

NORTHEASTERN UNIVERSITY LAW REVIEW

VOLUME 10

NUMBER 1

ARTICLES

Editors' Introduction

- 1 Human Rights Due Diligence and the Root Causes of Harm in Business Operations: A Textual and Contextual Analysis of the Guiding Principles on Business and Human Rights
Radu Mares
- 72 Collaborative Enforcement
Andrew Elmore
- 143 The Supreme Court's Constitutional "Bright Line": Preempting Authority of 47 of 50 States
Steven Ferrey
- 204 The Cost of Justice: The Importance of a Criminal Defendant's Ability to Pay in the Era of *Commonwealth v. Henry*
Cristina Rodrigues
- 281 Comity in US Courts
Thomas Schultz & Niccolò Ridi
- 367 Contract as Reliance, Promise, and Consent
Caleb B. Koufman
- 401 ADR in Japan's Financial Markets & The Rule of Law
Tomohiko Maeda & Andrew M. Pardieck

NORTHEASTERN UNIVERSITY
LAW REVIEW

VOLUME 10, NUMBER 1

© 2018 Northeastern University Law Review

Northeastern University School of Law
416 Huntington Avenue
Boston, Massachusetts 02115

Cite as 10 Ne. U. L. Rev.____ (2018).

Editorial Board

Editors-in-Chief

Andrew L. Bardetti Moayad T. Al-Suwaidan

Managing Editors

Angela J. Benoit Emily H. Miller

Executive Articles Editors

Tristan Sullivan-Wilson April Herleikson

Articles Editors

Michael P. Dickman	Jameson Calitri
Laurielle M. Howe	John M. Mazzuchi
Luciana Rivera	Brandon Sisson

Publications Editors

Kara E. Grogan Catherine A. Maronski

Symposium Editors

Melissa Ramos Phoebe A. Roth

Online Law Journal Editors

Daniel S. Guenther Ali J. Walendziak

Online Law Forum Editors

Karina I. Guzman Sarah J. Butson

Advisors

Roger I. Abrams, *Richardson Professor of Law*
Sarah Hooke Lee, *Associate Dean & Director of
Information and Research Services*
Rashmi Dyal-Chand, *Professor of Law*

Senior Staff Editors

Maraya Best
Jennifer L. Boyd
Devan C. Braun
Gianna M. Casola
Liam M. Collins
Amy R. Grenier
Maya Fe Holzhauer

Lindsay L. Jacobsen
Onyemma C. Obiekea
Jaquelyn Perez
Carly A. Perkins
Karen Phung
Jessica L. Rubine
Katherine Soule

Staff Editors

- | | |
|-----------------------|-----------------------|
| Nuryllen Aguasvivas | Susmita Gadre |
| Anna Maria Annino | Elizabeth L. Gardon |
| Meagan Antonellis | Lilian Giacoma |
| Queen Arsem-O'Malley | Munroe Graham |
| Moira C. Barry | Shelby M. Hecht |
| Nicole Beggiani | Gary Howell-Walton |
| Phil Benaszek | Jacob R. Jarred |
| Stephanie H. Berland | Jeffrey W. Kennedy |
| Amanda M. Bishop | Andrew L. Kinde |
| Samantha E. Bresler | Lauren A. Kopec |
| Jason L. Burrell | Madison L.T. Lang |
| Ariel C. Chang | Kelli McIntee |
| Beverly A. Cheng | Rachel Mills |
| Kelly E. Cooke | Michael C. Murphy |
| Jennifer R. Cullinane | Keith A. Nemeth |
| Caroline S. Daniels | Mark Overton |
| Richard E. Darst | Patricia A. Perez |
| Maria J. de la Motte | Ryan McGovern Quinn |
| Elena J. Depotopoulos | Kyle P. Rainey |
| Edwin F. Diaz | Timothy A. Roberts |
| Maxwell S. Dismukes | Michael A. Roderick |
| Anastasia Doherty | Frank B. Scardino |
| Ethan J. Dowling | Nina Schwartzman |
| Bruce Duggan | Taarika C. Sridhar |
| Angélica Durón | Nour Sulaiman |
| Elizabeth C. Egan | Amanda L. Varrichione |
| Sarah A. Eskreis | Annie M. Vozar |
| Diandra Y. Franks | |

Non-Discrimination Policy

Northeastern University does not condone discrimination on the basis of race, color, religion, religious creed, genetics, sex, gender identity, sexual orientation, age, national origin, ancestry, veteran, or disability status.

Faculty

Jeremy Paul

Dean and Professor of Law

Martha F. Davis

*Professor of Law and Associate Dean for
Experiential Education*

Sarah Hooke Lee

*Associate Dean and Director of
Information & Research Services*

Wendy E. Parmet[†]

*Matthews Distinguished University
Professor of Law, Director of the Center
for Health Policy & Law and Associate
Dean for Interdisciplinary Education &
Research Support; Professor of Public
Policy & Urban Affairs, Northeastern
University School of Public Policy &
Urban Affairs*

Roger I. Abrams

Richardson Professor of Law

Libby S. Adler

Professor of Law

Aziza Ahmed

Professor of Law

Brook K. Baker[†]

Professor of Law

Shalanda H. Baker

*Professor of Law, Public Policy and Urban
Affairs*

Leo Beletsky

*Associate Professor of Law &
Health Sciences*

Elizabeth M. Bloom

Teaching Professor

Lee P. Breckenridge

Professor of Law

Margaret A. Burnham

Professor of Law

Dan Danielsen

*Professor of Law and Faculty Director,
Program on the Corporation, Law &
Global Society*

Richard A. Daynard

University Distinguished Professor of Law

Melinda F. Drew

*Teaching Professor and Director of the
Academic Success Program*

Christine M. Durkin

Associate Teaching Professor

Rashmi Dyal-Chand

Professor of Law and Research Director

Peter D. Enrich

Professor of Law

Julian M. Fray

Assistant Teaching Professor

[†] Masters in Legal Studies Faculty

Faculty (cont.)

Hemanth C. Gundavaram
Associate Teaching Professor

James R. Hackney, Jr.[†]
*Professor of Law and Chief of Staff
and Senior Strategy Advisor, Office of
President Joseph E. Aoun*

Margaret Hahn-Dupont
Associate Teaching Professor

Stephanie R. Hartung
Teaching Professor

Woodrow Hartzog
Professor of Law and Computer Science

Wallace E. Holohan
*Senior Clinical Specialist and Director of
the Prisoners' Rights Clinic*

Karl E. Klare
*George J. and Kathleen Waters Matthews
Distinguished University Professor of Law*

Kandace Kukas
*Assistant Dean and Director of Bar
Admission Programs*

Mary E. Landergan
*Associate Teaching Professor and Co-
Director of the IP CO-LAB*

Margo K. Lindauer
*Associate Teaching Professor and Director
of the Domestic Violence Institute*

Kristin M. Madison
Professor of Law and Health Sciences

Carol R. Mallory
Associate Teaching Professor

Andrea M. Matwyshyn
Professor of Law

Susan A. Maze-Rothstein
*Teaching Professor and Director of the
Legal Skills in Social Context (LSSC)
Program*

Victoria McCoy
*Assistant Teaching Professor and Director,
Academic Success Program*

Daniel S. Medwed
*Professor of Law & Criminal Justice
and Faculty Director of Professional
Development*

Michael Meltsner
*George J. and Kathleen Waters Matthews
Distinguished University Professor of Law*

Susan Barbieri Montgomery
Executive Professor of Law & Business

Mary E. O'Connell
Professor of Law

David M. Phillips[†]
Professor of Law

[†] *Masters in Legal Studies Faculty*

Faculty (cont.)

Jason Potter

Associate Teaching Professor

Ira Sills

Lecturer in Law

Deborah A. Ramirez

Professor of Law

Emily A. Spieler

Hadley Professor of Law

Sonia Elise Rolland

*Professor of Law and Faculty Director,
LLM and International Programs*

Kara W. Swanson

Professor of Law

Rachel E. Rosenbloom

Professor of Law

Lucy A. Williams

*Professor of Law and Faculty Director,
Center for Public Interest Advocacy and
Collaboration*

James V. Rowan[†]

*Professor of Law and Director of the
Clinical Programs*

Margaret Y.K. Woo

*Professor of Law and Associate Dean for
Research and Interdisciplinary Education*

Peter B. Sessa

*Associate Teaching Professor and Director,
Community Business Clinic*

Rose Zoltek-Jick

*Associate Teaching Professor and
Associate Director, Civil Rights and
Restorative Justice Project*

Jessica M. Silbey

Professor of Law

Adjunct and Visiting Faculty

Joshua L. Abrams

Helena Alviar[†]

Brooke Arlington

Mark N. Berman

Jay D. Blitzman

Robert G. Burdick

Peter Campia

Gary Cooper

Kyle Courtney

Joshua Davis

Fernande RV Duffy

Angela Dugar

Dawn Effron

Elizabeth Fahey

Patricia Garin

Mark Gottlieb

Ilana Greenstein[†]

Joshua S. Grinspoon[†]

Elliot Hibbler

Patricia Illingworth

[†] Masters in Legal Studies Faculty

Adjunct and Visiting Faculty (cont.)

Ivana Isailovic	Stephen McJohn
David Ismay	William Mostyn
R. Marc Kantrowitz	Keith R. Osher
Michael Keating	Mark Paige
Melvin Kelley	Michael Pezza, Jr.
Peter M. Kelley	Alex G. Philipson
Kenneth King	Arnold R. Rosenfeld
Stephen M. Kohn	Stuart Rossman
Jootaek Lee	Alfreda Russell
Neil T. Leifer	Robert (Rusty) Russell
Sofia Lingos	Yuliya G. Scharf
Liliana Mangiafico	Amy Remus Scott

Masters of Legal Studies Faculty

Patricia Davidson	Natacha Thomas
Thomas Madden	Dan Urman
Sean Nolon	Kevin F. Wall
Rebecca Rausch	Marsha White

Editors' Introduction

Since its founding, the Northeastern University Law Review has grown and developed in exciting ways. Originally a series of issues around symposia, and then existing in the topic-based format typical of “law journals,” the Law Review now publishes articles covering an array of legal topics and on multiple platforms. As always, the Law Review is proud to remain loyal to the production and promotion of articles related to the realm of public interest law.

Volume 10, Issue 1 contains articles on a myriad of subjects. Professor Radu Mares explores the notion of human rights due diligence and what impact it can have on human rights infringements in business operations. Professor Andrew Elmore considers “collaborative enforcement” as a way to address enforcement gaps between state and local agencies and private, public interest organizations. Professor Steven Ferrey analyzes the *Huges v. Talen Energy* Supreme Court decision, showing the nuances between field and conflict pre-emption in the electric energy context. This issue also features a note by Attorney Cristina Rodrigues that examines the *Commonwealth v. Henry* Massachusetts Supreme Judicial Court decision by exploring its consequences on the collection of court-imposed payments from poor defendants. Professor Thomas Schultz and PhD Candidate Niccolò Ridi trace the evolution of comity throughout history. Professors Andrew Pardieck and Tomohiko Maeda look at Japan’s financial alternative dispute resolution system, noting meaningful benefits and challenges. Finally, another note by NUSL Alum Caleb Koufman considers different contract doctrines in light of an informal private order. These five articles and two notes present cutting-edge scholarship and thoughtful considerations for the future of our legal system.

Each year, the Law Review hosts a symposium exploring a contemporary topic of interest in the law. This annual event brings together practitioners and academics from across the country together to begin discussions which advance our understanding of that topic. This year, the topic was “International Law, Local Justice” in honor of the late Professor Hope Lewis. It is with great pride that our Volume 10, Issue 2 contains the product of this symposium which explored the impact of Professor Lewis’ work and illuminated the role human rights principles can play in protecting people from the volatility and extractive excess of inadequately regulated global markets. For Professor Lewis, who was both a leading academic and an

activist, research and scholarship were meant to inform the work of activists and the struggling communities for whom they speak. She mentored and supported generations of students whose work as lawyers and advocates today gives life to her vision of human rights, racial equality, social justice and engaged practice. We believe the flagship student-run publication of Northeastern University School of Law—a school founded on the principles of social responsibility—is especially well situated to encourage such a dialogue.

In addition to our print publication the Law Review has gradually increased its online presence, and publishes high-caliber student and alumni pieces on two online platforms: Extra Legal and the Online Forum. On Extra Legal, the Law Review published a variety of student notes and essays covering such topics as the *Case against Nathan Carman*, the *Legal Burden Copyright Places on the Hip-hop Community*, and the *Developing Legal Standard in Massachusetts for Teacher-On-Student Harassment*. With the constant news cycle this past year, the Online Forum allowed the Law Review to publish relevant, timely blog posts such as the *Future of the SJC’s Barbuto Decision Under the Trump Administration*, the *State of Bail in Massachusetts Following the Brangan Decision*, and *SCOTUS Addressing Whistleblowing Protections in Dodd-Frank*.

Finally, this volume would not have been possible without the tremendous efforts of our Staffers and Senior Staffers. We are extremely pleased and proud to know the new structure of this Law Review is being left in such capable hands. This decision to reformat the Law Review, would not have been possible without the hard work, dedication, and guidance we received from our Faculty Advisors—Associate Dean Sarah Hooke Lee, Professor Roger Abrams, and Professor Rashmi Dyal-Chand. We would also like to thank Professor Wendy Parmet and Professor Margaret Burnham for their strenuous efforts in organizing the symposium honoring Professor Lewis and assisting with the second issue in her honor. Lastly, we would like to express our gratitude to Dean Jeremy Paul and the entire faculty and staff of Northeastern University School of Law for their continued support of the Law Review and its mission.

Editorial Board
Northeastern University Law Review

Human Rights Due Diligence and the Root Causes of Harm in Business Operation: A Textual and Contextual Analysis of the Guiding Principles on Business and Human Rights

*Radu Mares**

Abstract

Following John Ruggie's UN mandate (2005-2011) the notion of human rights due diligence (HRDD) has become widely used by policymakers, corporations, NGOs and professionals. The question is whether this HRDD concept, as developed in the UN Guiding Principles (GPs), adequately addresses the deeper causes of human rights infringements in business operations. This article provides a textual and contextual analysis of the GPs and related documents with a focus on the concepts of mitigation and root causes. The GPs contain provisions that open the door for HRDD to be interpreted for a less demanding result. There are also drafting imperfections. The GPs refer repeatedly to mitigation of impacts which introduces redundancy and ambiguity in an instrument prized for its clarity and simplicity. This analysis of the GPs addresses concerns that the GPs propose an overly process-oriented and risk-management approach that leaves business too much flexibility and discretion. A closer look reveals that mitigation in the GPs entails multiple meanings, functions and organizational contexts. A surprisingly multifaceted concept is placed at the center of HRDD. To realize its human rights protection potential, the notion of HRDD must impress with utmost clarity that HRDD cannot be merely about reducing abuses and applying bandaids on symptoms, but should aim for not less than elimination of infringements of human rights from a company's operations and should address the underlying, deeper causes of abuses. Clarifying mitigation will ensure the internal consistency of the GPs and present HRDD as a rightholder-centered risk management approach suited to the human rights context.

* Senior Researcher, Raoul Wallenberg Institute of Human Rights and Humanitarian Law, Lund, Sweden; Associate Professor, Faculty of Law, Lund University, Sweden, radu.mares@rwi.lu.se

Table of Contents

I.	Introduction	4
II.	"Mitigation" of Adverse Impacts Between Redundancy and Ambiguity	8
	A. Opportunity to Pick and Choose	10
	B. Silences	13
	C. When Opportunity Meets Silence.....	17
III.	Corporate Responsibility to Respect Under the Shadow of Internal Inconsistency	18
	A. The Flow From the Abstract to the Specific.....	18
	B. A Multilayered Concept of RtR	29
IV.	Mitigation Clarified	31
	A. Meanings of Mitigation	32
	B. Contexts of Mitigation	32
	C. Types of Adverse Impacts	33
	D. Preventive Measures	34
	E. Remedial Measures	35
	F. Functions of Mitigation	40
V.	"Root Causes" of Adverse Impacts.....	43
	A. Options for Tightening HRDD	43
	B. "Root Causes" in the GPs and Subsequent Documents... ..	54
	C. Implications for HRDD.....	63
VI.	Conclusions.....	67

I. Introduction

This article analyzes one of the key concepts the United Nations Guiding Principles for Business and Human Rights¹ (GPs) put forward: human rights due diligence (HRDD). HRDD is a familiar concept to business executives, corporate lawyers and international law specialists. The heritage of the notion is clearly traceable in international law,² in international human rights law,³ in national law (e.g., compliance with legal obligations by undertaking due diligence steps)⁴ and in business risk management (e.g., as a method to identify and mitigate risks to companies and investors).⁵ In this light, U.N. Special Representative John Ruggie's choice to model the social responsibilities of companies around this HRDD concept appears as an inspired one. Due diligence has entered the day-to-day language of business and human rights (BHR), and corporate social responsibility (CSR) more broadly, shaping the way companies, civil society groups, policymakers, and professionals

1 John Ruggie (U.N. Special Representative of the Secretary General), *Guiding Principles on Business and Human Rights: Implementing the United Nations “Protect, Respect and Remedy” Framework*, U.N. Doc. A/HRC/17/31 (Mar. 21, 2011) [hereinafter *Guiding Principles*].

2 Treaty on Claims, Fisheries, Navigation of the St. Lawrence, etc., American Lumber on the River St. John, Boundary, Gr. Brit-U.S., art. VI, May 8, 1871, 17 Stat. 863 (commonly referred to as the Alabama Arbitration case) (indicating that a neutral State is bound to “use due diligence to prevent the fitting out, arming, or equipping, within its jurisdiction, of any vessel which it has reasonable ground to believe is intended to cruise or to carry on war against a power with which it is at peace ...”). See generally JOANNA KULESZA, *DUE DILIGENCE IN INTERNATIONAL LAW* (Brill 2016).

3 Velásquez-Rodríguez v. Honduras, Judgment, Inter-Am. Ct. H.R. (ser. C) No. 4, ¶ 172 (1988) (stating that “[a]n illegal act which violates human rights and which is initially not directly imputable to a State . . . can lead to international responsibility of the State, not because of the act itself, but because of the lack of due diligence to prevent the violation or to respond to it as required by the Convention.”).

4 Securities Act of 1933, § 11, 15 U.S.C. § 77k (1933) (providing that experts who conducted “reasonable investigation” have a due diligence defense against buyers of securities that suffered damages from a faulty registration statement). See also William K. Sjostrom Jr., *The Due Diligence Defense Under Section 11 of the Securities Act of 1933*, 44 BRANDEIS L.J. 549 (2006).

5 Transactional due diligence covers issues such as “business development, regional analysis, legal, tax and accounting matters, the environment, health and safety (EHS), human resources, security, and procurement.” MARGARET WACHENFELD ET AL., *STATE OF PLAY: THE CORPORATE RESPONSIBILITY TO RESPECT HUMAN RIGHTS IN BUSINESS RELATIONSHIPS* 41, 54 (2012), <http://www.ihrb.org/pdf/state-of-play/State-of-Play-Full-Report.pdf>.

think and talk about corporate responsibilities regarding human rights. But as it is said, familiarity should not breed content. While the overwhelming amount of work regarding HRDD currently goes towards specification in distinct CSR contexts—through new standards and tools,⁶ guidance,⁷ adaptations to diverse industrial⁸ and issue contexts⁹—and also in incorporating this concept in various bodies of regulation,¹⁰ little attention goes to critically examining the soundness of HRDD as outlined within the GPs.

On the one hand, HRDD—as referenced and reinforced in other main CSR instruments¹¹—is a useful reference point to stakeholders challenging business¹² around basic prudential steps thus enabling both collaborative work and criticism (e.g. failure to conduct assessments, to report, to take elementary corrective measures). On the other hand, the question is whether the GPs add limited value and might, overall, set the bar too low: a HRDD that lacks sufficient specificity of the expected action, lacks a reviewer

6 *Human Rights Compliance Assessment*, DANISH INST. FOR HUMAN RIGHTS, <https://hrca2.humanrightsbusiness.org/>. See also *Human Rights Due Diligence Info Portal*, GLOBAL COMPACT NETWORK GERMANY, <http://mr-sorgfalt.de/en/> (launched in 2016 by the Global Compact Network Germany and twentyfifty).

7 See Shift & Mazars LLP, UN GUIDING PRINCIPLES REPORTING FRAMEWORK (2015), http://www.ungreporting.org/wp-content/uploads/2015/03/UNGPReportingFramework_2017.pdf.

8 See European Commission, *European Commission Sector Guides on Implementing the UN Guiding Principles on Business and Human Rights* (2017), https://ec.europa.eu/anti-trafficking/publications/european-commission-sector-guides-implementing-un-guiding-principles-business-and-hum-0_en (covering three sectors: oil and gas, information and communication technology and employment and recruitment agencies).

9 GWENDOLYN REMMERT ET AL., *RESPECTING HUMAN RIGHTS - AN INTRODUCTORY GUIDE FOR BUSINESS* 28–53 (Global Compact Network, twentyfifty Ltd. & German Institute for Human Rights 2013) (going right by right to exemplify expected business conduct).

