Revenue was likely not intended to be considered an “element of income” under the Treaty.

One counterargument to the conclusion that the French tax is not within the scope of the Treaty is based on an understanding of the term “element of income.” Article 2 of the Treaty states that “all [French] taxes imposed on behalf of the State, irrespective of the manner in which they are levied, on total income, on total capital, or on elements of income or of capital” are subject to the Treaty. Revenue as a whole could be considered an element of income because income equals revenue less expenses. This would imply that a tax on revenue would be subject to the Treaty.

However, bilateral tax treaties that are based on the OECD’s model convention should be interpreted using the commentaries to the model convention. The commentaries to the 1992 OECD Model Tax Convention, which was the model in effect when the 1996 U.S.–France Tax Treaty was signed and ratified, refer to “items of income” several times, but never refer to revenue as an “item of income.” The Model Tax Convention commentaries explicitly acknowledge “dividends,” “interest,” “royalties,” “wages,” and “salaries” as items or classes of income that are subject to the Treaty. Each of these items is an element of income, in the sense that each is a component included in the calculation of income. Yet each of these items is much narrower

200 U.S.–France Treaty, supra note 1, art. 2 (emphasis added).
202 Id. at 12–13, 53.
203 Id. at 12.
204 Id.
205 Id. at 13.
206 Id. at 53.
207 Id.
than the broad term “revenue.” If the drafters of the Model Tax Convention had intended that taxes on revenue be considered taxes on “elements of income” and thus subject to the Treaty, it seems likely that the commentary would have listed “revenue” when it listed the other items and classes of income that are subject to the Treaty. This leads to the inference that, although revenue is involved in the calculation of income and therefore may technically be considered an element or item of income, the parties to the Treaty did not intend for revenue to be considered an element of income. Therefore, it seems likely that taxes on revenue were intended to be outside the scope of the Treaty and not subject to the permanent establishment requirement.

3. The Treaty would likely be interpreted to give effect to the parties’ intents, which would mean that the French tax is not within the scope of the Treaty.

A second counterargument to the conclusion that the French tax is outside the scope of the Treaty is based on the purpose of the Treaty. The purpose of this Treaty is the “avoidance of double taxation and the prevention of fiscal evasion.” One could argue that if France taxed the revenues of these companies, but those same revenues were later taxed as income by the U.S., one of the primary purposes of the Treaty would be frustrated. A broad reading of the Treaty in light of the Treaty’s purpose would imply that the French tax on revenue should be covered by the Treaty. The Vienna Convention on the Law of Treaties would seem to support such a broad reading. The Vienna Convention provides guidelines for the interpretation of international treaties and states that “[a] treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose.” This differs from the contractual, or intentionalism, approach by focusing on the broad goals of the treaty rather than the subjective intent of the parties. This broad reading could lead to an interpretation that, in light of the goal of preventing double taxation, the French digital tax is prohibited.

However, because a tax on revenues is likely to be passed onto the final consumer and is only collected by the taxing authority from the business for ease of collection purposes, there is arguably no double taxation in the typical sense because the tax burden is borne by different entities (the revenue tax is borne by the individual consumers and the income tax is borne by the business). Across a wide range of empirical analyses, it has been found that anywhere from 60 percent

---

208 U.S.–France Treaty, supra note 1; see Rebecca M. Kysar, Interpreting Tax Treaties, 101 IOWA L. REV. 1387 (2016) (suggesting that tax treaties should be interpreted in a “pragmatic fashion”).
210 Kysar, supra note 208, at 1402–03.
to over 100 percent of sales taxes are passed on to the final consumer. Shortly after the French tax was announced, Amazon announced that it would pass the tax completely on to its third-party sellers in the form of higher fees. A study by Deloitte about the ultimate effects of the French tax estimated that large digital businesses would be able to shift, on average, 96 percent of the tax burden to individual consumers and businesses who use their digital platforms. The study estimated that individual consumers would bear 57 percent of the ultimate tax burden. To the extent that the French tax is viewed as being imposed primarily on the final consumers, there would be no double taxation and thus no need to interpret the Treaty in such a broad manner.

Additionally, U.S. courts generally take a contractual, rather than statutory, approach to interpreting international tax treaties and, thus, attempt to interpret treaties to give effect to the parties’ intentions. The Vienna Convention was signed by the U.S. on April 24, 1970, but the U.S. Senate has never ratified the convention. The U.S. State Department notes that the U.S. “considers many of the provisions of the [Vienna Convention] to constitute customary international law on the law of treaties.” However, the U.S. Supreme Court “has only cited to [the Vienna Convention] twice and in an incidental fashion,” which leads some to question its authority in the U.S. Using the contractual approach favored by U.S. courts, the “primary object of interpretation is to ascertain the meaning intended by the parties rather than focus simply on the text.”