10 French Parliament has been discussing since 2013 a legislative proposal to impose mandatory human rights due diligence on large companies. Proposition de Loi 1519 du 6 novembre 2013 relative au devoir de vigilance des sociétés mères et des entreprises donneuses d'ordre [Proposed Law 1519 of Nov. 6, 2013 relating to the duty of vigilance of parent companies and undertakings giving orders], ASSEMBLÉE NATIONALE, www.assemblee-nationale.fr/14/propositions/pion1519.asp.

11 ORG. FOR ECON. CO-OPERATION AND DEV., *OECD GUIDELINES FOR MULTINATIONAL ENTERPRISES* 31, 34 (5th ed. 2011).

12 “Business” and “company” are terms used interchangeably in this article, unless the context indicates otherwise.

able to determine whether a company exercised sufficient diligence, and seemingly speaks of *mitigation* (i.e. reduction) of impacts as an acceptable aim of HRDD. This article focuses on the notion of mitigation as a threat to the internal consistency of the GPs and as an opening for watering down HRDD in various ways such as insufficient effort, long timeframe for improvements, or inappropriate techniques used to manage impacts. The GPs' emphasis on HRDD might result in an overly process-oriented and risk-management approach that leaves companies too much flexibility and discretion, especially as legal obligations and complaint mechanisms are still largely absent.¹³

The inquiry here examines whether there is full alignment between the abstract, normative responsibility to respect (RtR) and the operational, risk-management inspired HRDD. Two main questions are raised. First, do the repeated references to *mitigation*, understood as a reduction of impacts, diminish the stringency of the RtR's prescription that companies should eliminate adverse human rights impacts from operations? Mitigation means reduction of impact, potentially indicating that residual impact is tolerated, temporarily or permanently.¹⁴ This is problematic from a human rights perspective which is geared fundamentally towards ensuring rights and eliminating impacts. Analysis of GPs should clarify in what circumstances such residual impact might be acceptable. Second, do the GPs sufficiently address the danger that HRDD will be understood and implemented superficially instead of tackling the *root causes* of adverse human rights impacts?¹⁵ This analysis aims to clarify whether and how the GPs and subsequent materials from

13 Gwynne Skinner, et al., *THE THIRD PILLAR: ACCESS TO JUDICIAL REMEDIES FOR HUMAN RIGHTS VIOLATIONS BY TRANSNATIONAL BUSINESS* 86–91, 105–129 (2013).

14 United Nations Office of the High Commissioner of Human Rights, *The Corporate Responsibility to Respect Human Rights: An Interpretive Guide*, 7, 18 U.N. Doc. HR/PUB/12/02 (2012) [hereinafter *Interpretive Guide*] (indicating that “[t]he mitigation of adverse human rights impact refers to actions taken to reduce its extent, with any residual impact then requiring remediation.”).

15 Sharan Burrow, *Global Supply Chains: What does Labour Want?*, OPENDEMOCRACY (May 31, 2016), <https://www.opendemocracy.net/beyondslavery/ilc/sharan-burrow/global-supply-chains-what-does-labour-want> (noting that MNEs “constantly look to cut costs to maximise profits, putting considerable pressure on workers’ wages and working conditions. . . . Corporate and multi-stakeholder ‘corporate social responsibility’ schemes have had little if any positive impact on guaranteeing workers’ rights, and instead have deferred addressing underlying structural issues.”).

the GPs' drafters orient HRDD towards tackling the root causes of impacts and thus make HRDD a tool inherently capable of helping to eliminate abuses from corporate operations.

There is potential for the misinterpretation of the GPs as the implementation rush takes place.¹⁶ By their nature, the GPs operate at a higher level of abstraction and generality given that it was meant as a foundational document applicable to all human rights, industries, and countries.¹⁷ Inevitably, misinterpretation will occur and this might have to do with the way stakeholders grasp the GPs' message. However, to some extent, this might have to do with the way the GPs have been formulated. To facilitate the implementation work already under way, it is desirable that the potential for misinterpretation be addressed, particularly because the non-binding GPs rely explicitly on their persuasive force to frame and set in motion dynamics that push forward BHR governance. A textual and contextual analysis of the GPs and related materials is warranted.

This article draws on reports from the UN Special Representative of the Secretary General on business and human rights (SRSG) mandate of John Ruggie (2005-2011) as well as post-mandate documents to which Ruggie is associated, mainly the Interpretive Guide¹⁸ and reports from Shift.¹⁹ In terms of structure, the article has four main sections. Section 2 presents the problem raised by mitigation in a human rights context. Section 3 analyzes the corporate responsibility to respect human rights and the drafting of Pillar 2 of the GPs from the general to the more specific provisions. Section 4 clarifies the appropriate context and role of mitigation in the GPs. Section 5 examines the issue of root causes of adverse impacts and documents how the GPs and materials produced in the post-mandate period account for this issue.

16 John Ruggie, *The Shift from Principles to Practice*, SHIFT, (Mar. 2012), <http://www.shiftproject.org/article/shift-principles-practice-0> (noting the "proliferation of activities" following the GPs creates a potential risk for "divergent interpretations" and of "los[ing] much of the convergence" the BHR field has gained).

17 See Interview by Ron Popper with John Ruggie, *Six Questions for John Ruggie: Where is the Business and Human Rights Agenda Going?*, MIRIN (Oct. 24, 2008), <http://mirinpucrio.blogspot.se/2008/10/seis-perguntas-para-john-ruggie.html> (commenting that "[t]he sheer intellectual magnitude of the task has been the most daunting challenge.").

18 See *Interpretive Guide*, *supra* note 14.

19 Shift is a US-based non-profit organization promoting the implementation of the GPs chaired by John Ruggie. SHIFT, <http://www.shiftproject.org/> (last visited Sept. 8, 2017).

II. “Mitigation” of Adverse Impacts Between Redundancy and Ambiguity

According to the GPs, businesses have a responsibility to respect human rights, meaning “they should avoid infringing on the human rights of others and should address adverse human rights impacts with which they are involved.”²⁰ That responsibility requires business action on three main fronts: policy, due diligence, and remediation.²¹ In turn due diligence is a process composed of four main steps including: “assessing actual and potential human rights impacts, integrating and acting upon the findings, tracking responses, and communicating how impacts are addressed.”²² Each of the four steps is detailed in one principle and associated commentary.²³ The corporate responsibility boils down to preventive and corrective measures to address “adverse impacts.”²⁴ Adverse impacts are composed of both “actual” impacts—harms “that [have] already occurred or [are] occurring”—and “potential” impacts—risks “that may occur but [have] not yet done so.”²⁵

The notion of mitigation understood as reduction of impacts is entirely absent from the 2008 Protect-Respect-Remedy Framework.²⁶ However it is used frequently in the 2011 GPs. For instance, the notion is used to clarify the aims of HRDD: “In order to identify, prevent, *mitigate* and account for how they address their adverse human rights impacts, business enterprises should carry out human rights due diligence.”²⁷ Interestingly, the The Office of the United Nations High Commissioner for Human Rights (OHCHR) Interpretive Guide offers an extensive definition of HRDD without referring to mitigation or the aims of HRDD.²⁸ It seems HRDD can be clarified with or without references to mitigation, which fuels the inquiry into whether there might be something inappropriate with

20 *Guiding Principles*, *supra* note 1, Principle 11, at 13.

21 *Id.* Principle 15, at 15.

22 *Id.* Principle 17 at 16.

23 *Id.* Principles 18–21, at 17–20.

24 *Id.* Principle 13, at 14.

25 *Interpretive Guide*, *supra* note 14, at 5–7.

26 See John Ruggie (Special Representative of the Secretary-General on Human Rights and Transnational Corporations and other Business Enterprises), *Protect, Respect and Remedy: A Framework for Business and Human Rights*, ¶ 51, U.N. Doc. A/HRC/8/5 (Apr. 7, 2008) [hereinafter *Protect, Respect and Remedy*].

27 *Guiding Principles*, *supra* note 1, Principle 17, at 16 (emphasis added).

28 *Interpretive Guide*, *supra* note 14, at 6.

stressing mitigation in a human rights context.

The problems raised by mitigation circle around its relationships with prevention and remediation. What is the value added by including the notion of mitigation? What does it add to what prevention and remediation already say? Does it clarify anything—in terms of the aim, measures, effort, timescale, context—for our understanding of RtR and HRDD? Are we in the presence of a mere redundancy from imperfect drafting of the GPs or a deeper ambiguity requiring better specification of the GPs? Both redundancy and ambiguity could be alleviated through textual and contextual interpretations of GPs.

While the added value of mitigation is unclear, as it lives under the specter of redundancy, there is a danger that it dilutes what prevention and remediation already impress: passivity is not an option under the RtR at the ex ante or ex post stages, and that diligence has to be exercised to address adverse impacts. The danger is that a potent and operational concept as HRDD might unwarrantedly reopen the aim dimension that the RtR clarified at the general and abstract level in Principle 11. As argued herein, the aim of RtR is the elimination of human rights abuses from a company's operations and value chains. The drafting of Pillar 2 needs to be examined to clarify the relation between RtR and HRDD. However, the GPs are much more than a statement of RtR and HRDD as they contain a blueprint for transnational governance to reach the goal of elimination of human rights abuses in business. This will take the analysis of mitigation towards clarifying RtR in global value chains, a fragmented and dispersed organizational setting characterized by multiple causes of harm and weak rule of law in less developed countries. This setting leads to complex issues of distributing responsibility for human rights within value chains: upwards or downwards, and among public and private actors. The result is a partial or shared responsibility in complex situations where multiple actors are involved and root causes are exceedingly difficult to tackle. With concepts like mitigation and leverage,²⁹ the GPs seek to find a path between full responsibility and no responsibility for lead firms in value chains as stakeholders search for smart policy mixes to govern value chains where lead firms are part of the solution and carry a reasonable burden. Therefore an analysis of mitigation in the

29 The GPs explain that “[e]verage is considered to exist where the enterprise has the ability to effect change in the wrongful practices of an entity that causes a harm.” *Guiding Principles*, supra note 1, Principle 19 Commentary, at 18.

GPs should interpret the text of the GPs in context, with sensitivity to business organizational contexts and the challenges specific to the human rights governance context.

A. *Opportunity to Pick and Choose*

The Commentary accompanying HRDD explains that “potential impacts should be addressed through prevention or mitigation.”³⁰ Difficulties arise because less demanding action might be expected under mitigation than prevention. However, the Commentary shows little concern for redundancy or effort to separate distinct contexts where the two terms might apply. This deficiency is compounded by the explanation offered in the Interpretive Guide. Therein, mitigation is about the aim of reduction as it covers actions taken to reduce the extent and/or likelihood of impacts.³¹ In turn, prevention is about the aim of non-occurrence: “the prevention of adverse human rights impact refers to actions taken to ensure such impact does not occur.”³² Overall, and with their concise formulation, the GPs arguably give businesses the opportunity to pick and choose between different aims encapsulated in prevention and mitigation.

It is one thing to say that HRDD will at times fail to prevent occurrence and impact³³ and another thing to give companies the choice between aiming for elimination/non-occurrence and reduction of impact. The implication would be either a benign redundancy or a danger of watering down the aim of RtR. As to the redundancy, mitigation might mean nothing more than taking preventive and corrective measures. As to the danger, the appeal of the operational concept of HRDD over the more abstract RtR cannot be overlooked.³⁴ RtR, which carries the aim of elimination and appears as the abstract, general, and aspirational part, might become downplayed with attention going to HRDD, which speaks of mitigation and appears operational, more specific, and more familiar

30 *Guiding Principles*, *supra* note 1, Principle 17 Commentary, at 16.

31 *Interpretive Guide*, *supra* note 14, at 7.

32 *Id.*

33 *Guiding Principles*, *supra* note 1, Principle 22 Commentary, at 20.

34 Radu Mares, ‘Respect’ Human Rights: Concept and Convergence in LAW, BUSINESS AND HUMAN RIGHTS: BRIDGING THE GAP 38–41 (Robert C. Bird et al. eds., 2014) (arguing that HRDD overtook RtR as the key concept behind paragraph 13b where the company is responsible to act even if it did not cause/contribute to impacts in third party operations. HRDD was instrumental in expanding the scope of RtR when comparing the 2008 Framework Report with the 2011 GPs operationalizing the 2008 Report).

to risk managers and decision-makers. Such reversal between the operational HRDD and the normative RtR makes the specificities of the human rights context harder to grasp and eventually misdirects the HRDD risk management process.

Whether companies can pick and choose between reduction and non-occurrence of impact is troublesome because the human rights context has its specificities. Reduction means *residual impact* on human rights in business operations is a given and is an accepted social cost of doing business, whether only temporary or even permanently. However, the human rights perspective does not accommodate residual impact easily. The normativity of international human rights standards aims for elimination of infringements of such rights, not for their reduction. First, human rights instruments always emphasize elimination by referring to rightholders through terms such as “Everyone,” “No one,” “All [persons],” and so on.³⁵ These instruments do not say “some people” or a “majority” to aim merely for reducing infringements of human rights. Second, by their very nature, human rights are *minimum* standards of treatment that are meant to ensure human dignity.³⁶ Just offering fractional protection does not make sense. So it is not just formulations in treaties but the entire normative thrust of human rights standards that indicates elimination rather than reduction as the sole legitimate aim.³⁷ Surely, under international law, state obligations correlating to human rights are at times formulated with qualifiers like progressive realization and the maximum available resources; for instance, a UN Covenant lays down the obligation that “[e]ach State Party to the present Covenant undertakes to take steps . . . to the maximum of its available resources, with a view to achieving progressively the full realization of the rights recognized in the present Covenant”³⁸

35 G.A. Res. 217 (III) A, Universal Declaration of Human Rights (Dec. 10, 1948).

36 See *id.* at art. 1 and 22.

37 U.N. Office of the United Nations High Commissioner for Human Rights, *Principles and Guidelines for a Human Rights Approach to Poverty Reduction Strategies*, at 25-26, U.N. Doc. HR/PUB/06/12 (2012), <http://www.ohchr.org/Documents/Publications/PovertyStrategiesen.pdf> (explaining how the human rights based approach emphasizes elimination, see for instance how the “key targets and indicators” are formulated).

38 G.A. Res. 2200 (XXI), International Covenant on Economic, Social and Cultural Rights at Article 2.1 (Dec. 16, 1966) (“[e]ach State Party to the present Covenant undertakes to take steps, individually and through international assistance and co-operation, especially economic and technical, to the maximum of its available resources, with a view to achieving progressively the full realization of the rights recognized in the present Covenant”).

Furthermore, the Sustainable Development Goals,³⁹ which often replicate human rights aspirations, contain aims of significantly reducing the quantity⁴⁰ as well as aims of outright elimination.⁴¹ This, however, does not signify tolerance of residual harm. Rather, it is a twofold acknowledgement that first, there is no magic wand to instantaneously eliminate infringements of human rights and second, maximum efforts must be pursued and even measured against precise timescales to ensure those rights.⁴²

The inherent aim of human rights standards can only be elimination of abuses. Duty-bearers take measures that should have the capacity and design to achieve that aim, and not a lesser aim of reducing abuses, which entails acceptance of residual harm.⁴³ In the

39 See generally G.A. Res. 70/1, Transforming Our World: The 2030 Agenda for Sustainable Development (Sept. 25, 2015).

40 For example, one aim is “[b]y 2020, halve the number of global deaths and injuries from road traffic accidents” (Goal 3.6). G.A. Res. 70/1, at 16 (Sept. 25, 2015). Or “[b]y 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty” *Id.* at 15. Another aim is “[b]y 2030, reduce the global maternal mortality ratio to less than 70 per 100,000 live births” *Id.* at 16. Another aim is “[b]y 2030 . . . substantially reduce the number of people suffering from water scarcity” *Id.* at 18.

41 For example, one aim is “[b]y 2030, eliminate gender disparities in education” *Id.* at 17. Or, “[e]liminate all forms of violence against all women and girls” *Id.* at 18. Another aim is “[t]ake immediate and effective measures to eradicate forced labour, end modern slavery and human trafficking and secure the prohibition and elimination of the worst forms of child labour, including recruitment and use of child soldiers, and by 2025 end child labour in all its forms.” *Id.* at 20.

42 Mariëtte van Huijstee et al., *How to Use the UN Guiding Principles on Business and Human Rights in Company Research and Advocacy*, SOMO (Nov. 2012), <https://www.somo.nl/wp-content/uploads/2012/11/English-version.pdf> (commenting on residual impact as follows: Mitigation of adverse impacts “refers to actions taken to reduce the extent of an impact by a third party, with any residual impact then requiring remediation. This *should not be interpreted as meaning that it is appropriate for a company to reduce a human rights abuse by a third party to some extent and that a little bit of remaining abuse is acceptable*. All the company’s efforts should be focused on ceasing the human rights impact by the third party that is causing the harm, but since the company does not have full control over the third party, a full stop cannot be guaranteed by this company.”) (emphasis added).

43 ON COMMON GROUND CONSULTANTS INC., HUMAN RIGHTS ASSESSMENT OF GOLDCORP’S MARLIN MINE, CANADA, at 83 (Susan Joyce, et. al. eds. 2010), http://hria-guatemala.com/en/docs/Human%20Rights/OCG_HRA_Marlin_Mine_May_17.pdf (“Some impacts to the environment may be tolerated from a human rights perspective so long as these are within the established standards for human health. The overarching standard for

GPs, both states and businesses are asked to respect human rights.⁴⁴ Notably though, states and companies have different options to achieve elimination of human rights abuses from their jurisdiction or operations. Thus, companies have the option of eliminating abuse from their operations by either protecting rightholders or separating themselves from the rightholders, that is, by cutting links with abusive third parties.⁴⁵ In contrast, sovereign states only have the option of increasing the protection for their populations (rightholders) to meet their elimination goal. The GPs are more than a code of corporate conduct; they are meant as a foundational treatment of transnational governance insights into how overlapping responsibilities of companies and states relate to each other in global value chains. Therefore, the GPs have no space for ambiguity on such a fundamental dimension as the aim of RtR and HRDD. In human rights, speaking of elimination, as opposed to mitigation, is an indication that doing business or governing through abusing rights (repression) is inherently illegitimate and cannot be accepted as a systemic feature of business or government.

Referring to measures that reduce impact should not obscure the aim of elimination of abuse. Indeed, the GPs sensibly refer to “measures” for the prevention, mitigation, and remediation when explaining the corporate responsibility to respect human rights,⁴⁶ but the aim of elimination risks being lost as references to mitigation and due diligence pile up and the complexities of protecting rights in global value chains become apparent. Insisting on the aim of elimination is not simply a matter of formulation where one might lean toward either the aspirational or the pragmatic. It is a matter of directing effort in the proper direction and designing governance responses capable of delivering on that aim. It is about an effort to identify, reflect, and address root causes of abuse, not about complacent and half-hearted efforts aimed at tackling symptoms with simplistic mitigation measures. It is about detecting misdirected and substandard business measures that do not comply with the GPs and are inappropriate in the human rights context, but might still be promoted as in line with HRDD and RtR.

B. Silences

human rights is an absolute: ‘do no harm.’”).

44 See *Guiding Principles*, *supra* note 1, Principle 6, at 10.

45 See *id.* Principle 19 Commentary, at 18.

46 *Guiding Principles*, *supra* note 1, Principle 11 Commentary, at 13.

The above ambiguity resulting in pick and choose options is further compounded by some silences in the GPs. The GPs neither indicated specific actions tailored for the particularities of each human right nor guided inquiries into the reasonable amount of effort a business should put in addressing impacts. Thus, on the one hand, rather than pursuing specificity, Ruggie opted for a foundational document applicable to all contexts and rights, admittedly a general but necessary document in the deeply contested field of BHR.⁴⁷ On the other hand, Ruggie did not make the GPs dependent on a monitoring or adjudicative body: that would have added not only teeth to the GPs, but would have also endowed them with a mechanism to specify HRDD in context by establishing facts and assessing them against a normative standard of reasonableness. Through precedent, HRDD would have been progressively clarified in an authoritative manner. These are limitations of the GPs that were well accepted by the drafters and that arguably do not detract from the GPs' value added.⁴⁸ But without ex ante or ex post ways of specifying HRDD, the risk of indeterminacy about the due diligence effort expected arises. Indeed, fine assessments of due diligence, reasonableness, and prudence cannot be done in the abstract and specification is inherently difficult ex ante; due diligence can only be assessed in light of specific circumstances of the case⁴⁹ for which a reviewer—such as a court or other investigator—is indispensable. Mindful of the failed precedent of the UN Norms that were built on the standard-reviewer coupling, Ruggie instead sought a persuasive account of the BHR field by combining public and private governance in a “polycentric governance” system⁵⁰ and harvesting the leverage of multiple stakeholders. So the GPs are the results of a mandate aiming to deliver a widely agreed upon document on which subsequent efforts to specification and accountability could build.

47 See JOHN GERARD RUGGIE, *JUST BUSINESS: MULTINATIONAL CORPORATIONS AND HUMAN RIGHTS* xi-xii (W. W. Norton & Co. 2013).

48 Radu Mares, *Business and Human Rights After Ruggie: Foundations, the Art of Simplification and the Imperative of Cumulative Progress*, in THE UN GUIDING PRINCIPLES ON BUSINESS AND HUMAN RIGHTS – FOUNDATIONS AND IMPLEMENTATION 1-50 (Radu Mares ed., 2012) (outlining the ways in which the SRSG contributed to the BHR debate and the reference point that GPs have become among other CSR standards).

49 *Interpretive Guide*, *supra* note 14, at 6.

50 John G. Ruggie, *Life in the Global Public Domain: Response to Commentaries*, JAMES G. STEWART BLOG (Feb. 5, 2015), <http://jamesgstewart.com/author/john-g-ruggie> (explaining polycentric governance).

The question however is whether the problematic presence of *mitigation* negatively affects the HRDD notion, which is already under the specter of indeterminacy given that the GPs are silent on the amount of effort and specific measures. HRDD is grounded in risk management, which brings familiarity and increases the traction of the GPs with businesses executives.⁵¹ However, the GPs emphasize that HRDD refers to risks to rightholders and not merely risks to business.⁵² Given that HRDD address this specific type of risk, some options raised by risk management do not apply in the human rights context. The enterprise will not have the liberty to choose among different risk management techniques that it has for other risks – to simply take the risk and do nothing, to reduce the risk to a tolerance level the company might deem acceptable based on its calculations, or to insure against it.

The GPs offer valuable insights in each of these situations. First, taking the risk and doing nothing is not acceptable; a company that has a “crucial” relationship with a third party might choose to continue dealings as opposed to terminating the relationship, but the GPs require that “[i]n any case, for as long as the abuse continues and the enterprise remains in the relationship, it should be able to demonstrate its own ongoing efforts to mitigate the impact and be prepared to accept any consequences—reputational, financial or legal—of the continuing connection.”⁵³ Similarly, taking the risk and doing nothing about risks posed by third party operations while expecting that the government would detect and act to reduce harm is similarly unacceptable: “The corporate responsibility ‘exists independently of States’ abilities and/or willingness to fulfil their own human rights obligations . . . [and] exists over and above compliance with national laws and regulations protecting human rights.”⁵⁴ Second, reducing risks to a tolerance level chosen by the company cannot be informed solely by calculations on risks to business because HRDD is about risks to human rights and the GPs repeatedly emphasizes transparency and participation in identifying

51 See RUGGIE, *supra* note 47, at 141–8.

52 *Guiding Principles*, *supra* note 1, Principle 17 Commentary, at 16 (“[H]uman rights due diligence can be included within broader enterprise risk-management systems, provided that it goes beyond simply identifying and managing material risks to the company itself, to include risks to rights-holders.”).

53 *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 18.

54 *Id.* Principle 11 Commentary, at 13.

and addressing impacts.⁵⁵ Third, there is no insurance to be gained from “offsetting” impacts as the GPs indicate that good deeds in one area (e.g. social contributions, economic development) cannot offset adverse impacts on human rights in another area: “Business enterprises may undertake other commitments or activities to support and promote human rights, which may contribute to the enjoyment of rights. But this does not offset a failure to respect human rights throughout their operations.”⁵⁶ Also, adherence to process offers no insurance in the event harm occurs. There are no guarantees that harm will be excused if the company followed process (process legitimacy), and there is not a full and automatic defense based on HRDD: “business enterprises conducting such due diligence should not assume that, by itself, this will automatically and fully absolve them from liability for causing or contributing to human rights abuses.”⁵⁷

Where the GPs speak of *prevention* and *mitigation* there is the danger that companies will see this twofold conceptualization—prevent and mitigate—as giving a choice between the two and consider mitigation as being fulfilled by any of the usual risk management techniques. The GPs counteract this danger by offering parameters of HRDD throughout the GPs.⁵⁸ This is an important contribution towards presenting HRDD for what it genuinely is: a rightholder-centered risk management approach. Are these parameters sufficiently emphasized and exhaustively identified? This article highlights the troublesome implications of *mitigation* for the GPs, emphasizes the above parameters of HRDD, and adds a new general parameter of *root cause* orientation for HRDD to reduce its relative indeterminacy in a foundational instrument like the GPs.

55 For example, Principle 18 holds that “business enterprises should identify and assess human rights impacts through ‘meaningful consultation with potentially affected groups and other relevant stakeholders.’” *Id.* Principle 18 at 17. Principle 20 explains how the tracking of performance should “draw on feedback from both internal and external sources, including affected stakeholders.” *Id.* Principle 20 at 20. Principle 21 holds that business should be transparent about their efforts and performance as such communication would provide “transparency and accountability to individuals or groups who may be impacted and to other relevant stakeholders, including investors.” *Id.* Principle 21 at 20.

56 *Id.* Principle 11 Commentary, at 13.

57 *Id.* Principle 17 Commentary, at 16–17.

58 E.g., *id.* Principle 17, at 16. See also Section III.A.

C. When Opportunity Meets Silence

In sum, the concerns about the appearance of mitigation in the very definition of HRDD have to do with residual impact, and how the GPs manage to handle this residual harm. This requires analysis of the GPs for internal consistency (coherence of GPs dimension). It also requires reflection about how HRDD applies in a diversity of organizational contexts where abuses have to be eliminated from global value chains (organizational dimension) and how an instrument like GPs seeks to drive change in a global governance context through regulatory and non-regulatory means (governance dimension). With clarity on these three dimensions, HRDD will genuinely appear as a distinct, rightholder-centered risk management approach.

By bringing together the three dimensions, the specter of indeterminacy hovering over HRDD in the GPs will be reduced. This requires highlighting both the parameters contained in the GPs as well as reflecting on the options for further specification of HRDD in the post-GPs period, on which Ruggie counted.⁵⁹ The insights and general approach to BHR taken in the GPs will become clearer. In turn this will hinder self-serving interpretations and superficially compliant measures by businesses accustomed to regular risk management techniques.⁶⁰ Insisting that the reasonable efforts of businesses are channeled in the right direction (not merely at the symptoms but at the root causes of abuse), the spotlight is unavoidably put on the actions and inactions of other public and private stakeholders who have a bearing on global value chains. This can only facilitate reflection on the complementary or shared

59 *Guiding Principles*, *supra* note 1, at 5 (“[The Guiding Principles] will mark the end of the beginning: by establishing a common global platform for action, on which cumulative progress can be built, step-by-step, without foreclosing any other promising longer-term developments.”).

60 Int'l Trade Union Confederation, *The United Nations “Protect, Respect, Remedy” Framework for Business and Human Rights and the United Nations Guiding Principles for Business and Human Rights - A Guide for Trade Unionists*, 15 (May 2012) (illustrating that “[s]ome companies are claiming that their existing CSR initiatives or practices are the ‘due diligence’ that they need. Indeed, there is a real danger that companies will recast existing CSR practices as ‘best practice’ applications of the Guiding Principles. Many of these claims and ‘best practice’ examples are not credible. Often they overlook important ideas in the Guiding principles: that the amount of due diligence must be related to risk and the severity of the impact for instance. The fundamental obligation to avoid adverse impacts is often absent in the ‘best practices’ as is remediation which is often confused with philanthropy.”).

responsibilities of public and private actors to act for further specification of HRDD in context (i.e. define the reasonable, prudent company's conduct in the circumstances) and stronger enforcement of corporate responsibilities and state obligations (i.e. through monitoring and remedial mechanisms).

The next section contains an analysis of Pillar 2 and the multilayered concept of RtR across the three above-mentioned dimensions.

III. Corporate Responsibility to Respect Under the Shadow of Internal Inconsistency

Pillar 2 of the GPs—corporate responsibility to respect human rights—offers a multilayered concept of RtR with HRDD being the core operational concept. Pillar 2 offers layers of explanations regarding what is expected from companies in terms of aims, required actions, and operational contexts. This is a valuable treatment of RtR moving progressively to deeper levels of specificity providing, in the process, practical guidance as well as indispensable clarifications about the nature of the tasks at hand. In this way, Pillar 2 offers not only a wealth of actionable prescriptions but also dispels misunderstandings that could creep in at different junctures. A 100-page Interpretive Guide followed shortly after the release of the GPs and further augments these contributions.⁶¹

With the concerns about the definition of HRDD in Principle 17 noted above, this section expands analysis to other principles in Pillar 2, particularly Principles 11, 13, 15 and 19. Attention goes to the specific dimensions captured in each principle and the flow from the general to the more concrete. This will reveal whether inconsistencies exist, whether formulations could be interpreted for greater clarity and better flow, and why current difficulties appeared. Attention will be paid to the aim dimension of RtR (the aim of elimination or reduction of impact), organizational context (own conduct or conduct of third parties), and responsibilities of multiple actors in value chains (lead firms, suppliers, governments).

A. *The Flow From the Abstract to the Specific*

Principle 11

The responsibility to respect is the most general level

61 *Interpretive Guide*, *supra* note 14, at 3–4.

at which the GPs begin addressing the issue of corporate social responsibilities. Principle 11 uses a two-pronged structure to explain *respect* as “avoid infringing” rights and “address” adverse impacts with which a company is “involved.”⁶² This two-pronged explanation is a change from the initial single-pronged definition used in the 2008 Framework report which wrote: “To respect rights essentially means not to infringe on the rights of others—put simply, to do no harm.”⁶³ Going for the two-pronged definition increases complexity, however, and can foster misinterpretation as two dimensions are overlapping: the first dimension is the *aim* (do not infringe) and the second dimension is the *action* of a proactive and reactive type expected (address impacts). In other words, “avoid” refers to what to aim for (elimination rather than something less demanding like reduction, which allows for residual harm), while “address” refers to what to do (act to minimize risks ex ante and take corrective/remedial measures ex post). The two dimensions are consistent but could easily skip out of alignment.

The drafting of the GPs is a step back in simplicity from the 2008 Framework which had just one concept—“do no harm”—that relayed simply the aim of RtR and captured both *aim* and *act* in that short formulation.⁶⁴ This drafting choice is explainable by what appears as Ruggie’s overriding concern for the RtR to be understood not merely as a negative obligation—not infringe, no harm—to refrain from action. Ruggie emphasized the RtR is not a negative obligation.⁶⁵ Understanding the RtR as a responsibility to refrain, to remain passive would have shortcut the entire message of HRDD as a proactive, systematic process of identifying and addressing impacts.

Between “avoid infringing” and “do no harm” (or “not to infringe”) there is full equivalence and both point towards elimination of human rights abuse from business operations. So, the aim of the RtR is clear and unchanged between 2008 and 2011: the RtR is about the elimination of human rights impacts from an enterprise’s operations and taking preventive and corrective actions

62 *Guiding Principles*, *supra* note 1, Principle 11, at 13.

63 *Protect, Respect and Remedy*, *supra* note 26, ¶ 24.

64 *Id.*

65 *Id.* ¶ 55 (clarifying that “‘doing no harm’ is not merely a passive responsibility for firms but may entail positive steps—for example, a workplace anti-discrimination policy might require the company to adopt specific recruitment and training programmes.”).

able to achieve this aim. This point regarding the aim of RtR should not be obscured or watered down as RtR is explained at more concrete levels in subsequent principles.

Principle 13

Principle 13 explains what “involved,” mentioned in Principle 11, means: the company is involved when it caused, contributed, or is linked to impacts through its business relationships.⁶⁶ This dimension clarifies the scope of business involvement. However, Principle 13 is drafted in a complex way that overlaps several additional dimensions. Besides stating *when to act* (cause/contribute/linkage), Principle 13 explains generally what to do in each situation: “avoid causing or contributing . . . and address” in paragraph 13a and “seek to prevent or mitigate” in paragraph 13b.⁶⁷ Thus, it adds some specificity to the dimension of *how to act*. Furthermore, Principle 13 also uses formulations seemingly indicating different levels of stringency of responsibilities (“avoid” versus “seek to”).⁶⁸ It seems this formulation adds a third dimension as it qualifies the *amount of effort* expected from the enterprise, and sets the bar for the paragraph 13b situation a notch lower. Several observations could be made regarding this complex drafting of Principle 13.

First, there are two conceptual faultlines in Principle 13. The first faultline is along the actions expected from a company depending on whether it caused harm or not: paragraph 13a covers involvement by causation (company conduct is the sole cause, respectively one of the causes of impact) while paragraph 13b expressly covers impact where the company has not caused impact in any way, but where RtR applies merely because of association with a wrongdoer.⁶⁹ This is the *causality faultline*. The second faultline is along own conduct and the conduct of a third party. Different ambitions in terms of elimination or reduction of impacts attach to each context: for own

66 *Guiding Principles*, *supra* note 1, Principle 13, at 14.

67 *Id.*

68 *Id.*

69 The responsibility under paragraph 13b—a responsibility by association—is not as expansive as it sounds. The company has a responsibility to act only for those impacts directly linked to its products, services and operations, and not for all impacts of its business partners. A further limitation results from the company not having to remediate the impacts, but just to prevent and mitigate its involvement first through leverage, and then by cutting links. See *Guiding Principles*, *supra* note 1, Principle 19, at 21–22; see also discussion *infra* section III.A.

conduct the company can eliminate the impact, but for the conduct of third parties, the company might simply not be able to eliminate the impact due to lack of control. This is the *organizational faultline*. How the GPs square this difficulty and reconcile it with the RtR's aim of elimination becomes evident only in Principle 19 depicting "appropriate action" (leverage to minimize impacts is expected, and as last resort, termination of relationship).⁷⁰

Second, the question is why Principle 13's drafting was so ambitious to combine multiple dimensions with multiple organizational contexts. The fact that Principle 13 covers clearly with its structure the first faultline (causality faultline) is not only appropriate but of extreme importance for BHR: it does not draw the boundaries of RtR so narrowly to leave out value chain responsibility in a time where global businesses pursue expansion strategies based on outsourcing, subsidiarization and joint ventures.⁷¹ The second faultline (organizational faultline) cuts the same pie differently, but the formulation exposed by Principle 13 cannot keep up. Indeed, mitigation is equally relevant in the "contribute" (13a) and "linkages" (13b) scenarios—as shown clearly by Principle 19—but is only mentioned in the latter (13b).⁷²

Table 1

Principle 13	Scenario	Action (GP 13 & 19) (Aim of RtR Dimension)	Third party (Organizational Dimension)	Involvement (Causality Dimension)
Para. 13a	Cause	Eliminate (by minimizing impact)	-	Causality
Paras. 13a 13b	Contribute	Eliminate(same) and Mitigate (by reducing impact)	Third party	Causality
Para. 13b	Linkages	Mitigate (same) and Eliminate (by cutting links)	Third party	Association

70 *Guiding Principles*, *supra* note 1, Principle 19, at 20–21.

71 See U.N. CONFERENCE ON TRADE AND DEV. [UNCTAD], *World Investment Report 2013: Global Value Chains: Investment and Trade for Development* 27 (2013); ILO, *Decent work in global supply chains*, International Labour Conference, 105th Session (2016) http://www.ilo.org/wcmsp5/groups/public/---ed_norm/---relconf/documents/meetingdocument/wcms_468097.pdf.

72 *Guiding Principles*, *supra* note 1, Principle 13, at 14.

The current formulation of Principle 13 does not accommodate both these organizational and causality dimensions. The structure of Principle 13 (paragraphs a and b) is not problematic: the drafters chose the causality dimension over the organizational one. That would have been perfectly fine if the action expected from the company was left at a general level instead of attempting to specify it. However, the drafting of Principle 13 attempts just that—specificity—resulting in undesirable consequences.

One negative consequence is that by being overly ambitious in drafting and adding two other dimensions, Principle 13 ends up endangering the aim dimension exposed in Principle 11. The complex drafting of Principle 13 and the appearance of *mitigation* introduce a tension with the aim of elimination in RtR. By speaking of avoiding, addressing, preventing, and mitigating, Principle 13 reopens the aim dimension with this complex terminology. Paragraph 13a speaking of “avoid causing or contributing” and Principle 11 speaking of “avoid infringing” dovetail; “avoid” became the vector for “do no harm” in Principle 11 and gave the RtR the elimination goal. The fact that the drafting structure of Principles 13a and 11 are identical (“avoid” and “address”) while paragraph 13b uses a new terminology (“seek” to “mitigate”) compounds the perception that the responsibility under paragraph 13b is the odd one out, the less demanding one. The dissonance in the drafting of paragraphs 13a and 13b might indicate that something has happened to the responsibility in 13b and that might have to do with the very goal of elimination. And mitigate is the vector carrying that message. It is true that paragraph 13b is less demanding regarding remediation,⁷³ but that should not encourage interpretation that the responsibility is less demanding for other dimensions, such as the aim dimension.

Another negative consequence is an imprecise, needless and premature specification of expected actions in cause, contribute and linkages scenarios. Mitigating adverse impacts makes an appearance in this effort to specify expected action. Principle 13 crucially settles the scope dimension of RtR, that is, the situations where the RtR exists. However, it is not clear why Principle 13 had to be so complex, so ambitious. And all this in a context where paragraph 13b is already the weakly justified area of RtR as it runs counter to general legal principles of liability based on fault and causality.⁷⁴

73 Thus, no remediation is expected from a company merely linked to harm; the remedy component of RtR is dropped in paragraph 13b. *Id.* at 16, 24.

74 See Radu Mares, *Responsibility to Respect: Why the Core Company Should Act When*

Principle 13 adds a second layer of clarification to the RtR, but appears counterproductive, premature, and imprecise in spelling out the more concrete actions businesses should take in varying contexts. There are subsequent principles in the GPs that offer such concrete guidance: Principle 19 (discussed further below).

Principle 15

Principle 15 begins the essential discussion of the main elements and operational steps that the RtR requires. The GPs are the result of the Human Rights Council's request for operationalization of the Protect-Respect-Remedy Framework of 2008.⁷⁵ The SRSG fulfilled this task of operationalization by delivering structure, simplicity, and absence of jargon to a field ripe with ambiguities and controversies.⁷⁶

Principle 15 highlights the three main components of RtR a company should have in place: policy, due diligence systems, and a process for remediation of harm. In this way, Principle 15 gives concreteness to Principle 11's abstract requirement to "address" human rights impacts. In addition, this Principle indicates that a company does not have to offer remediation in the Principle 13b situation (but only when the company caused or contributed to impact, paragraph 13a). Principle 15 sets the stage for the next operational Principles 16-22 and the flow to these principles is commendable. It is a Principle covering one dimension only and accomplishes this task with clarity.

Principle 17

Principle 17 expands on the definition of HRDD already offered in Principle 15 as a "process to identify, prevent, mitigate and account for how they address their impacts on human rights."⁷⁷ It covers well, just like Principle 15 did, the operational steps expected from a diligent company and outlines four such steps: impact assessment, appropriate action, tracking, and communication.⁷⁸ Each step will be allocated one dedicated principle (Principles 18-

Affiliates Infringe Rights in THE UN GUIDING PRINCIPLES ON BUSINESS AND HUMAN RIGHTS – FOUNDATIONS AND IMPLEMENTATION 1-50 (Radu Mares ed., 2012).

75 Human Rights Council, Res. 8/7 U.N. Doc A/HRC/RES/8/7 (June 18, 2008).

76 Mares, *supra* note 48.

77 See *Guiding Principles*, *supra* note 1, Principle 17, at 16–18.

78 *Id.* at 16–17.

21), while Principle 17, with its lengthy Commentary, reflects on some characteristics (“parameters”) of HRDD: its on-going nature (as opposed to one-off transactional due diligence (DD)), its adaptation to operational context, its relation to broader enterprise risk-management systems, the possibility of prioritizing action in the value chain context, HRDD as a defense, and secondary liability aspects of complicity.⁷⁹ Such useful observations are continued in Principles 23-24.

The difficulties with Principle 17 are twofold. First, the Principle uses the concepts of prevention and mitigation side by side in an enumeration displaying an inflation of terms with uncertain value added. Second, the Commentary’s discussion of potential and actual impacts⁸⁰ unhelpfully links mitigation to potential impacts only—this does not aid the understanding of either prevention or remediation in the GPs.

Regarding the first aspect, one wonders what is gained by this inflation of terms—prevention and mitigation—in the very definition of HRDD in a document prized for simplicity and clarity. By comparison, the 2008 report explained HRDD as “a process whereby companies . . . manage the risk of human rights harm with a view to avoiding it.”⁸¹ Also, HRDD is a “concept [that] describes the steps a company must take to become aware of, prevent and address adverse human rights impacts.”⁸² These formulations appear preferable for the purpose of Principle 17 as the first general explanation of HRDD. Such formulations would not problematize the aim dimension settled in Principle 11—but actually reinforce it by mentioning “with a view to avoiding it”—and would delay description of the more concrete action expected under HRDD in each of the cause-contribute-linkages contexts until Principle 19. Such a delay would be welcome as it would leave the task of Principle 17 unique: outline the four main steps and provide general “parameters” of HRDD to grasp correctly its distinct nature among other risk management approaches. It appears that the definition of HRDD in Principle 17 amalgamates the peculiar formulation in Principle 13b—featuring prevention and mitigation—with other elements of DD—“identify”

79 *Id.* Principle 17 Commentary, at 16.