France is neither a party nor a signatory to the Vienna Convention, and French judges do not explicitly refer to the Vienna Convention when interpreting treaties. Although scholars have found it difficult “to determine positively the

212 Chandler, supra note 199.
213 Pellefigue, supra note 211, at 27–28.
214 Id. at 27.
216 Kysar, supra note 208, at 1402.
218 Kysar, supra note 208, at 1402.
219 Id. at 1397 (quoting RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW OF THE UNITED STATES § 325, reporters’ note 4 (A.L.I. 1986) (internal quotation marks omitted)).
220 Vienna Convention, supra note 209.
methods of interpretation of French judges,\footnote{Id. at 455 n.290 (quoting Emmanuel Decaux, Pierre Michel Eisemann, Valerie Goessel-Le Bihan & Brigitte Stem., France, in L’INTÉGRATION DU DROIT INTERNATIONAL ET COMMUNAUTAIRE DANS L’ORDRE JURIDIQUE NATIONAL: ÉTUDE DE LA PRATIQUE EN EUROPE 241, 269 (Pierre Michel Eisemann ed., 1996)).} French courts do consider the intent of the treaty framers\footnote{Id. at 450.} as well as the treaty’s purpose.\footnote{Id. at 455; Vienna Convention, supra note 209, art. 31 (stating that “[a] treaty shall be interpreted in good faith . . . in the light of its object and purpose.”).}

Using an intentionalism approach, it seems that the U.S.–France Treaty would not be read to apply to the French digital tax. The Treaty makes no mention of any consumption taxes levied by either the U.S. or France and does not seem to be intended to disrupt consumption taxes in either countries (for example, neither the French value added tax nor U.S. sales taxes are affected by the Treaty). Furthermore, it is generally understood that consumption taxes can be imposed by the state in which the consumption occurs.\footnote{Walter Hellerstein, Jurisdiction to Tax Income and Consumption in The New Economy: A Theoretical and Comparative Perspective, 38 GA. L. REV. 1, 52–53 (2003).} Therefore, it seems likely that the intent of the parties was that consumption taxes would fall outside of the scope of the U.S.–France Treaty and that each country would be free to impose consumption taxes without the restraints of the Treaty (and therefore without the permanent establishment requirement).

4. The French tax does not come within the scope of the U.S.–France Treaty and therefore does not violate the Treaty.

The French tax seems most aptly characterized as a consumption tax and therefore outside the scope of international tax treaties. Despite the counterarguments, it seems most likely that the French digital service tax does not constitute a tax on “business profits” under Article 7 of the Treaty. Under the typical meaning of “profits” and given the specific terms of the Treaty, the French digital services tax does not fall within this understanding. Even considering the Treaty’s inclusion of an “element of income” within the Treaty’s scope, it seems most likely that the tax imposed on revenue does not fall within the meaning of an element of income. Application of an intentionalism approach to interpreting the Treaty would seem to lead to the conclusion that the Treaty was not intended to reach consumption taxes such as a tax on revenues. The French digital tax seems to be most aptly characterized as a consumption tax on the provision of digital services that falls outside the scope of the U.S.–France Treaty and therefore is not in violation of international tax treaties.\footnote{See generally, Patricia A. Brown, Professor, U. Miami L. Sch., Do Tax Treaties Matter in a Post-BEPS World?, Remarks at The University of Chicago Law School’s 72nd Annual Federal Tax Conference, Session 1 (Nov. 8, 2019).}
The prevalence of digital services in the modern economy has generated controversy about where and how much digital businesses should be taxed. As multinational organizations, such as the OECD and EC, work toward agreeable solutions, many countries (notably France) have become impatient and continue to enact—or threaten to enact—their own unilateral taxes on digital businesses.\textsuperscript{227} Such unilateral action has drawn criticism for allegedly further complicating this already muddled area of international tax law, increasing compliance burdens, increasing double taxation, and adversely impacting global investments and technological developments.\textsuperscript{228} Additionally, scholars have noted that these digital taxes may violate existing law by violating international tax treaties, constituting illegal state aid, violating non-discrimination clauses of national constitutions, and/or violating the E.U.’s fundamental freedom to provide services.\textsuperscript{229} Questions of illegal state aid, constitutionality, and European fundamental freedom are beyond the scope of this Comment,\textsuperscript{230} but this Comment argues that the French digital tax does not violate existing international tax treaties. Because the French tax is a tax on revenues rather than profits, it is most aptly characterized as a consumption tax rather than an income tax and thus does not fall within the scope of the international treaties for the prevention of double taxation. Because many bilateral tax treaties are based on the OECD Model Tax Convention (including the U.S.–France Treaty), the analysis in this Comment can be applied to a wide array of international tax treaties and could be useful if other countries follow France in instituting taxes on digital businesses. Although there are still open questions about the legality of the French tax and similar taxes, it seems likely that they do not violate existing international tax treaties.

\textsuperscript{227} Wamsley, \textit{supra} note 5; OECD, \textit{Programme of Work}, \textit{supra} note 17, at 7.

\textsuperscript{228} OECD, \textit{Programme of Work}, \textit{supra} note 17, at 7.

\textsuperscript{229} Govindarajan et al., \textit{supra} note 23; Gottlieb & Ali, \textit{supra} note 23.