80 *Id.* (“Potential impacts should be addressed through prevention or mitigation, while actual impacts—those that have already occurred—should be a subject for remediation (Principle 22.”).

81 *Protect, Respect and Remedy*, *supra* note 26, ¶ 25.

82 *Id.* ¶ 56.

which is the first step in any DD process, and “account for how they address,” which is a CSR-specific requirement grounded in transparency and accountability that is uncommon to transactional DD.⁸³ The result is a step back from the 2008 Report formulation; as with Principle 13, “mitigation” carries the danger of reopening the goal of RtR dimension (elimination), especially as Principle 17 is the first in-depth explanation of HRDD in the GPs.

Regarding the second aspect, Principle 17 Commentary links mitigation to potential impacts only: “Potential impacts should be addressed through prevention or mitigation, while actual impacts—those that have already occurred—should be a subject for remediation (Principle 22).”⁸⁴ This is a simplistic statement that endangers the internal consistency of the GPs regarding both the ex-ante (prevention) and ex-post (remediation) stages of HRDD. The Commentary on potential and actual impacts fails to impress the full relevance of mitigation regarding both potential and actual impacts. Linking mitigation with potential impacts only runs contrary to the text of GP 19 Commentary (“leverage to prevent or mitigate the adverse impact”)⁸⁵ as well as the OHCHR Guide’s definition of mitigation (of adverse impacts in general, not only potential impacts).⁸⁶

Regarding the ex ante stage, the reference to “mitigation” of risks (potential impacts) adds little if anything to what “prevention” already says; the danger is one of redundancy. It is the OHCHR complex explanation of mitigation that delivers insight and clarifies what mitigatory measures are legitimate and which might not be.⁸⁷ The Commentary however achieves not more than overlap with “prevention” and carries the potential to water down the elimination aim of RtR. Regarding the ex post stage, the Commentary is inadequate again. By stating that actual impacts “should be a subject for remediation,”⁸⁸ the Commentary seemingly denies the applicability of mitigation with regard to remedies. The danger is one of partialism. Concerned not to water down the core idea that infringements of rights should be remediated fully, the Commentary

83 *Guiding Principles*, *supra* note 1, Principle 17, at 16. For more on transactional due diligence, see *supra* note 5.

84 *Guiding Principles*, *supra* note 1, Principle 17 Commentary, at 16.

85 *Id.* Principle 19 Commentary, at 19.

86 *Interpretive Guide*, *supra* note 14, at 6–7.

87 See discussion *infra* Section IV. D.

88 *Guiding Principles*, *supra* note 1, Principle 17 Commentary, at 16.

loses sight of the specific context of third party operations to which RtR does apply. The company has to exercise leverage over third parties to respect human rights, even if the company has no responsibility to offer itself remediation in the “linkages” scenario (Principle 13b). Whether that leverage covers only re-occurrence or also remediation to rightholders is not fully clear from Principle 19, which explains the notion of leverage. So, regarding the ex post stage, Principle 17 represents an unsuccessful attempt at simplification of HRDD as it fails to acknowledge the multiple dimensions of mitigation in the BHR setting.

In sum, Principle 17 offers a needless complication in the formulation of HRDD (by comparison with previous SRSG reports) by referring to “mitigation” side by side with “prevention” while the Commentary offers a partial and therefore deficient framing of “mitigation” (at both ex ante (potential impacts) and ex post (actual impacts) stages). The drafting of Principle 17 again falls prey to the attraction of premature and de-contextualized specification. The result is an undesirable potential reopening of the aim dimension settled in Principle 11—because of the potency of HRDD as an operational concept compared to the abstract and normative RtR—and compounds the deficiencies raised by Principle 13 with its overly complex drafting. Caught between redundancy and incomplete specification, the formulations in Principle 17 and its Commentary are opaque and a weak link between Principles 11 and 19.

Principle 19

Principle 19 offers a description of “appropriate action” to respect human rights structured along the cause-contribute-linkages dimension introduced in Principle 13.⁸⁹ It offers much needed clarification by specifying what is expected in each setting. Thus, for the cause scenario, the company should “cease or prevent the impact.”⁹⁰ For the contribute scenario, the company should “cease or prevent its contribution and use its leverage to mitigate any remaining impact.”⁹¹ For the linkage scenario, the company should use its “leverage to prevent or mitigate the adverse impact,” and if this leverage is lacking or failing, the company “should consider ending the relationship.”⁹²

⁸⁹ *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 18.

⁹⁰ *Id.*

⁹¹ *Id.*

⁹² *Id.*

From the language of Principle 19 it becomes clear that “mitigation” of impacts is a notion relevant only when third party conduct is present: when the company contributes or is linked to a wrongdoer. As the company has no control over a third party’s operations, mitigation (reducing the impact) applies. This reduction of impacts is presented as a legitimate goal for appropriate actions and appears to be in tension with the elimination aim of general RtR in Principle 11. Principle 19 Commentary is important because it indicates that elimination of impacts can be achieved in two different ways and that mitigation (reduction goal) is tolerated conditionally and temporary.⁹³

Thus, for the contribute scenario, the company is expected to eliminate its own contribution—by preventing it ex ante, by ceasing it while it is on-going, and by remediating it ex post—as well as to not remain passive regarding what its partner is doing. Regarding the latter, the company should aim to mitigate the impact produced by the third party; leverage should be used to at least reduce that impact. Is such residual impact tolerated indefinitely in contradiction with Principle 11, which speaks of elimination? The answer is negative: if the relationship with the third party continues, the prescriptions of the linkage scenario begin to apply due to the ongoing association of the company with the wrongdoer.

For the linkage scenario, the company, by definition, has no causal conduct of its own to preempt, to terminate (cease), or to remediate. What the enterprise has is a linkage that needs to be managed in two main steps: using leverage to address impacts, and, should that fail, terminating the relationship. The Commentary to Principle 19 clarifies some important points. The order is leverage first, termination second in order to preempt quick disengagement that would be more harmful to victims.⁹⁴ The decision to terminate should itself be scrutinized for adverse impacts on rightholders, again to preempt unintended consequences for rightholders becoming worse off. Exercising leverage can take multiple forms, ranging from applying direct pressure on the perpetrator to involving other public and private stakeholders to address impacts.⁹⁵ Having little leverage should lead first to efforts to increase that leverage before terminating the relationship. An exception for “crucial” relationships with strategic suppliers is allowed, but under the

93 *Id.* at 18-19.

94 *Id.*

95 *Interpretive Guide*, *supra* note 14, at 49.

condition that mitigatory efforts still be continued and with warning that the company assumes the risk as such efforts offer no cover against potential legal, financial, or reputational consequences.⁹⁶

Mitigation and leverage are tightly linked in Principle 19 which calls for use of “leverage to prevent or mitigate the adverse impact.”⁹⁷ The disjunctive “or” appears to allow the company to choose between the aim of a full elimination or mere reduction; this is problematic for both the internal consistency of the GPs (compare with Principle 11’s aim of elimination) and the specificity of the human rights context. Thus, the dangers would come from two directions: premature disengagement, which would be inconsistent with the interest of rightholders, or indefinite tolerance to residual harm, which would be inconsistent with the normative thrust of the human rights system. Through Principle 19, the GPs define a path towards elimination that impresses two points. First, residual impact is tolerated conditionally and temporarily. Mitigation (reduction as the aim) is conditional because the company has to continuously use its leverage to minimize impacts, and it is temporary because if continuous reduction stalls, the company should terminate the relationship. Second, the goal of elimination of impacts resulting from associated third party operations can be achieved either through minimization of impacts with a view to their full elimination or through elimination of the third parties from the value chain.

The GPs emphasize the two options of elimination and the order among them to impress that HRDD is a rightholder-centred risk management approach.⁹⁸ Its aim is not immediate elimination of impacts from a company’s value chains per se—which would be a traditional risk management approach—but increasing the protection for rightholders. The GPs sought to define RtR and HRDD in a way that is consistent with the aim of elimination of impacts and still can summon the leverage of lead firms throughout value chains.⁹⁹ In this way, GPs remain aligned with human rights normativity by not tolerating residual harm indefinitely and unconditionally, as well as seeking to activate new sources of power to drive change throughout value chains.

It appears that Principle 19 is well aligned with Principle 11 and consistent with the aim of elimination of impacts despite the

96 *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 18.

97 *Id.*

98 *Id.*

99 *See id.* at 18-19.

appearance of *mitigation* as reduction of impacts. The Commentary makes it clear that the relevance of mitigation is strictly in relation to third party conduct.¹⁰⁰ This is the sole organizational context where mitigation—the aim of reduction—is legitimate and where residual impact is tolerated. Mitigation introduces a temporal and conditional exception from the aim of elimination of human rights abuses. This is due to a value chain context where companies can achieve elimination of impacts by terminating the relationship with third parties instead of achieving elimination by securing rights through protective measures capable of preventing and redressing impacts on rightholders. The GPs try to secure leverage by delegitimizing cut-and-run responses to risk management (i.e. abruptly cutting linkages with wrongdoers) and to maximize a company’s leverage—in conjunction with actions of other actors—to actually deliver on the elimination goal of RtR. The elaborations in Principle 19 are essential for grasping how HRDD is distinct from other risk management approaches. It is in Principle 19 that the added value of *mitigation*—together with the notion of *leverage*—and its applicability to a distinct organizational context—third party harmful operations—becomes clearer. Mitigation hints at the governance dimension (polycentric governance to global value chains) and organizational dimension of the GPs (illegitimate risk-management approaches to third party non-compliance).

B. A Multilayered Concept of RtR

The RtR concept applies in a multitude of operational and organizational contexts, and the GPs offer a careful progression of the notion of corporate responsibility from the abstract “avoid infringing” to the specific “appropriate action.”¹⁰¹ The GPs gradually specify the RtR but each new layer of explanation and specificity should not problematize the aim of RtR (elimination of impacts) by casually using the notion of mitigation and unnecessarily complicated drafting. Based on this analysis, the value added by each principle can be pinpointed. Level 1 (Principle 11) is fundamentally about the aim of RtR: elimination of impacts from an enterprise’s operations. Level 2 (Principle 13) is fundamentally about the scope of RtR—a responsibility to act when a business causes, contributes, or is linked to adverse impact. Level 3 of specification (Principle

100 See *id.*

101 *Id.* Principles 11 and 19, at 13, 18.

19) indicates how to act in these three organizational contexts explained along the temporal dimension of before-, during-, and post-harm. Level 4 of specification is captured in Principle 15, which synthesizes and streamlines the key components of the RtR: policy, due diligence, and remediation. Level 5 of specification comes in Principle 17 that further disaggregates the key steps HRDD requires: impact assessment, action, tracking, and communication. Level 6 of specification contains parameters and insights that are indispensable for understanding and implementing the RtR, and are to be found in Principles 14, 17, 23, 24 and throughout Pillar 2. At this level, Principle 17 Commentary arguably misses the chance to impress strongly enough a key parameter: HRDD needs to tackle root causes of abuse in own operations and value chains. Only measures aimed at root causes are capable to deliver on the aim of RtR: elimination as non-occurrence of impact and abuse.

In sum, Pillar 2 offered important layers of specification as follows:

Level 1 (Principle 11)—*aim* of RtR: elimination (human rights dimension)

Level 2 (Principle 13)—*scope* of RtR: cause/contribute/linkage (causality dimension)

Level 3 (Principle 19)—*action* in the organizational context: own conduct versus third party conduct (organizational dimension)

Level 4 (Principle 15)—*action* to discharge RtR: three components of RtR (content of RtR dimension)

Level 5 (Principle 17)—*action* to discharge HRDD: four steps of HRDD (content of HRDD dimension)

Level 6 (Principles 11-24)—*parameters* (nature of the risk-management dimension)

The importance of aligning CSR standards is well accepted, as this can stabilize expectations and increase leverage brought by different stakeholders and governance schemes.¹⁰² During his mandate, the SRSG did emphasize external alignment, that is, between the GPs and other CSR instruments.¹⁰³ This insistence on

¹⁰² European Commission, *A Renewed EU strategy 2011-14 for Corporate Social Responsibility*, at 3–7, 13, COM (2011) 681 final (Oct. 25, 2011), [http://www.europarl.europa.eu/meetdocs/2009_2014/documents/com/com_com\(2011\)0681/com_com\(2011\)0681_en.pdf](http://www.europarl.europa.eu/meetdocs/2009_2014/documents/com/com_com(2011)0681/com_com(2011)0681_en.pdf).

¹⁰³ RUGGIE, *supra* note 47, at 159–166.

alignment opened the door for the GPs to be incorporated in other binding and non-binding mechanisms, which now regularly refer to HRDD. That increased the leverage of the GPs, and of the SRSG efforts, way beyond the UN halls. This article however mulls over internal alignment, meaning the extent to which the general and more specific levels of RtR are synchronized and flow from general to specific. Were misalignments to take place, this would create possibilities for misunderstanding, less clarity, more complexity, and eventually diminish the persuasive potential of the GPs. That would reduce a key source of leverage—persuasion—on which a non-binding instrument like the GPs is bound to count on.

As the previous subsection substantiated, the flow in Pillar 2 is shortcut at times by complex drafting. Principle 13 merges a number of dimensions in such a way that terminology can simply not keep up with; the asymmetric drafting between paragraphs 13a and 13b ends up presenting the responsibility in 13b as the odd one out, and potentially out of tune with RtR's aim of elimination. Also, there is premature and needless specification when *prevention* and *mitigation* are placed side by side in the definition of HRDD in Principles 15 and 17, which risks watering down the aim of RtR. There is also amalgamation of organizational contexts and partial explanations regarding adverse impacts in Principle 17 Commentary that introduce the charged concept of mitigation accompanied by clarifications that compound redundancy and confusion regarding contexts and impacts. Finally, Principle 19 Commentary offers important insights but still omits a full explanation of the relation between *leverage* and remediation aspects.¹⁰⁴ The next section seeks to alleviate such drafting deficiencies by clarifying the precise applicability of mitigation and its added value in the economy of the GPs.

IV. Mitigation Clarified

The previous section indicated that drafting choices and the introduction of the notion of mitigation decrease clarity in Pillar 2. More specifically, mitigation understood as reduction of impacts threatens to reopen the aim of RtR (elimination of impacts) and creates new complications for both prevention and remediation. These difficulties can be explained by highlighting the multidimensional character of mitigation in the GPs: the two meanings of mitigation, the two organizational contexts where

104 *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 18-19.

mitigation applies, and the two types of adverse impacts it covers. By identifying and untangling these dimensions, the GPs can be interpreted to justify the presence and explain the value added of mitigation in the GPs (two functions mitigation fulfills in defining the RtR), and the operational implications (legitimate and illegitimate risk-management measures). The implication is that mitigation is a charged concept with wide ramifications that should rather not be used casually to explain corporate responsibilities under the GPs.

A. Meanings of Mitigation

Mitigation of impacts could refer to mitigation as *aim* (reduction as opposed to elimination) and mitigation as *action* (mitigatory measures to minimize impact as opposed to inaction). The problem with mitigation as aim is that in the human rights context, it is incompatible with the RtR's aim of elimination of impact (human rights harm) from business operations. That raises the question of if, and under what conditions, some contexts of the RtR could accommodate this aim of reduction. The problem with mitigatory measures is that it creates redundancy with simpler notions of preventive and corrective actions, with uncertain value added. To the extent that the distinction between measures and aim fades, the risk is that the aim of elimination specific to the RtR context—as different from other risk-management contexts—will be confused and watered down as businesses are asked to mitigate human rights abuses.

B. Contexts of Mitigation

It should be recalled that RtR has broad coverage through the cause-contribute-linkages dimensions established in Principle 13.¹⁰⁵ Principle 13 covers operations and decisions over which the company has control, as well as operations of third parties over which the company has stronger or weaker leverage, if at all. It is the context of third party operations that makes the situation more complex. Ensuring human rights in third party operations boils down to exercising leverage or terminating relationships (linkages scenario), as well as removing own harmful contributory decisions (contribution scenario). Leverage and termination are both fraught by the reality that elimination cannot be achieved satisfactorily; leverage might not be enough to deliver observance of human rights,

¹⁰⁵ See *supra* Section III.A., Principle 13.

and termination will eliminate impacts from the value chains but not from the life of rightholders that might be worse off following the lead firm cutting the relationship with the supplier.¹⁰⁶ That means that mitigation can become a legitimate aim of RtR in third party operations—in the contribution and linkages scenarios.¹⁰⁷ However, this aim is conditional and temporal.

Whether mitigation in third party operations is an aim or a measure is in the eye of the beholder. However, given the complexities of long value chains, the period of time during which residual impact is tolerated is bound to be long as change will unavoidably take time. Expecting companies to address root causes of abuse through HRDD, as proposed herein,¹⁰⁸ increases the complexity of the task even more, which entails that some residual impact will be present for the long term. Overall, complications appear because the GPs use the notion of mitigation in the context of human rights, in the transnational setting of global value chains, and in different organizational contexts involving third parties. The aim here is to prevent the risk of contamination coming from insufficient attention to context that would end up watering down the RtR and give businesses unwarranted flexibility and discretion.

C. Types of Adverse Impacts

Discussing mitigatory measures also requires distinguishing between potential and adverse impacts—between risk of harm and harm that has occurred. Although the GPs refer to the mitigation of adverse impacts, both potential and actual,¹⁰⁹ nowhere is mitigation defined. That definition is provided by the OHCHR Guide, which shows this definition is not a self-evident task in the human rights context. The Guide offers a twofold definition. On the one hand, “[t]he mitigation of adverse human rights *impact* refers to actions taken to reduce its extent, with any residual impact then requiring remediation.”¹¹⁰ On the other hand, “[t]he mitigation of human

106 CENTRE FOR RESEARCH ON MULTINATIONAL CORPS, *Conflict Due Diligence by European Companies*, SOMO 3 (November 2013), <https://www.somo.nl/nl/conflict-due-diligence-by-european-companies/> (“In the case of conflict minerals from the DRC, it is recognised that such due diligence should be undertaken in a manner that does not contribute to the de facto embargo of minerals from the region.”).

107 See *supra* Section III.A., Principle 13.

108 See *infra* Section V.

109 *Guiding Principles*, *supra* note 1, Principle 17, at 16.

110 *Interpretive Guide*, *supra* note 14, at 7. (emphasis added)

rights *risks* refers to actions taken to reduce the likelihood of a certain adverse impact occurring.”¹¹¹ It is this latter definition that points out that some risk-management measures are not legitimate in the human rights context. Such measures cover both the *ex ante* (prevention) and *ex post* (remediation).

D. Preventive Measures

The OHCHR Guide offers an improvement over the GPs that simply stated that “[p]otential impacts should be addressed through prevention or mitigation, while actual impacts – those that have already occurred—should be a subject for remediation.”¹¹² The GPs link mitigation to potential impacts only—a clarification smacking of redundancy with prevention and failing to draw an essential operational implication. The Guide, however, defined mitigation of potential impacts as a reduction in likelihood only.¹¹³ Indeed, in traditional risk assessment, a risk is reduced by diminishing the damage (extent) and/or its probability.¹¹⁴ What the Guide indicates is that a reduction of risks in extent without a change in the likelihood (same probability) that harm will happen is an illegitimate form of mitigation under the GPs. An unchanged likelihood of harm—even after HRDD is applied—is tantamount to a permanent acceptance of residual harm, which is not acceptable even at the *ex ante* stage. This is due to the human rights context and the Guide captures its normativity well. Thus, preventive measures that only reduce the extent of potential harm are illegitimate under the GPs; a reduction in likelihood is indispensable.

Mitigation thus requires elaboration at the *ex ante* stage, a clarification that is however absent from the GPs and only added in the Guide. Further reflection on mitigation is needed even more at the *ex post* stage. Indeed, the GPs linking mitigation with potential impacts only will require further interpretation as it runs contrary to the text of GP 19 (“leverage to prevent or mitigate the adverse impact”)¹¹⁵ and contrary to the Guide’s definition of mitigation (of adverse impacts in general, not only potential impacts).

E. Remedial Measures

111 *Id.* (emphasis added).

112 *Guiding Principles*, *supra* note 1, Principle 17 Commentary, at 16.

113 *Interpretive Guide*, *supra* note 14, at 7.

114 *Id.* at 83.

115 *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 19.

The GPs state that “[p]otential impacts should be addressed through prevention or mitigation, while actual impacts—those that have already occurred—should be a subject for remediation.”¹¹⁶ By stating that actual impacts “should be a subject for remediation,” the GPs seemingly deny the applicability of mitigation with regard to remedies. Not only this, but the GPs¹¹⁷ indicate that a company does not have to offer remediation for mere linkages with abusive third parties: RtR includes “[p]rocesses to enable the remediation of any adverse human rights impacts they cause or to which they contribute.”¹¹⁸ That means that the linkages scenario of Principle 13b loses the remediation component: the company does not have a responsibility to offer remedies to rightholders for the impacts caused solely by its partners without a contribution from the company. Does that mean that mitigatory measures regarding remediation are not expected? That measures to facilitate or offer remediation are optional rather than imperative, desirable but not expected?

One interpretation is that a company does not have a responsibility to offer remedies under Principle 13b situations. Principle 22 states:

Where adverse impacts have occurred that the business enterprise has not caused or contributed to, but which are directly linked to its operations, products or services by a business relationship, the responsibility to respect human rights does not require that the enterprise itself provide for remediation, though it may take a role in doing so.¹¹⁹

The GPs thus indicate it is optional and desirable for a company to engage in remediation efforts. It is legitimate too, which can be an important aspect to the extent the company takes a critical stance regarding local regulatory failures.

Another interpretation would begin by noticing that even if the company does not have a responsibility to offer remedies directly, this does not necessarily mean it should be unconcerned by whether rightholders get remediation from the third party or the state. What is then the content of appropriate action regarding remediation under

116 *Id.* Principle 17 Commentary, at 16.

117 *See id.* Principles 13b, 15c, 22, at 14, 15, 20.

118 *Id.* Principle 15c, at 15.

119 *Id.* Principle 22 Commentary, at 20–21.

the GPs, and is that action really optional in character? Principle 19 Commentary speaks about a responsibility to apply leverage to mitigate impacts.¹²⁰ Nowhere in the text are these impacts limited to potential impacts. Actually, regarding contribution, companies are asked to “mitigate any remaining impact.”¹²¹ And regarding linkages, companies are asked to “prevent or mitigate the adverse impact”;¹²² similarly Principle 13 indicates companies should “[s]eek to prevent or mitigate adverse human rights impacts.”¹²³ Under Principle 19, which asks companies to exercise leverage to mitigate adverse impacts, it would be curious if the company was diligently using its leverage *ex ante* to prevent and, during abuses, to stop them, but desist once the harm occurred and remain indifferent to whether the stakeholders it aimed to protect have access to remedy or not. Therefore, the imperative responsibility to exercise leverage to address impacts should expand to cover remediation by third parties.

If leverage over the third party exists, why should it be used to mitigate potential impacts only, and not actual impacts as well? This remains fully compatible with Principle 22. Such interpretation is also mindful that the distinction between linkages and contribution—despite the dramatically different prescription regarding remediation—is more a term of art and science, because stakeholders almost always disagree on the correct characterization of contribution or linkages in value chains. Lead firms are seen as (a) sometimes contributing directly through their purchasing decisions or (b) maybe encouraging by failing to take a stronger stance (possibly by adopting an outsourcing business model that systematically enhances competitive market pressures on suppliers throughout the value chains) and, ultimately, (c) always benefiting from cost reductions from exploitative value chains.¹²⁴ So the imperative nature of responsibility to exercise leverage towards remediation can be accepted, but what are its operational implications?

So far, the analysis points that taking mitigatory measures regarding remediation is imperative. Could there be mitigatory measures regarding remediation that are illegitimate? For instance,

120 *Id.* Principle 29 Commentary, at 18–19.

121 *Id.* Principle 19 Commentary, at 19.

122 *Id.*

123 *Id.* Principle 13, at 14.

124 See *Interpretive Guide*, *supra* note 14, at 5 (defines both “legal” and “non-legal complicity”).

could a company terminate relationships with its supplier for failure to remediate actual impacts? Would that be legitimate or not? Recall what Principle 19 depicts as “appropriate action” in the linkages scenario: first apply leverage over the third party, then terminate the relationship.¹²⁵ There is a sequencing of steps—an escalation in the response of the company. The fact that the supplier has not remedied abuses, the fact that rightholders did not obtain remediation for harms, is by itself not enough reason to terminate the relationship. Conceivably, as long as using leverage makes progress in continuously reducing the risk (making gains in preventing harm and curtailing on-going abusive practices), it would not be legitimate for the lead firm to justify termination based solely on the third party’s failure to remediate past harms. Clearly, failures of the supplier to remediate harm counts—as one factor among others, with all of them assessed in a local context—in a possible escalation from leverage to termination. However, the GPs provide the company a conditional and temporary exception from the goal of elimination in third party contexts and emphasize pursuing a sequence of steps against uncompliant third parties.¹²⁶ This risk-management approach is appropriate to the human rights context where the rightholders’ interests count against facile risk-management measures (cut and run). Therefore, the Commentary to Principle 19 read together with Principle 22 should be interpreted as a rightholder-centered approach to risk management that denies the legitimacy of measures that use remediation failures to prematurely escalate from using leverage (step 1) to terminating the relationship (step 2).

It follows that some unremedied harm, the residual impact inherent in mitigation, would be acceptable as an exception from the elimination aim of Principle 11. This is due to the choice between the two mitigatory measures—elimination (of impact from value chain through cutting links) and reduction (of impact through staying engaged and exercising leverage)—that can only be legitimately made with attention to the local context and following an informed and participatory process. Failure of third parties to remediate severe harms would change the discussion and tilt the balance towards termination. Also, total unresponsiveness by third parties to remediation concerns, particularly where feasible

125 *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 18–19.

126 *Id.*

measures are possible, would provide a stronger ground for the lead firm to legitimately terminate the relationship. Failure of suppliers to remediate harm can become a root cause of ongoing and future impacts that lead firms have to address through their HRDD process.

Having the lead firm tolerate residual harm (unremedied harm) in third party operations is not tantamount to partial remediation (reduced harm) as a legitimate way of doing business under the GPs. It merely applies to one organizational context—third party operations. Remediation in this context is a composite one: third parties (suppliers and host state) offer full remediation while the company offers partial remediation by exercising leverage over the third parties. Indeed, under the GPs, full remediation is within the responsibility of the wrongdoer (supplier) to *respect* (Pillar 2) together with the obligation of the state to *protect* (Pillar 1).¹²⁷ Companies linked to the wrongdoer have a role in *mitigating* these actual impacts by exercising their leverage on third parties to obtain remediation for rightholders. This is a responsibility not to remain passive, even when third parties impact rights and even at ex post stage.

A conclusion can be derived by bringing together the three angles proposed here: first, it is imperative to exercise leverage to secure remediation by others; second, unsuccessful leverage over remediation is usually not a sufficient reason to terminate the relationship; and third, the ability to justify the choice between the two mitigatory measures of leverage and termination depends on the context. The conclusion is that the GPs should not be interpreted as offering companies in the Principle 13b context blank permission to stay out of remediation and focus only on prevention.¹²⁸ While relieved of the burden of offering remediation themselves, companies are expected to use their leverage to mitigate harm, including mitigating the lack of remedies for rightholders. This point is made explicit

127 *Id.* Principles 1, 11, at 6, 13.

128 One can disagree here with the interpretation offered by Shift. “O&G companies do not have to remediate . . . Impacts they have neither caused nor contributed to: it is the responsibility of those who have contributed to the impacts to provide for or cooperate in their remediation. However, where the impacts are nevertheless linked to the O&G company’s operations, it has a responsibility to use its leverage to prevent or *mitigate the risk* of the impacts continuing or recurring.” European Commission [EC], *Oil and Gas Sector Guide on Implementing the UN Guiding Principles on Business and Human Rights*, at 70 (2013), https://www.ihrb.org/pdf/eu-sector-guidance/EC-Guides/O&G/EC-Guide_O&G.pdf [hereinafter *Oil and Gas Guide*] (emphasis added).

in the relevant ILO document on corporate responsibilities which has been revised to incorporate the GPs. It reads: “Multinational enterprises should use their leverage to encourage their business partners to provide effective means of enabling remediation for abuses of internationally recognized human rights.”¹²⁹

Mitigation here is a legitimate aim under RtR, not only merely a mitigatory measure. It is a genuine exception from the RtR’s elimination aim that is justified by the composite remediation in the third party context. By eliminating remediation from the linkages scenario, the GPs drafters seem to have overreacted in their attempts to, on the one hand, present to businesses a precise RtR that is not overly broad in the supply chain context, and, on the other hand, to not dilute remediation through partial solutions (full remediation being a human right). However, in the value chain context, remediation is a composite one and a human rights imperative is to summon the leverage of lead firms to clean value chains and increase protection of rights throughout the chains. That means this context of remediation (ex post stage) is where mitigation has the highest value added in the GPs: it depicts a genuine exception from the aim of elimination and corrects the overreaction of the GPs to the idea of partial remediation. It is an exception that applies only to the lead firm (not to suppliers and host states), only in the context of third party operations (not causation and contribution), and only to remediation (not prevention).

This analysis allows a comparison of mitigation at the ex post and ex ante stages. At the ex post stage of remediation, mitigation is a genuine exception from the elimination aim of RtR. And it refers to the very aim of reduction and not merely to mitigatory measures. Actually, the GPs have defined the RtR to not cover remediation in linkages situations.¹³⁰ Through as argued above, and explained in the section below, mitigation can actually correct the overreaction of the GPs drafters take on the RtR.¹³¹ In contrast, at the ex ante stage of prevention, mitigation is an apparent, not a genuine, exception, from the elimination aim given that it is a conditional and temporal exception mandating some steps a company should take before

129 Int’l Labour Org., *Tripartite Declaration of Principles concerning Multinational Enterprises and Social Policy*, at 65 (2017), http://www.ilo.org/wcmsp5/groups/public/---ed_emp/---emp_ent/---multi/documents/publication/wcms_094386.pdf.

130 *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 19.

131 *Infra* Section IV.F.

cutting links with third parties.¹³²

These prevention and remediation complexities in the multiple contexts of RtR bring the analysis to the functions that mitigation actually plays in the GPs. If mitigation cannot be shown to add any value in explaining the RtR, then this charged concept should be dismissed as redundant, decried as diminishing the simplicity and clarity of the GPs, and deemed counterproductive in the human rights and business context.

F. Functions of Mitigation

The appearance of mitigation throughout the GPs introduced a level of ambiguity. The ambiguity can be explained by the multidimensional character of mitigation applied in a human rights and transnational governance context. With further interpretation, it appears that this notion serves a number of functions of high importance for RtR.

First, mitigation has a distinct place in the GPs by carving out two exceptions—a seeming exception and a genuine exception—from the goal of elimination. Furthermore, both exceptions apply only in the context of third party operations. One the one hand, the seeming exception has to do with the sequence of steps (use leverage, increase it, consider impacts of terminating relations) a company should take before cutting links with third parties. In the GPs, priority goes to reducing impacts and ultimately eliminating them through improved third party practices, not through cutting links.¹³³ In this way, referring to *mitigation* prevents the goal of elimination from taking over through the wrong means (elimination by cutting links) and legitimizes the use of leverage as the appropriate means. This exception from the elimination goal is also a temporal and conditional one.

On the other hand, the genuine exception from the elimination goal has to do with remediation. Remediation in third party operations is composite remediation and the GPs, with their polycentric governance approach, seek to mobilize leverage from multiple duty-holders—host state, supplier, and lead firm—to ensure human rights.¹³⁴ Principle 19 can be interpreted to justify an imperative responsibility for lead firms to use leverage to increase remediation, even if this in itself can only amount to partial or

132 *Infra* Section IV.F.

133 *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 18–19.

134 See Ruggie, *supra* note 50.

uncertain remediation.¹³⁵ This is a genuine and permanent exception from the elimination aim. In this way, referring to *mitigation* prevents the goal of elimination pursued by some corporate accountability advocates through ambitious means (strict liability of lead firms for failure to prevent or remediate) from taking over more modest, but potentially more realistic and transformative contributions by lead firms exercising leverage and contributing to remediation by third parties.

Second, the appearance of mitigation in HRDD seemed to have fulfilled another function. References to leverage and mitigation in Principle 19 appear to somehow counterbalance an overly narrow formulation of RtR and create an opening to expand the scope of RtR regarding remediation. Here, mitigation works not to imperil the elimination aim of RtR, but to expand the scope of RtR. Rather than a notion that confuses and potentially waters down the aim of RtR, mitigation, together with the notion of leverage, prevents a narrow interpretation of the GPs, which explicitly dropped remediation from a company's RtR in the paragraph 13b context.¹³⁶ The GPs references to mitigating adverse impacts can be used to impress that companies have an imperative rather than a discretionary role in remediation in third party operations, despite the text of Principle 22 Commentary. Thus, the rightholders' access to remedies is enhanced.

Third, introducing mitigation in the definition of RtR carries connotations of reasonableness and familiarity with established risk-management practices. Mitigation might communicate that RtR imposes a reasonable burden on companies. The operational rather than aspirational character of the entire GPs is strengthened. Mitigation points out that companies cannot truly guarantee (ensure) non-occurrence. Instead, mitigation indicates that reduction is a commendable short-term achievement for a company embarked on the long and hard road of achieving full elimination. Even in own operations, mitigation communicates that a responsible company involved in inherently hazardous activities should prepare for emergency situations with contingency planning to mitigate (reduce) the scale of a disaster, should it happen. Thus, the company would reduce unfolding impacts through diligent planning and action before the impacts escalate, rather than being unprepared and inactive. However, these correct connotations do not need to

135 See *supra* Section IV. E.

136 *Guiding Principles*, *supra* note 1, Principles 15c, 22, at 15, 20 (stating that remediation is expected only where a company cause or contributed to adverse impacts, and not for mere linkages to abusive third parties).

be communicated obliquely through the term “mitigation”—placed side by side with prevention—as the GPs adopt a tone and a range of explanations that achieve the task more directly and effectively. Thus, this function of mitigation, though appealing at first sight, adds little value to the GPs and would be well replaced by references to preventive and corrective measures to minimize harm, particularly in the text preceding Principle 19 where mitigation makes its genuine contribution. Explaining HRDD as being generally about mitigation of impacts is least helpful in the human rights context and obscures the rightholder-centered risk management approach that the GPs genuinely put forward.

In sum, this analysis pinpoints the applicability of mitigation (aim of reduction of impact) to one organizational context only: that of third party conduct, which comprises the “contribute” and “linkages” scenarios.¹³⁷ Only in this context can mitigation remain consistent with the elimination aim of RtR, as stated in Principle 11. The value mitigation has added to the GPs follows from the two specific functions discernable in Principle 19. First, the function to pinpoint—together with the notion of leverage—a key HRDD step, and a stop gap towards the aim of elimination.¹³⁸ Second, the function to expand—again, together with the notion of leverage—the scope of RtR by having the company play a role in securing rightholders’ access to remedies.¹³⁹ Finally, the analysis herein points out that *mitigation of impacts* is a charged and multidimensional notion that can confuse and water down the elimination aim of the RtR stated in Principle 11. The watering down of RtR from a language emphasizing mitigation is generated by drafting imperfections in the GPs and mainly by the weight of the HRDD notion relative to the abstract and normative RtR—less in the economy of the GPs themselves, but more in the way that GPs are read by stakeholders. Clarifying mitigation in the GPs ensures consistency between the RtR and HRDD, and affirms HRDD as a rightholder-centered risk-management approach geared towards elimination of impacts.

The next section draws attention to a general parameter of RtR and HRDD. Arguably a discussion about *root causes* of impacts and about using a purpose-built methodology to identify and address deep causes would have been well placed among the parameters in Principle 17. Furthermore, this parameter is consistent with the GPs

¹³⁷ See *id.* Principle 13, at 14.

¹³⁸ *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 18–19.

¹³⁹ *Id.*

and is observable in post-mandate documents involving Ruggie and his team.

V. “Root Causes” of Adverse Impacts

One danger hovering over an instrument that places so much weight on the notion of HRDD is that this risk management process is rather indeterminate and might ultimately not be effective due to the flexibility and discretion it affords companies. This is not only because mitigation used in HRDD created ambiguities that could be exploited to lower compliance, but also because the GPs lack a centralized monitoring or dispute resolution mechanism, are a general instrument devoid of specific guidance for specific rights and specific contexts, and are silent on the amount of effort expected from companies implementing HRDD. These limitations can be explained against the strengths of the GPs as a foundational treatment for the BHR field,¹⁴⁰ but arguably it makes it even more important for HRDD to be explicitly geared towards tackling root causes of adverse impacts rather than mere symptoms and contributing factors.

This section presents options for tightening HRDD to ensure meaningful rightholder-centred risk-management, highlights references to root cause analysis in the GPs-inspired documents from Ruggie and Shift, and presents implications for prevention and remediation that follow from this analysis of mitigation and root causes.

A. *Options for Tightening HRDD*

With no centralized oversight mechanism and no specificity of expected measures, there is an inbuilt looseness in the HRDD, which can result in a HRDD that is largely symbolic, generates limited improvements, and fails to address underlying issues.¹⁴¹ A

140 Radu Mares, *A Rejoinder to G. Skinner’s Rethinking Limited Liability of Parent Corporations for Foreign Subsidiaries’ Violations of International Human Rights Law*, 73 WASH. & LEE L. REV. ONLINE 117, 117–58 (2016), <http://scholarlycommons.law.wlu.edu/cgi/viewcontent.cgi?article=1045&context=wulr-online> (explaining the evolution of regulatory thinking in the BHR field through three ‘baselines’).

141 Judy Gearhart, *Global supply chains: time for a new deal?*, OPENDEMOCRACY (June 2, 2016), <https://www.opendemocracy.net/beyondslavery/ilc/judy-gearhart/global-supply-chains-time-for-new-deal> (considering that “[t]he problem with the UNGPs as they are currently conceptualised, however, is that they risk simply creating a new generation of voluntary programmes because they do not define minimum performance requirements or set regulatory

good starting point for analysis is the definition of HRDD, which in the GPs emphasizes the steps of the management process a business should employ.¹⁴² However, in a report preceding the GPs, Ruggie covered a different dimension of HRDD: “Due diligence is commonly defined as ‘diligence reasonably expected from, and ordinarily exercised by, a person who seeks to satisfy a legal requirement or to discharge an obligation.’”¹⁴³

Subsequently, the OHCHR more precisely linked HRDD to the “reasonable person” which is the notion commonly used in the law of negligence.¹⁴⁴ Actually, the Guide offered two definitions. A general due diligence definition indicated that:

Due diligence has been defined as “such a *measure of prudence*, activity, or assiduity, as is properly to be expected from, and ordinarily exercised by, a reasonable and prudent [person] under the particular circumstances; not measured by any absolute standard, but depending on the relative facts of the special case.”¹⁴⁵

Specifically for HRDD, the Guide wrote: “In the context of the Guiding Principles, human rights due diligence comprises an *ongoing management process* that a reasonable and prudent enterprise needs to undertake, in the light of its circumstances (including sector, operating context, size and similar factors) to meet its responsibility to respect human rights.”¹⁴⁶ The necessity to have two definitions could be explained by the different emphasis: the first definition indicates a “measure of prudence” while the second indicates “management process.” This is significant given that the GPs are a type of instrument unable to specify the amount of effort expected from companies; it can only depict the risk-management

standards.”).

¹⁴² *Guiding Principles*, *supra* note 1, Principle 17, at 16. See also *supra* text accompanying note 77.

¹⁴³ John Ruggie (Special Representative to the Secretary General), *Business and Human Rights: Towards Operationalizing the “Protect, Respect and Remedy” Framework*, ¶ 71, U.N. Doc. A/HRC/11/13 (Apr. 22, 2009).

¹⁴⁴ Mayo Moran, *The Reasonable Person: A Conceptual Biography In Comparative Perspective*, 14 LEWIS & CLARK L. REV. 1233, 1238 (2010).

¹⁴⁵ *Interpretive Guide*, *supra* note 14, at 6 (emphasis added)(citation omitted) (quoting *Due Diligence*, BLACK’S LAW DICTIONARY (6th ed. 1990)).

¹⁴⁶ *Id.* (emphasis added).

process that has some obligatory steps.¹⁴⁷ This is the contribution of the GPs, as well as their limitation. The looseness comes from the impossibility to specify the amount of effort beyond generalities like: the process has to be “effective,” companies should “take the necessary steps,” take “every reasonable step,” and that the diligence used to address impacts is “due.”¹⁴⁸ This is expected from the reasonable person whose compliance can only be determined by investigating the particular circumstances.

The HRDD notion in the GPs does not prejudge the amount of effort question in either paragraph 13a or 13b contexts. Paragraph 13a has a more stringent formulation for own conduct (cause and contribute scenarios);¹⁴⁹ however, that does not imply that the company is, or could be, asked by the GPs to adopt the highest levels of responsiveness (akin to a zero-tolerance approach) for all human rights risks. The fact that paragraph 13a is not worded with the qualifier “seek to” that accompanies paragraph 13b,¹⁵⁰ should not be seen as a mark of higher stringency of effort. Indeed, Principle 19 still speaks of “take the necessary steps to cease or prevent” own harmful conduct.¹⁵¹ Noteworthy though, the GPs do indicate a higher stringency for severe impacts and gross abuses.¹⁵²

The GPs were not meant to settle the amount of effort appropriate in specific circumstances. Instead, they were meant to arrive at a widely accepted understanding of corporate responsibilities (e.g. wide scope expanding throughout value chains, essential components of RtR) and move the BHR field beyond deadlock and polarization. Based on these achievements, further specification and remedial mechanisms would more easily arise. To its critics, the GPs were a missed opportunity precisely because new oversight mechanisms and more precise responsibilities did not emerge from

147 *Guiding Principles*, *supra* note 1, Principles 17–22, at 16–22 (detailing the steps of the risk management process).

148 *Id.* Principle 19 Commentary, at 18–19.

149 See *supra* text accompanying note 67.

150 See *supra* text accompanying note 67.

151 *Guiding Principles*, *supra* note 1, Principle 19 Commentary, at 18. See also Occupational Safety and Health Convention (No. 155), art. 16, June 22, 1981, 1331 U.N.T.S. 279 (asking employers “to ensure that, so far as is reasonably practicable, the workplaces, machinery, equipment and processes under their control are safe and without risk to health.”) (emphasis added).

152 *Guiding Principles*, *supra* note 1, Principle 14, 23, at 14, 18 (stating that businesses should prioritize addressing severe impacts and treat gross abuses as a matter of legal compliance).

the SRSG mandate.¹⁵³ To its supporters, the strength of the GPs came from proposing this new polycentric governance approach to dealing with regulatory gaps in the globalized economy, further augmented by HRDD parameters and insights that created a new rightholder-centred risk management approach.¹⁵⁴ These parameters are important for reducing the indeterminacy of the HRDD notion and for facilitating stakeholder dialogues around strategic and specific measures companies should undertake.

One such parameter is an emphasis on root causes. HRDD would be an effort oriented towards root causes rather than symptoms.¹⁵⁵ Thus while the amount of effort still remains open, the correct orientation of that effort towards deeper causes is added. An emphasis on root causes, which could well be one general parameter among others mentioned in Principle 17 and on Level 6 of the RtR, is one option to further tighten HRDD. Other options are review mechanisms, specification of concrete measures in specific contexts, and theories of strict liability. These three other options are essential for the BHR field and attracted attention and activity before and after the SRSG mandate;¹⁵⁶ however, none of these options seemed feasible for Ruggie to act on at the time and instead he chose to add value with the GPs as a short, foundational, non-legally binding instrument in a polarized BHR field.

Review Mechanism

An authoritative reviewer is indispensable to determine whether a sufficient amount of diligence is exercised. Such determination can only be made with a view to specific circumstances. Indeed, an authoritative reviewer (e.g. court) assesses whether the company displayed the diligence and prudence needed in the circumstances as measured against a legal standard of “reasonable

153 See Chris Albin-Lackey, *Without Rules: A Failed Approach to Corporate Accountability*, HUMAN RIGHTS WATCH (2013), https://www.hrw.org/sites/default/files/related_material/business.pdf.

154 Mares, *supra* note 48.

155 For example, the Fair Labor Association developed its “Sustainable Compliance methodology” which “is designed to move the field of social compliance beyond policing and band-aid fixes [and toward] uncovering root causes of problems and providing systemic, sustainable solutions so that problems are fixed in a lasting way.” *Our Methodology*, FAIRLABOR.ORG, <http://www.fairlabor.org/our-methodology> (last visited Sept. 9, 2017).

156 See Mares, *supra* note 140, at 118–20.

person.”¹⁵⁷ However, Ruggie approached his mandate persuaded not to repeat the failure of the UN Norms,¹⁵⁸ which envisaged a monitoring mechanism as well as a national remedial mechanism, in addition to proposing an expansive corporate responsibility.¹⁵⁹ The GPs, therefore, were not only more cautious about the scope of corporate responsibility, but Ruggie also did not risk sinking the entire instrument by making it dependent on establishing a monitoring and enforcement mechanism for their implementation. Instead, the GPs affirm the necessity of a remediation and grievance mechanism for the polycentric governance system to work and map the types of remedial mechanisms (containing judicial and non-judicial mechanism).¹⁶⁰

Ex Ante Specification (Action Plans)

Having the GPs outline specific measures to specific human rights in specific industries and geographies would have possibly been beyond the capacity of a SRSG mandate and impossible to concentrate in a HRC report. It would have also been premature if the general conceptual foundation of the field was lacking and stakeholders were in disagreement over fundamentals in a contested field like BHR. Still, the SRSG mandate conducted targeted research and published more specific guidance on a variety of issues such as impact assessments,¹⁶¹ non-judicial grievance mechanisms,¹⁶² state-investor contracts,¹⁶³ and other topics.¹⁶⁴ The quest for specifying

157 Moran, *supra* note 144.

158 John Ruggie, *Business and Human Rights: The Evolving International Agenda*, 101 AM. J. OF INT'L L. 819, 822 (2007) (explaining the failure of the Norms).

159 Sub-Commission on the Promotion and Protection of Human Rights, Res. 2003/16, UN Doc. E/CN.4/Sub.2/2003/12/Rev.2, arts. 16–18 (Aug. 13, 2003).

160 *Guiding Principles*, *supra* note 1, Principles 25–31, at 22–26.

161 See John Ruggie (Special Representative to the Secretary General), *Human Rights Impact Assessments – Resolving Key Methodological Questions*, U.N. Doc. A/HRC/4/74 (Feb. 5, 2007).

162 John Ruggie (Special Representative to the Secretary General), *Piloting Principles for Effective Company/Stakeholder Grievance Mechanisms: A report of Lessons Learned*, A/HRC/17/31/Add.1 (May 24, 2011).

163 John Ruggie (Special Representative to the Secretary General) *Principles for Responsible Contracts: Integrating the Management of Human Rights Risks into State-Investor Contract Negotiations: Guidance for Negotiators*, U.N. Doc. A/HRC/17/31/Add.3 (May 25, 2011).

164 For a complete list of reports issued by the SRSG mandate, see *U.N. Secretary-General's Special Representative on Business & Human Rights*, BUSINESS & HUMAN RIGHTS RESOURCE CTR., <https://business-humanrights.org/en/>

corporate responsibilities has defined the CSR movement, before and since the GPs; it increasingly delivers indicators, tools, and guidance in specific contexts and for specific HRDD steps.¹⁶⁵ The ISO 26000 is another guideline that goes more in-depth to specify expected action on a number of human rights issues.¹⁶⁶

Action plans are a method increasingly employed in business and human rights to specify HRDD actions.¹⁶⁷ This method is appealing as it can work in the absence of legal requirements but can also be supported or even mandated through law.¹⁶⁸ Such action plans happen at multiple levels. First, even in the absence of a reviewer or detailed legal requirements, companies adopt “corrective action plans” (CAPs) detailing the targets, measures, timelines, and resources, as needed, to address adverse impacts in local circumstances.¹⁶⁹ This applies to a company’s own operations as well as third parties where lead firms agree with suppliers on CAPs.¹⁷⁰ Such CAPs are a modality to settle the specific amount of

un-secretary-generals-special-representative-on-business-human-rights (last visited Sept. 8, 2017).

165 See THE DANISH INSTITUTE OF HUMAN RIGHTS, *supra* note 6; Global Compact Network Germany, *supra* note 6; Shift & Mazers LLP, *supra* note 7; European Commission, *supra* note 8; GWENDOLYN REMMERT ET AL., *supra* note 9.

166 INT’L ORG. FOR STANDARDIZATION, ISO 26000: GUIDANCE ON SOCIAL RESPONSIBILITY 19-69 (2010) (providing “[g]uidance on social responsibility core subjects”).

167 See, e.g., Claire Methven O’Brien et al., DANISH INST. FOR HUMAN RIGHTS & INT’L CORP. ACCOUNTABILITY ROUNDTABLE, *National Action Plans On Business And Human Rights: A Toolkit for the Development, Implementation, and Review of State Commitments to Business and Human Rights Frameworks* (June 2014), <https://static1.squarespace.com/static/583f3fca725e25fc45aa446/t/5865d59fe6f2e17f4f0cb629/1483068841826/DIHR-ICAR-National-Action-Plans-NAPs-Report3.pdf>

168 For example, France has passed legislation imposing a duty of vigilance on French companies, which requires them to establish and implement a ‘vigilance plan’. Sarah A. Altschuller & Corentin Chevallier, *French National Assembly Adopts Corporate Duty of Vigilance Law*, FOLEY HOAG LLP: CORPORATE RESPONSIBILITY AND THE LAW (Feb. 28, 2017), <http://www.assemblee-nationale.fr/14/pdf/ta/ta0924.pdf>. See also ASSEMBLÉE NATIONALE, *supra* note 10.

169 See *Remediation*, THE ACCORD ON FIRE AND BUILDING SAFETY IN BANGLADESH, <http://bangladeshaccord.org/remediation> (last visited Sept. 8, 2017) (providing an explanation on the inspection and remediation process under the The Accord on Fire and Building Safety in Bangladesh).

170 See, e.g., FAIRLABOR.ORG, *supra* note 155.

effort expected in specific circumstances of individual companies.¹⁷¹ Second, at higher levels, states adopt National Action Plans (NAPs) for BHR that are expected to specify further regulatory and business measures fitted for that jurisdiction. In August 2017, there were 13 NAPs that states have adopted even though this is not a legal obligation under an international treaty.¹⁷² Third, a recent trend and outgrowth of such NAPs are industry-level compacts where public and private stakeholders in a specific industry agree on plans and concrete measures to address typical impacts in that industry.¹⁷³ In sum, such action plans bring specificity to HRDD at company, industry, or jurisdiction levels. Finally, should there be legal obligations, prosecutors and judges may compel companies to set up compliance systems and plans with a view to a company's specific circumstances and in order to bring itself into compliance with law.¹⁷⁴ The French Parliament has recently passed legislation

-
- 171 See, e.g., MICROSOFT, *Citizenship Report*, 84 (2013), http://admin.csrwire.com/system/report_pdffs/1323/original/Microsoft_Citizenship_2013_v2.pdf (explaining the company "works closely with suppliers on corrective action plans to resolve issues. We require suppliers to identify the root cause, the corrective course of action, and future preventive actions for all the issues found. Suppliers must correct issues within specific deadlines based on the severity of the non-conformance found to avoid restrictions on new Microsoft business or the possibility that we'll terminate our business with them.").
- 172 The Office of the United Nations High Commissioner for Human Rights (OHCHR) maintains a repository of NAPs. *State National Action Plans*, UNITED NATIONS HUMAN RIGHTS: OFFICE OF THE HIGH COMMISSIONER, <http://www.ohchr.org/EN/Issues/Business/Pages/NationalActionPlans.aspx> (last visited Sept. 8, 2017).
- 173 The Netherlands government aims to conclude International Corporate Social Responsibility (ICSR) covenants with companies in 13 Dutch sectors, identified as high risk: agriculture, chemicals, construction, electronics, energy, finance, food, metals, oil and gas, retail, textiles/apparel, wholesale, and wood and paper products. *Progress on Voluntary Agreements and Legislation*, MVO PLATFORM, (Sept. 12, 2016), <http://mvopplatform.nl/news-en/developments-in-the-netherlands>. The Dutch government adopted a root-cause approach with its NAP when it wrote: "the challenge in the next few years will be timely identification of risks in Dutch companies' supply chains. The government wants to work on structural solutions within international chains, not incident management." Ministry of Foreign Affairs, *National Action Plan on Business and Human Rights*, 14 (2014), <https://business-humanrights.org/sites/default/files/documents/netherlands-national-action-plan.pdf>.
- 174 For instance, under anti-trafficking US Federal Acquisition Regulation (2015), federal contractors are required to develop compliance plans with minimum components such as an "awareness program," "a reporting process for employees to use," "a recruitment and wage plan," "a housing plan," and "procedures . . . to monitor, detect, and terminate agents; and

to makes HRDD mandatory; this would require large companies to adopt and implement systems in line with the GPs.¹⁷⁵

Theories of Strict Liability

A way to bypass the entire *amount of effort* difficulty is to make companies liable for harm irrespective of fault. Determining whether the company was prudent in preventing harm and exercised effort would become unnecessary.¹⁷⁶ Such theories of strict liability—as opposed to fault—or negligence-based liability—would be applicable for own operations as well as third party operations, covering all scenarios in Principle 13.¹⁷⁷ Conceivably, even lead firms that outsourced production would be deemed liable for suppliers' harmful operations. However, new legislation would be needed to create such liability, which could furthermore still allow for a defense for a company that demonstrated it undertook rigorous HRDD.¹⁷⁸ Allowing a HRDD defense, of a narrow or broad scope, would in effect position liability on a continuum between the no-fault and fault-based liability.

This new legal regime would result in further specification of what HRDD entails because companies would be pressed to urgently specify what compliance with the law entails and case-law would progressively specify what concrete HRDD is acceptable under the law. This solution would eschew difficulties regarding offering ex ante specification of HRDD. However, this liability theory would be of exceptional application as it overturns the rule in civil and criminal liability legal regimes: fault-based liability (as opposed to strict liability) and limited liability based on separate corporate

subcontractors.” Sarah A. Altschuller, *Federal Acquisition Regulation: Amendments to Strengthen Prohibitions Against Trafficking*, FOLEY HOAG LLP: CORPORATE SOCIAL RESPONSIBILITY ALERT (Feb. 2, 2015), <http://www.foleyhoag.com/publications/alerts-and-updates/2015/february/federal-acquisition-regulation-amendments-to-strengthen-prohibitions-against-trafficking>.

175 ASSEMBLÉE NATIONALE *supra* note 10. FAIR LABOR ASSOCIATION *infra* note 211.

176 See 3 LEAH HECTOR, CORPORATE COMPLICITY AND LEGAL ACCOUNTABILITY: CIVIL REMEDIES — REPORT OF THE INTERNATIONAL COMMISSION OF JURISTS EXPERT LEGAL PANEL ON CORPORATE COMPLICITY IN INTERNATIONAL CRIMES, 10 (2008).

177 See *supra* Section III.A.

178 See, eg., Proposition De Loi Relative Au Devoir De Vigilance Des Sociétés Mères Et Des Entreprises Donneuses D’ordre, Texte Adopté n°708, 23 mars 2016, www.assemblee-nationale.fr/14/pdf/ta/ta0708.pdf.

personhood (as opposed to joint and several liability).¹⁷⁹

The GPs are silent on theories of liability as Ruggie indicated his mandate was never aimed to arrive at a grand theory of corporate responsibility.¹⁸⁰ Should he have built the GPs on theories of strict liability, the opposition to the entire project would have been guaranteed. In fact, the UN Norms proposed a broad corporate responsibility formulated in a manner assuming a stricter form of liability and indicating the primacy of human rights law over other “lesser” bodies of law.¹⁸¹ This clearly contributed to the defeat of the Norms in the UN. The GPs being agnostic on theories of liability leaves the decision to regulatory processes at the national level, where regulators should design the appropriate liability regime. Notably, the GPs do not exclude any form of liability. Thus, while the GPs emphasize HRDD, there is a warning for businesses that undertaking HRDD is not automatically a defense:

Conducting appropriate human rights due diligence should help business enterprises address the risk of legal claims against them by showing that they took every reasonable step to avoid involvement with an alleged human rights abuse. However, business enterprises conducting such due diligence should not assume that, by itself, this will automatically and fully absolve them from liability for causing or contributing to human rights abuses.¹⁸²

179 HECTOR, *supra* note 176, at 10, 45–49.

180 Ruggie indicated he never intended to put forward “a robust moral theory or a full scheme for the attribution of legal liability to underpin the Framework.” RUGGIE, *supra* note 47 at 107. He further wrote, “I did not set out to establish a global enterprise legal liability model. That would have been a purely theoretical exercise”. *Id.* at 189.

181 Larry Catá Backer, *Multinational Corporations, Transnational Law: The United Nations’ Norms on the Responsibilities of Transnational Corporations as a Harbinger of Corporate Social Responsibility in International Law*, 37 COLUM. HUMAN RIGHTS L. REV. 336, 357, 363-64 (2005) (writing that “[s]ince the definition of TNC does not recognize the distinct legal personalities of the corporations that together constitute the TNC, the Norms essentially pierce the corporate veil . . . The Norms produce a standard incompatible with the domestic corporate laws of a majority of states. Yet the Norms make little effort either to recognize or resolve this conflict . . . The problem, of course, is that, as a matter of the domestic law of most states, the autonomous legal personality of a corporation matters. Most states have developed very strong public policies in favor of legal autonomy.”) (references omitted).

182 *Guiding Principles*, *supra* note 1, Principle 17 Commentary, at 17.

Root Cause Orientation

The twofold definition of the Guide indicates a concern for HRDD as a “process.”¹⁸³ The GPs are not about *amount* of effort but about the *type* of effort wherein certain steps are expected. A further emphasis, as proposed herein, on the *direction* of efforts towards deeper underlying causes of harm would impress the idea that superficial measures to mitigate impacts are inconsistent with HRDD and the GPs. Orienting HRDD towards root causes could be incorporated as one of the parameters of HRDD highlighted in Principle 17. It would be in tune with the general but valuable observations that are appropriate in a foundational and generic instrument like the GPs. This is consistent with lessons from supply chain management where efforts of brands to address suppliers’ harmful operations have not achieved expected results because root causes—such as own purchasing decisions and insufficient worker empowerment—have not been addressed properly or at all.¹⁸⁴ Such root causes put the spotlight on both the lead firm’s blameworthy conduct (paragraph 13a, “contribute”)¹⁸⁵ as well as on the complexities of exercising influence over third parties (paragraph 13b, “linkages” and Principle 19 on “leverage”).¹⁸⁶ Mitigatory measures would then have to be assessed for their ability to address such root causes with a view to eliminating human rights harms from value chains.

There are methodologies, such as Root Cause Analysis (RCA) that are purpose-built for thoroughness of assessments and corrective measures to prevent (re)occurrence of harm.¹⁸⁷ Their necessity derives from well known problems, as Duke Okes writes: “People and organizations often don’t believe they have the time to perform the in-depth analyses required to solve problems. Instead, they take remedial actions to make the problem less visible and implement a patchwork of ad hoc solutions they hope will prevent

183 See *Interpretive Guide*, *supra* note 14, at 6 (“human rights due diligence comprises an ongoing management process”).

184 See, e.g., FAIR LABOR ASSOCIATION, *supra* note 155.

185 Mark Anner et al., *Toward Joint Liability in Global Supply Chains: Addressing the Root Causes of Labor Violations in International Subcontracting Networks*, 35 COMP. LAB. L. & POL’Y. J. 1-3 (2013) (examining the sourcing practices of the brands and retailers that coordinate these supply chains as a root cause of sweatshop conditions in international subcontracting networks).

186 See *Interpretive Guide*, *supra* note 14, at 49; see also *Guiding Principles*, *supra* note 1, Principle 13(b), at 14, Principle 19, at 18.

187 See *infra* note 193 and accompanying text.

recurrence . . . ”¹⁸⁸

This article uses the notions of root causes and root cause analysis to convey the importance of a thorough investigation—in-depth and comprehensive—of visible and less visible factors leading to infringements of human rights. It is not essential for the current purposes to define root causes precisely and distinguish them clearly from contributing factors. Indeed, HRDD should be able to tackle both root causes—without which the impact would not take place—and contributing factors—without which the impact would still occur but to a reduced degree—through appropriate action. What matters is for the HRDD process, namely the impact assessment¹⁸⁹ and tracking stages,¹⁹⁰ to not stop prematurely in identifying and addressing deeper causes.

As summarized by James Rooney, the RCA is a four-step process involving: (1) data collection; (2) causal factor charting (“a sequence diagram . . . that describes the events leading up to an occurrence, plus the conditions surrounding these events. . . . Causal factors are those contributors . . . that, if eliminated, would have either prevented the occurrence or reduced its severity”); (3) root cause identification (a “Root Cause Map to identify the underlying reason or reasons for each causal factor”); and (4) recommendation generation and implementation.¹⁹¹ For instance, RCA was fruitfully used in the area of public health. This method was chosen for its ability to pinpoint the underlying causal relationships associated with a global health incident and to propose recommendations for preventing recurrence:

RCA allows a comprehensive system-wide perspective that breaks down complex global health problems into increasingly smaller components, enabling in-depth analysis from one level and dimension to the next. RCA rests on the premise that getting at the root cause of a problem is more effective than addressing “immediately obvious” symptoms, and that a problem typically has more than one root cause.

188 DUKE OKES, ROOT CAUSE ANALYSIS: THE CORE OF PROBLEM SOLVING AND CORRECTIVE ACTION 1 (2009).

189 See *Guiding Principles*, *supra* note 1, Principle 18, at 17–18.

190 See *id.* Principle 20, at 19.

191 James J. Rooney & Lee N. Vanden Heuve, *Root Cause Analysis for Beginners*, 37 QUALITY PROGRESS 45, 46–49 (2004).

[The] central features of the methodology include recursive questioning to identify causal factors and root causes related to a problem, and identifying effective solutions to prevent recurrence.¹⁹²

The GPs, especially Principle 17, missed the opportunity to impress that a root cause orientation should be an explicit and indispensable parameter of HRDD. The question is whether other provisions in the GPs hint in this direction, even if maybe only obliquely so. Further reference will be made to subsequent documents to which Ruggie is related, such as the Interpretive Guide and Shift reports. Such subsequent materials are a useful interpretive aid in an attempt to compensate for this oversight in the GPs themselves.

B. “Root Causes” in the GPs and Subsequent Documents

Nowhere do the GPs use the notion of root causes. The Interpretive Guide however explicitly indicates that:

[W]herever a significant human rights impact has occurred, the enterprise is well advised also to undertake a root cause analysis or equivalent process to identify how and why it occurred. This kind of process can be important if the enterprise is to prevent or mitigate its continuation or recurrence. A root cause analysis can help pinpoint what actions by which parts of the enterprise, or by which other parties related to the enterprise, played a role in generating the impact, and how.¹⁹³

This is the only reference to root cause in the Guide and it is confined to Principle 20, which captures one out of the four stages of HRDD, namely tracking policy effectiveness. A range of other reports issued by Shift, the non-profit organization chaired by Ruggie and dedicated to the implementation of the GPs, refer repeatedly to root causes.¹⁹⁴ A report written for the European Commission promotes

¹⁹² Jennifer Prah Ruger, *Control of Extensively Drug-Resistant Tuberculosis (XDR-TB): A Root Cause Analysis*, 3 GLOBAL HEALTH GOVERNANCE, April 2010, at 2–3.

¹⁹³ *Interpretive Guide*, *supra* note 14, at 54 (emphasis omitted).

¹⁹⁴ See e.g., SHIFT, RESPECTING HUMAN RIGHTS IN GAP INC.’S GLOBAL SUPPLY CHAIN, 2 (2013), <http://gapinc.com/content/attachments/sersite/Shift-UNGPpdf>.

“a Systematic Approach to Tracking” and mentions RCA among the “Possible Approaches”:

Where a severe human rights impact has occurred, or lesser impacts occur repeatedly, O&G [oil & gas] companies should consider a deeper analysis of the underlying or “root” causes of the incident. Initial impressions may suggest that the company’s own actions or decisions had nothing to do with the impacts; but in some cases a deeper analysis might reveal that it did in fact play a role, and show how it could help prevent the same thing from recurring. Many O&G companies already have experience of applying root cause analysis to major health and safety or environmental incidents. They may be able to adopt and adapt these methodologies for human rights issues as well.¹⁹⁵

This report also addresses the supply chain context. There, despite monitoring and auditing systems currently being used by many lead firms, “[t]hey often miss issues due to their brief nature; [and] may fail to grasp the bigger picture or root cause of repeated human rights impacts.”¹⁹⁶ More collaborative approaches between brands and suppliers include “[s]upporting or analysing the root cause(s) of significant impacts. This can test the conclusions drawn from audits and find any underlying problems.”¹⁹⁷

Another Shift report comments on the clothing company GAP’s approach to supply chain management and commends its use of RCA:

Notably, root cause analysis is an integral part of Gap Inc.’s processes. This helps Gap Inc. understand the underlying causes of failures to respect workers’ human rights and helps Gap Inc. see where its own actions might contribute to negative human rights impacts. Gap Inc.’s transparency in recognizing the relevance of purchasing practices for its vendors’ ability to respect human rights is commendable, and

195 *Oil and Gas Guide*, *supra* note 116, at 55.

196 *Id.* at 59.

197 *Id.*

provides a powerful example to other companies.¹⁹⁸

This statement reflects well the emphasis the GPs place on RtR in value chains and recognizes that brands are not only linked, but contribute to suppliers' non-observance of labor rights. These purchasing decisions of brands are a "contribution" under paragraph 13a;¹⁹⁹ they are a significant contributing factor, if not a root cause, to sweatshop practices in supplier factories.²⁰⁰ For a long time, this was a contributing factor that had not been acknowledged by companies and by much of the CSR discourse, which emphasized the auditing efforts of brands as a commendable but seemingly doomed effort in front of determined non-compliance from suppliers.²⁰¹ Indeed, there has been "an implicit assumption that most non-compliances occur in spite of the efforts of Participating Companies."²⁰² This position of brands as contributors to non-compliance is now openly and belatedly acknowledged by leading companies, including Nike, where its own procurement decisions appeared as a root cause, or at least a key factor of risk, for supplier non-compliance—a risk factor of staggering size: 68%.²⁰³ The same happens when suppliers' workers are deprived of freedom of association; the Fair Labor Association (FLA) and the Ethical Trading Initiative (ETI) have been emphasizing this since mid-2000s,²⁰⁴ and the Bangladesh Accord

¹⁹⁸ SHIFT, RESPECTING HUMAN RIGHTS IN GAP INC.'S GLOBAL SUPPLY CHAIN, 2 (2013), <http://gapinc.com/content/attachments/sersite/Shift-UNGP.pdf>.

¹⁹⁹ See *supra* Section III.A.

²⁰⁰ See, e.g., RUGGIE, *supra* note 47, at 1 (referring to Apple's decision to change product specifications on short notice creating difficulties for its supplier and its workforce).

²⁰¹ See Helle Bank Jørgensen et al., *Strengthening Implementation of Corporate Social Responsibility in Global Supply Chains*, WORLD BANK GROUP [WBG] 53–54, 101 (2003).

²⁰² SHIFT, IMPLICATIONS OF THE UN GUIDING PRINCIPLES ON BUSINESS AND HUMAN RIGHTS FOR THE FAIR LABOR ASS'N, 18 (July 2012).

²⁰³ Nike indicates that, "During FY11, more than two-thirds (68 percent) of the excessive overtime incidents identified and analyzed through audits of 128 factories were attributable to factors within Nike's control, primarily forecasting or capacity planning issues, shortened production timelines and seasonal spikes." NIKE, INC., FY10/11 SUSTAINABLE BUSINESS PERFORMANCE SUMMARY 53 (2012).

²⁰⁴ See ETHICAL TRADING INITIATIVE, FREEDOM OF ASS'N IN COMPANY SUPPLY CHAINS – A PRACTICAL GUIDE, (2013), http://s3-eu-west-1.amazonaws.com/www.ethicaltrade.org.files/shared_resources/foa_in_company_supply_chains.pdf?j11porNelxmYq49FHPmB3IMLySLe54k8.

involves unions in its governance and has supported a strengthening of freedom of association regulations.²⁰⁵ The failure to identify and address such root causes in global value chains explains the failure (or limited success) of mitigation measures during the last 20 years.²⁰⁶ These mitigation measures cannot be made more effective by simply enhancing audits and offering training for management on labor rights.

In another report, Shift reviewed the FLA, which started in the 1990s and is one of the flagship initiatives on labor rights in supply chains.²⁰⁷ Shift commended FLA for emphasizing the abovementioned risk that brands contribute to human rights abuses through their sourcing practices. It added: “Also noteworthy is the FLA’s recent move away from pure compliance auditing towards the incorporation of more root cause analysis and capacity building approaches, aimed at more effective and sustainable mitigation of risks to workers’ rights.”²⁰⁸ The report also remarked on FLA’s own assessment of a switch “from ‘Independent External Monitoring’ visits and reports—which followed classic auditing approaches—to a new approach of ‘Sustainable Compliance’, which involves Independent External Assessments focused on ‘uncovering root causes of problems and providing systemic, sustainable solutions so that problems are fixed in a lasting way.’”²⁰⁹ Currently, the FLA uses its Sustainable Compliance methodology,²¹⁰ which emphasizes root causes, and its reports on affiliate companies include “a description of the root causes of violations, recommendations for sustainable and immediate improvement, and the corrective action plan for each risk or violation as submitted by the company.”²¹¹

205 The governance of the Accord involves companies and trade unions in equal numbers with the ILO being a neutral Chair. Bangladesh Accord Secretariat, *Introduction to the Accord on Fire and Building Safety in Bangladesh*, THE ACCORD ON FIRE AND BUILDING SAFETY IN BANGLADESH, (Jan. 2015), <http://bangladeshaccord.org/wp-content/uploads/Brochure-about-the-Accord.pdf>.

206 Int’l Labor Org., *Decent work in Global Supply Chains*, at 47-51 (2016) (analyzing the potential and limitations of private governance and CSR).

207 SHIFT, *supra* note 183. Fair Labor Association is ‘a collaborative effort of socially responsible companies, colleges and universities, and civil society organizations’ started in 1999. FAIR LABOR ASS’N, <http://www.fairlabor.org> (last visited Oct. 16, 2017).

208 SHIFT, *supra* note 183, at 4.

209 *Id.* at 7.

210 See Mares, *supra* note 123, at 118–20.

211 FAIR LABOR ASS’N, INDEPENDENT EXTERNAL ASSESSMENT REPORT: HUGO BOSS (2015), <http://www.fairlabor.org/affiliate/hugo-boss>.

Shift also noted the increasing use of RCA by brands. The clothing company Phillips-Van Heusen (PVH) indicates that they

Now engage factories in corrective action planning (CAP) development earlier in the assessment process so more time is spent on implementing the action plans. During this process, our team works with factories to identify root causes, offer suggested actions, and develop feasible remediation plans. This dialogue also provides an opportunity to understand how our purchasing practices may hinder our factories abilities' to comply.²¹²

Finally, a Shift report places RCA in the context of building credibility with stakeholders by enhancing transparency and demonstrating a rigorous approach:

Many complex human rights challenges within supply chains do not have immediate or easy solutions. Time may be required for root cause analysis, for industry-wide collaboration, for increasing company leverage to enable action with a supplier, and for identifying appropriate and effective remediation measures. In such instances, increased transparency can be particularly beneficial, enabling a company to convey internally and externally the seriousness with which it is treating challenging issues . . .²¹³

Nike indicates that to address excessive overtime, the company is focusing on a “continued analysis of root causes, which has led us to identify and address key business processes upstream from the factory.”²¹⁴ The search for root causes can go as far as challenging business models, such as overly flexible supply chains.

212 SHIFT, EVIDENCE OF CORPORATE DISCLOSURE RELEVANT TO THE UN GUIDING PRINCIPLES ON BUSINESS AND HUMAN RIGHTS 12 (2014), https://www.shiftproject.org/media/resources/docs/Shift_EvidenceCorporateDisclosureUNGPs_2014.pdf.

213 SHIFT, RESPECTING HUMAN RIGHTS THROUGH GLOBAL SUPPLY CHAINS, 18 (2012).

214 NIKE, INC., *supra* note 184.

Nike speaks of optimizing its factory base: “We have moved toward establishing long-term relationships with fewer factories as trusted partners, rather than having short-term transactional relationships with a larger number of factories.”²¹⁵ Indeed, short-term relationships appear as a risk factor, and clearly are a contributing factor of high importance if not a root cause:

When supply chains consolidate, it is an opportunity to align shared values. Several companies participating in the research detected a growing trend to consolidate supply chains. Companies are creating deeper relationships with fewer suppliers, particularly strategic ones. In the process, purchasers focus on long-term value, grounded in expertise more than price. This implies closer and longer-term relationships, and sharing of standards and systems. It may be that this process will make it easier to integrate human rights, ethical values and good practices in supply chains.²¹⁶

In his writings, Ruggie took note of Professor Richard Locke’s work dedicated to improving labor standards in supplier factories.²¹⁷ Employing a root cause orientation, Locke’s approach was developed as a response to the auditing/compliance approach.²¹⁸ The focus of Locke’s approach is on uncovering, analyzing, and correcting root causes.²¹⁹

In 2010, three prominent organizations active in CSR issued the Guide to Human Rights Impact Assessment and Management, with an introductory quote from Ruggie.²²⁰ This work introduces the “mitigation hierarchy” concept (avoid, reduce, restore, compensate). Companies are expected to develop Appropriate Mitigation Action Plans in order to “accurately address previously identified human

215 *Id.* at 31.

216 THE INST. FOR HUMAN RIGHTS AND BUS. & THE GLOB. BUS. INITIATIVE ON HUMAN RIGHTS, STATE OF PLAY: THE CORPORATE RESPONSIBILITY TO RESPECT HUMAN RIGHTS IN BUSINESS RELATIONSHIPS 109 (2012).

217 See RUGGIE, *supra* note 47, at 206 n.12.

218 RICHARD LOCKE, THE PROMISE AND LIMITS OF PRIVATE POWER: PROMOTING LABOR STANDARDS IN A GLOBAL ECONOMY 2, 174 (2013).

219 *Id.* at 78, 123–25.

220 DÉSIRÉE ABRAHAMS & YANN WYSS, GUIDE TO HUMAN RIGHTS IMPACT ASSESSMENT AND MANAGEMENT (HRIAM), 5 (2010).

rights risks and impacts” by employing the “mitigation hierarchy.”²²¹ The Guide leaves companies discretion in adopting measures suitable to the four levels in the mitigation strategy.²²² For instance, the wording emphasizes reduction of impact: “Guided by the mitigation hierarchy, companies may consider designing measures that reduce any negative human rights impact.”²²³ Even for severe impacts, companies seem to retain some discretion on measures falling short of the goal of elimination (*avoid impacts*): “The severity and magnitude of the human rights risk and impact will determine which type of mitigating approach to pursue. Human rights impacts that pose significant and immediate risk to the health, safety and the lives of stakeholders will benefit from avoidance, reduction and/or restoration measures.”²²⁴ The Guide also indicates that “appropriate actions” should be based on the “technical and financial feasibility of the actions and measures required.”²²⁵ That technical and financial aspects are relevant is indisputable, but such acknowledgement should be doubled by an explicit prescription to identify and address root causes, especially for severe potential impacts (high risk to human rights). In sum, the mitigation hierarchy concept is useful because it explicitly mentions the aim of avoidance (elimination) under the umbrella term of mitigation, which otherwise could have been misconstrued as requiring reduction only. However, the Guide speaks casually of companies moving among the levels of the hierarchy to *reduce* impacts (for “financial feasibility” reasons,²²⁶ for example) and does not impress the importance of designing mitigatory measures capable of identifying and addressing root causes.

Overall, the Guide does not appear fully attuned to the human rights context and able to explain the specificities of HRDD as a rightholder-centred risk management process. By contrast, other instruments issued by the International Labor Organization employ a more precise formulation. Thus, the ILO Safety and Health

221 *Id.* at 49.

222 See *id.* at 12–13; See also DANISH INST. FOR HUMAN RIGHTS & IPIECA, INTEGRATING HUMAN RIGHTS INTO ENVIRONMENTAL, SOCIAL, AND HEALTH IMPACT ASSESSMENTS 17 (2013) (seeking to “apply a human rights lens to impact mitigation and management” and promoting the same hierarchy).

223 ABRAHAMS & WYSS, *supra* note 201, at 49.

224 *Id.*

225 *Id.* at 50.

226 *Id.*

in Mines Convention refers to the responsibilities of employers to take preventive and protective measures:

[T]he employer shall assess the risk and deal with it in the following order of priority: (a) eliminate the risk; (b) control the risk at source; (c) minimize the risk by means that include the design of safe work systems; and (d) in so far as the risk remains, provide for the use of personal protective equipment, having regard to what is reasonable, practicable and feasible, and to good practice and the exercise of due diligence.²²⁷

A final remark is reserved for the relevance of a root cause orientation at both the ex ante and ex post stages of HRDD. HRDD should address deeper underlying causes in order to be genuinely capable of reaching the elimination goal of RtR. This prescription applies at the ex ante stage of human rights potential impacts, which is human rights impact assessments (HRIAs),²²⁸ as well as the ex post stage of actual impacts, which is tracking the effectiveness of measures taken.²²⁹ As shown, the Interpretive Guide refers expressly to RCA in relation to “tracking” so that a company analyzes root causes to prevent re-occurrence.²³⁰ It is an investigation with hindsight. Is there an equivalent notion to root causes identifiable in Principle 18 dealing with HRIA? Given that there are no impacts that occurred it follows that there are no root causes of damage for a RCA to investigate. At the ex ante stage, where risks of harm are to be identified, one properly speaks of *factors of risk*. It is an investigation that requires foresight. Identifying these factors of risk—some more hidden than others—and their interrelations requires a thorough analysis likeminded with the RCA’s search for root causes and contributory factors. Is there a dedicated notion in the GPs that impresses the importance of deeper factors of risk and of special methodologies to identify them?

In the GPs, there are no explicit references to significant factors of risk equivalent to root causes as the terminology is uniformly

227 Int'l Labor Org., Safety and Health in Mines Convention, art. 6, June 22, 1995, No. 176.

228 See *Guiding Principles*, *supra* note 1, Principle 18, at 17–18.

229 *Id.* Principle 20, at 19.

230 *Interpretive Guide*, *supra* note 14; *Oil and Gas Guide*, *supra* note 116, and accompanying text.

about “risk of adverse human rights impacts” and “potential adverse human rights impacts.”²³¹ Rather, the GPs expect HRIAs to identify all factors of risk.²³² However, the GPs explicitly refer to importance of foresight efforts and methods: HRIAs should be “projecting how the proposed activity and associated business relationships could have adverse human rights impacts on those identified.”²³³ As early as 2007, Ruggie issued a report dedicated to HRIA methodologies where he emphasized foresight efforts: “After describing those baseline conditions, HRIAs should put forth a view of what is likely to change because of the business activity. This is a difficult and subjective exercise; one approach is to construct multiple scenarios, while another might predict outcomes based on varying levels of intervention.”²³⁴ Thus, Ruggie highlighted some methods able to address root causes at the ex ante stage of HRDD. Even in 2017 there is no authoritative methodology dedicated to HRIAs, while efforts are on-going.²³⁵

In addition to methodology, the GPs offer another angle to get to the deeper factors of risk: emphasizing the level of effort placed in HRIAs. Thus, as Shift noted in high-risk contexts, where there are significant risks to human rights, the company should pursue a deeper impact assessment. The level of effort of the inquiry needs to be enhanced to match the severity of adverse impacts and this should also bring to the surface deeper factors of risk. “Operating in high risk contexts . . . requires greater attention, effort and resources at every step of the process [of assessing human rights impacts].”²³⁶ Thus companies, for example, should seek “to understand the root causes of a conflict (for example ethnic tensions or access to resources) and their implications for human rights and for company operations.”²³⁷

Both these ex ante and ex post investigations of human rights

231 *Guiding Principles*, *supra* note 1, Principle 17–18, at 16–18.

232 *Id.* Principle 18, at 17 (“business enterprises should pay special attention to any particular human rights impacts on individuals . . .”).

233 *Id.*

234 John Ruggie (Special Representative to the Secretary General), *Human Rights Impact Assessments – Resolving Key Methodological Questions*, ¶ 14 U.N. Doc. A/HRC/4/74 (Feb. 5, 2007).

235 Dylan Tromp, *Assessing Business-Related Impacts on Human Rights - Indicators and Benchmarks in Standards and Practice*, INEF-Report, (2016), <http://inef.uni-due.de/cms/files/report110.pdf>.

236 *Oil and Gas Guide*, *supra* note 116, at 29.

237 *Id.*

impacts seek to uncover hidden and less understood elements—factors of risk or root causes—that, if overlooked, reduce or cancel the effectiveness of HRDD efforts. As the Guide states, “the processes for assessing human rights impact should be systematic so that the various elements add up to a coherent overview of actual and potential human rights impact associated with an enterprise’s activities and relationships and can accurately inform the subsequent steps in the due diligence process.”²³⁸ This emphasis on root causes also permeates the efforts of regulators in high-risk sectors, such as conflict minerals. The European Commission wrote:

One of the objectives of the EU’s proposal is to break the link between minerals extraction, minerals trading, and the financing of armed conflicts. The root causes of the problems must be identified, as should the triggers of conflicts and structural fragility, their dynamics, and the roles of the various actors involved.²³⁹

Such emphasis on root causes orientation of HRDD allows some final observations regarding operational measures expected under the GPs.

C. Implications for HRDD

The analysis so far has clarified when mitigation (reduction of impacts) exceptionally is a legitimate aim (i.e. third party operations), and highlighted the peril that mitigatory measures could be misdirected if root causes are not addressed. In the latter case, HRDD would be less effective or could be used strategically—with negative effects for the credibility of the GPs—to create a semblance of compliance with the GPs, an illusion that a company’s efforts address impacts rigorously. The emphasis on deeper causes prevents HRDD from becoming a partial and misdirected response and from disguising the absence of more demanding measures needed to achieve the aim of RtR of elimination. This subsection brings

238 *Interpretive Guide*, *supra* note 14, at 41.

239 EUROPEAN COMMISSION, FREQUENTLY ASKED QUESTIONS - RESPONSIBLE SOURCING OF MINERALS ORIGINATING CONFLICT-AFFECTED AND HIGH-RISK AREAS: TOWARDS AN INTEGRATED EU APPROACH (March 5, 2014), http://europa.eu/rapid/press-release_MEMO-14-157_en.htm.

together the *mitigation* and *root causes* characteristics of the HRDD effort so it does not succumb to easy compliance and inadequate risk-management techniques.

The goal of elimination is neither about offering a guarantee that impacts will not occur nor about aiming to avoid activities that might have adverse impacts. These clear-cut prescriptions are neither feasible nor achievable in practice, except for exceptional situations.²⁴⁰ As Linda Spedding notes, “[i]t may not be cost effective, or just unachievable, to remove risk altogether by risk management.”²⁴¹ Risk avoidance can be achieved “through strategic decisions such as withdrawing from a market sector or region,” but this form of risk treatment—by avoiding risks—is deemed “very limited” in practice.²⁴²

HRDD is not simply about reduction of adverse impacts, but about reduction with a view to full elimination—not only with the view, but also with the capacity to achieve this. Given that avoidance is a limited option in practice,²⁴³ addressing root causes becomes the main option to minimize impacts with a view to their elimination. It is about the reasonable person’s design of corrective measures that have the intrinsic capacity to eliminate rights abuses from its operations. Thus, compliance is not merely a function of effort (enough resources) and time (fast enough)—it is the design of measures and their intrinsic capacity to eliminate or not harm that also matter. So, reduced impacts—through mitigatory measures—can actually constitute an abuse of human rights when prevention (factors of risk unaddressed) and correction (root causes unaddressed) were substandard due to misdirected effort, insufficient amount of effort, or inappropriate risk tolerance level. There are implications for HRDD both in terms of prevention and remediation, at both the ex ante and ex post stages of HRDD.

At the prevention stage, there is rarely the chance to offer guarantees that human rights impacts will not occur throughout their operations, or even that they will not reoccur. What can be reasonably done in terms of prevention, in terms of managing risks (potential impacts) of occurrence and/or reoccurrence? The

240 For example, in case of grave abuses, the risks would be better shouldered by companies, from a policy point of view.

241 LINDA S. SPEDDING, DUE DILIGENCE AND CORPORATE GOVERNANCE 158 (Butterworth-Heinemann, 1st ed. 2005).

242 *Id.* at 284–85.

243 *Id.* at 157–58.

definition in the Interpretive Guide that prevention refers to “actions taken to *ensure* such impact does not occur”²⁴⁴ appears aspirational. How can a company ensure non-occurrence? At first sight, the only way that prevention, so defined, can be achieved is through avoidance of operations in a risky environment. At a second glance, however, prevention can be achieved by addressing root causes. It is even possible to remove some root causes from the chain of factors causing harm. For example, to prevent harm to Amazonian tribes, an oil company in Camisea, Peru, flies personnel to the site instead of building roads through the forest; these measures were chosen because those roads were previously used by illegal loggers to expand operations and affect tribes with deadly consequences.²⁴⁵ Another example could refer to inherently dangerous technological processes; for the sake of argument, instead of using cyanide in large scale mining, other chemicals could be used and could thus altogether remove the hazards generated by that specific chemical. Such an emphasis on tackling root causes gives a fuller meaning to prevention as defined by the Guide and does not severely confine the applicability of this concept to either a—still rare—prescription to avoid altogether planned activities in risky environments, or to a—still exceptional—prescription to offer guarantees (insurance) to rightholders.²⁴⁶ Attention to root causes and deeper factors of risk should be an essential trait of rightholder-centred risk-management in the GPs.

Regarding remediation, the presence of effective grievance mechanisms that enterprises might offer guarantees, at best, that impacts do not turn into unremedied impacts (abuse). Actual impacts get addressed. However, all grievance mechanisms—

244 *Interpretive Guide*, *supra* note 14 (emphasis added); *see also supra* text accompanying note 14 (emphasis added).

245 *Drilling in the Wilderness - Energy Extraction can Coexist with Native Peoples and Forests*, THE ECONOMIST, (Apr. 26, 2014), <http://www.economist.com/news/americas/21601267-energy-extraction-can-coexist-native-peoples-and-forests-drilling-wilderness>.

246 Both are of limited applicability: the former because business, developmental and even human rights considerations make such avoidance undesirable and/or unfeasible, and the latter because asking lead firms to offer guarantees by assuming responsibility for harm in supplier operations is at odds with foundational business law principles of separate corporate personality and limited liability. See Radu Mares, *Legalizing Human Rights Due Diligence and the Separation of Entities Principle*, in BUILDING A TREATY ON BUSINESS AND HUMAN RIGHTS: CONTEXT AND CONTOURS (S. Deva & D. Bilchitz eds., Cambridge Univ. Press, forthcoming Oct. 2017).

judicial or non-judicial—have obstacles that reduce access to justice for rightholders.²⁴⁷ So, setting up effective remedial mechanisms and reducing such obstacles becomes the main option to minimize actual impacts with a view to their elimination at the ex post stage of HRDD. It is as close as one gets to a guarantee that impacts—or abuses—are eliminated. Parenthetically, the GPs explain RtR in terms of adverse “impacts” but also refers to “abuses” of human rights.²⁴⁸ When does impact turn into abuse? A right might have been infringed, but as long as reasonable prevention and full remediation are offered, impact does not turn into abuse. Abuse is unprevented and/or unremedied impact. So, it is correct for Pillar 2 to speak of impacts as the entire pillar is about preventive and remedial measures.

The goal of elimination is not about an aspiration contained in an abstract RtR that in practice, in real life, boils down to mitigating (reducing) impacts, which would appear as a worthwhile step forward anyway. This is not the way to judge the adequacy of HRDD and is inappropriate in a human rights context. This is not how the relationship between the RtR and HRDD in the GPs should be understood. The GPs are about a correctly designed HRDD that aims and has the potential to avoid impacts, even if in practice such HRDD might fall short. The GPs are grounded in human rights normativity and cover a special type of adverse impact. Although the GPs speak of “impacts” and not of “abuses,” and employ a risk-management language colored by mitigation terminology hinting at a reduction of impact, the GPs do not tolerate residual abuse of human rights as the accepted cost of doing business in a global economy riddled by governance gaps. The RtR aim of reduction is tolerated in one context only—third party operations—based on the understating that summoning the leverage of lead firms is a strategic opportunity for the human rights system to deliver enhanced protection. Thus, the GPs as a governance framework and as a corporate risk-management framework are geared towards elimination of abuses and require a HRDD concept that is designed

²⁴⁷ *Guiding Principles*, *supra* note 1, Principle 26, at 23–24. (speaking of such obstacles). See also United Nations High Commissioner for Human Rights, *Improving accountability and access to remedy for victims of business-related human rights abuse*, U.N. Doc. A/HRC/32/19 (May 10, 2016).

²⁴⁸ See *Guiding Principles*, *supra* note 1, Principles 11–15 at 17–22 (containing many references to “impacts”); *Guiding Principles*, *supra* note 1, Principles 17, 19, 23 Commentaries (referring to “abuse”).

to deliver that operationally.

VI. Conclusions

Mitigation was a notion not used once in the 2008 Protect-Respect-Remedy Framework.²⁴⁹ In 2011, the GPs explained the RtR and HRDD by making repeated references to mitigation. Mitigation of adverse impacts means *reduction* of impact, which potentially entails that residual impact is tolerated. This opens the troublesome possibility that businesses will understand the RtR as demanding something less than *elimination* of human rights impacts from their operations. This article started from the concern that having HRDD refer to mitigation might result in redundancy—no value added to what prevention and remediation already say)—or misunderstanding—water down the RtR aim of elimination)—and thus reduce the persuasive force of the GPs. Therefore, the article sought to clarify mitigation in terms of application to organizational contexts, its functions, its meanings, and the specificities of the human rights context.

This analysis pinpoints the applicability of mitigation (aim of reduction of impact) to one organizational context only: that of third party conduct, which comprises the “contribute” and “linkages” scenarios in Principle 13. Second, mitigation of impacts is a notion overlapping two meanings—aim and measures. Mitigation measures is just another term for preventive and remedial measures. However, mitigation is a charged concept introduced into a context where residual abuse cannot be accepted as the normal price of doing business, as the entire human rights system pursues the aim of elimination of infringements of human dignity. Therefore, as a difference from other contexts, some traditional risk-management strategies are illegitimate. Businesses and risk management professionals do not have the liberty to choose among different risk management techniques that they have for other risks—to simply take the risk, or to insure against it, or to reduce the risk/impact. The aim here is to eliminate adverse human rights impacts from a business’ operations. Clarifying mitigation helps HRDD become a genuine rightholder-centred risk management approach to human rights. HRDD means taking measures that reduce impact with an aim to its elimination and with the capacity to achieve that aim too. Third, mitigation together with the notion of leverage, serves

249 See *Protect, Respect and Remedy*, *supra* note 26.

two specific functions in the GPs, both in the context of third party operations. On the one hand, mitigation pinpoints a key HRDD step—leverage first, termination of relationship second—in dealing with non-compliant third parties, a stop-gap towards the aim of elimination. On the other hand, mitigation expands the scope of RtR regarding remediation by making it imperative for the company press for securing rightholders' access to remedies.

The GPs offer numerous parameters and insights into what a DD approach entails in a human rights context. One that is missing from the GPs is an emphasis on *root causes* of adverse impacts. It is essential for risk management to tackle the root causes of impacts, rather than the symptoms and contributing factors. Mitigatory measures that are misdirected indicate measures that lack the capacity, if not the declaratory aim, to eliminate human rights abuses from business operations. Tackling root causes should be a parameter of HRDD in Principle 17. Subsequent documents released by Ruggie and his colleagues refer explicitly to root causes and root cause analysis methodologies designed to ensure a systematic and thorough pursuit of deeper causes. This parameter is also important because of the silences in the GPs on “amount of effort” (impossible to define in abstract, *ex ante*), the absence of a centralized reviewer to monitor and enforce the GPs (unattainable politically at the time), and the unfeasibility of strict liability theories under the RtR (of exceptional application only due to conflict with default principles of civil and criminal liability). Therefore, insisting that the HRDD effort gets at least directed correctly towards root causes allows the assessment of whether effort has been misdirected or not and thus whether mitigatory measures inherently lacked the very capacity to achieve elimination.

This analysis helps alleviate the problem of complacency of HRDD and the potential misalignment of HRDD and RtR in the GPs. On the one hand, businesses routinely insist they make effort, that it takes time, that impacts have been reduced, and that they are not perfect but trying hard. The peril of complacency is there, as companies design their actions to demonstrate reduction in negative impacts and claim credit for the results for their efforts. So, is the danger that stakeholders witness an elaborate display of preventive and corrective actions with little indication that the root causes of impacts are actually being considered and addressed. There are already concerns from CSR—and GPs-skeptics that the GPs have put forward an overly process-oriented and risk-management approach

that leaves companies too much flexibility and discretion. On the other hand, HRDD and RtR might get out of sync if the notion of mitigation is not clarified and understood in the economy of the GPs. In the GPs, RtR, which carries the aim of elimination, appears as the abstract, general, and aspirational part, might be overlooked with attention going to HRDD, which speaks of mitigation and appears operational, more specific, and more familiar to risk managers and decision-makers. Indeed, the HRDD terminology has gone mainstream following the GPs, is recognizable in other CSR instruments and in casual explanations that companies have to mitigate human rights impacts. In sum, preventing complacency and misalignment within the GPs, helps ensure the internal consistency of the GPs, clarify HRDD as a rightholder-centred risk management approach appropriate to the human rights context, and ultimately enhance the persuasive power of the GPs.

* * *

Collaborative Enforcement

*Andrew Elmore**

Abstract

Labor standards enforcement in the low-wage workplace has long suffered from a lack of capacity, expertise and remedies that blunt the impact of public and private enforcers alike. The question of how to address these pathologies in state and local workplace regulation has gained new urgency with the virtual explosion of regional labor lawmaking and the deregulatory impulses of the new federal administration.

This Article identifies collaboration between state and local agencies and private, public interest organizations (“PIOs”) as one pathway to address these enforcement gaps, by amplifying the deterrent effect of public and private enforcement and by improving legal remedies. This Article offers this form of public-private regulatory experimentation, which it calls “collaborative enforcement,” as a conceptual framework that can (a) effectively and efficiently address enforcement gaps by integrating a range of enforcement tools that public and private enforcers cannot access independently; (b) subject public agency enforcement priorities to political accountability; and (c) facilitate sophisticated types of tripartite regulation championed by earlier scholarship.

Private delegations in collaborative enforcement, however, can create a risk of PIO abuse of the delegation and of public agency cooptation of PIOs, which will require measures to protect public agency and PIO independence.

* Acting Assistant Professor, New York University School of Law. The author was a Section Chief in the New York Attorney General’s Labor Bureau until June 2015. The views expressed in this Article are the author’s alone and do not reflect the views of the Bureau or of the Office. A special thank you is extended to Scott Cummings, Cynthia Estlund, Jennifer Gordon, Roderick Hills, Margaret Lemos, Deborah Malamud, Jon Michaels, and Noah Zatz for providing detailed and insightful feedback on the thoughts and ideas expressed in this Article, and for helpful comments and suggestions, the author thanks Kate Andrias, Miriam Baer, Susan Carle, Adam Cox, Nestor Davidson, Janice Fine, Michele Gilman, Daryl Levinson, Joseph Mastrosimone, Benjamin Sachs, and Deborah Anne Widiss. Thank you also to faculty members at the CUNY School of Law, Miami School of Law, and UCLA School of Law, and to participants in the Colloquium on Scholarship in Employment and Labor Law and the AALS Emerging Voices in Workplace Law program, for thoughtful feedback. The author also thanks Bing Le and James Yang, and the editors of the Northeastern University Law Review, particularly Luciana Rivera and Brandon Sisson, for outstanding research assistance. All errors are the author’s.

National Labor Relations Act preemption and state nondelegation doctrine do not threaten the core requirements of collaborative enforcement, but do constrain the scope of its delegations and legislative aims. The techniques described in this Article may be applied to other areas of civil enforcement in which underdeterrence is a result of similar enforcement pathologies.

Table of Contents

I.	Introduction	75
II.	Public and Private Enforcement Pathologies.....	82
	<i>A. The Enforcement Pathologies of Independent Enforcement</i>	83
	1. Public Agency Enforcement	84
	2. Private Enforcement	88
	<i>B. The Limitations of Independently Addressing PIO and Public Agency Enforcement Pathologies.....</i>	93
	<i>C. Tripartism to Deter Rather Than to Encourage Self-Regulation.....</i>	96
III.	Public-Private Regulatory Experimentation to Improve the Deterrent Effect of Enforcement.....	101
	<i>A. Collaborative Enforcement.....</i>	102
	1. Remedial Enforcement	102
	2. Grant-based Enforcement	107
	<i>B. Evaluation of Collaborative Enforcement</i>	110
	1. Collaborative Enforcement Can Effectively and Efficiently Address Enforcement Pathologies	110
	2. Collaborative Enforcement Injects Political Accountability into Public Enforcement and Prevents Capture by the Regulated Entities	113
	3. Collaborative Enforcement Improves Effectiveness and Governance of Public-Private Enforcement Through Tripartism	115
	<i>C. Proposal: Collaborative Enforcement to Amplify the Deterrent Effect of Enforcement</i>	117
IV.	The Risk of Eroding Public and Private Independence and Doctrinal Limitations to Collaborative Enforcement	121
	<i>A. The Erosion of Independent Enforcement.....</i>	122
	1. Public Enforcement Independence	124
	2. Harm to Private Enforcement	127
	<i>B. NLRA Preemption and State Nondelegation.....</i>	129
	1. NLRA Preemption	130
	2. State Nondelegation Doctrine	135
V.	Conclusion.....	139

I. Introduction

Structural changes in the nature of work have made labor and safety and health violations in the low-wage workplace prevalent and difficult to detect and resolve. Declining unionization, the increasing use of temporary and contingent workers, and the “fissuring” of industries in which lead firms shed labor costs by outsourcing them to small, often fly-by-night companies, has led to the proliferation of small, decentralized and geographically dispersed workplaces where noncompliance with legal protections is an entrenched social norm.¹

Public agencies and private, public interest organizations (PIOs) often share the goal of strengthening labor standards through enforcement but lack the capacity, expertise or remedies to change entrenched norms of noncompliance in the informal economy. Adding to this challenge is the recent proliferation of state and local lawmaking to lift standards and enact newly-minted protections, such as mandatory paid sick leave and restrictions on on-call scheduling, that will only achieve their intended effects if they are effectively enforced.² With the retreat of federal agencies from labor standards enforcement,³ and the limited role of the private bar in sectors where

1 See DAVID WEIL, THE FISSURED WORKPLACE: WHY WORK BECAME SO BAD FOR SO MANY AND WHAT CAN BE DONE TO IMPROVE IT 54, 70–73, 245 (2014).

2 Nearly forty cities and states passed local minimum wage ordinances higher than the federal minimum wage since 2012. NAT'L EMP'T LAW PROJECT, CITY MINIMUM WAGE LAWS: RECENT TRENDS AND ECONOMIC EVIDENCE (2016), <http://www.nelp.org/content/uploads/City-Minimum-Wage-Laws-Recent-Trends-Economic-Evidence.pdf>. Seven states and twenty-eight local jurisdictions have passed legislation requiring paid sick leave, and many others restrict employers from requiring employees to call in the day of their shift to determine their daily hours. See Russ Wiles, *More States Adopt Tough Paid Sick-Leave Laws*, USA TODAY (July 5, 2017), <https://www.usatoday.com/story/money/business/2017/07/05/more-states-adopt-tough-paid-sick-leave-laws/103336610/>; Charlotte Alexander, Anna Haley-Lock & Nantiya Ruan, *Stabilizing Low-Wage Work*, 50 HARV. C.R.-C.L. L. REV. 1, 41–46 (2015).

3 Over the past year, the federal administration proposed slashing the United States Department of Labor (USDOL) budget by over 20%, rescinded guidance providing a broad definition of “joint employers” under wage-and-hour law, and has declined to advocate for the Obama-era salary basis threshold expanding overtime coverage for low-wage employees classified as managers. Bourree Lam, *The Jobs Programs Trump's Budget Would Cut*, THE ATLANTIC (Mar. 17, 2017), <https://www.theatlantic.com/business/archive/2017/03/trump-budget-dol/519933/>; Daniel Weissner, *U.S. Labor Department Rescinds Obama-Era Rule on ‘Joint Employment’*, REUTERS (June 7, 2017) <https://www.reuters.com/article/us-usa-labor-jointemployment-idUSKBN17H0JL>.

employers are often small and judgment proof,⁴ this enforcement burden will rest squarely on state and local public agencies and PIOs such as unions, worker centers and community-based organizations.

The thesis of this Article is that to address these new challenges, state and local agencies will need to address longstanding enforcement gaps that undermine the effectiveness of public enforcement. It identifies collaboration between state and local public agencies and PIOs as one pathway to do this, and offers collaborative enforcement a conceptual framework to (a) effectively and efficiently address enforcement gaps by integrating enforcement tools and creating new remedies; (b) strengthen the political accountability of public agency enforcement; and (c) facilitate sophisticated types of tripartite regulation championed by earlier scholarship.

While scholars of public-private regulatory experimentation have traditionally proposed delegating monitoring responsibilities to the regulated entities,⁵ this Article shows that public-private

reuters.com/article/us-usa-labor/u-s-labor-department-rescinds-obama-era-rule-on-joint-employment-idUSKBN18Y2PZ; Nevada v. U.S. Dep’t of Labor, No. 16-41606, 2017 WL 3149309, at *17 (5th Cir. June 30, 2017).

4 Employment law noncompliance tends to be concentrated in small workplaces. Janice Fine & Jennifer Gordon, *Strengthening Labor Standards Enforcement through Partnerships with Workers’ Organizations*, 38 POL. & SOC’Y 552, 555 (2010) (finding that the industries “at greatest risk of FLSA violations are overwhelmingly composed of establishments of fewer than twenty employees”). Employers in small, low-wage workplaces are often judgment proof and more likely to respond to notice of enforcement by closing and reopening under a new name, or by disappearing altogether. Noah Zatz, *Working Beyond the Reach or Grasp of Employment Law*, in THE GLOVES-OFF ECONOMY 47 (2006). Class-action suits by private attorneys can theoretically afford workers a relatively costless and anonymous means to challenge unlawful practices. *Id.* at 46. But they are unrealistic in small workplaces, and federal wage-and-hour lawsuits typically cannot be brought by class action. Catherine K. Ruckelshaus, *Labor’s Wage War*, 35 FORDHAM URB. L.J. 373, 383 (2008). Private, aggregate litigation in employment law, moreover, may soon be closed off entirely in workplaces that require employees to waive participation in class actions. See Transcript of Oral Argument, N.L.R.B. v. Murphy Oil USA, Inc., WL 4882790 (U.S. argued Oct. 2, 2017) (Oral argument regarding whether required class action waivers violate the right for employees to engage in “concerted activities” in pursuit of their “mutual aid or protection” under the National Labor Relations Act.).

5 The best-known discussion of public-private collaboration in regulation is enforced self-regulation. See John Braithwaite, *Enforcement Self-Regulation: A New Strategy for Corporate Crime Control*, 80 MICH. L. REV. 1466, 1467–73 (1982). In this model, public agencies delegate monitoring to the regulated entities, reserving government enforcement resources for monitoring the regulated entities’ internal compliance regimes. *Id.* This model assumes the existence of regulated entities with the willingness and expertise to

regulatory experimentation can also be useful in instances in which self-regulation by the regulated entities is unlikely in the near term. In these instances, instead of enforced self-regulation, public agencies and PIOs can change the behavior of regulated entities through collaboration, with the goal of amplifying the deterrent effect of enforcement, as exemplified by collaboration between PIOs and public agencies in the Fight for Fifteen movement.⁶

Scholarship that has theorized this possibility⁷ has not yet addressed administrative law literature cautioning that privatizing public regulation can create incentives for abuse, defeat accountability and undermine democratic values.⁸ Collaborative enforcement can

design and implement compliance plans. *See generally* IAN AYRES & JOHN BRAITHWAITE, RESPONSIVE REGULATION: TRANSCENDING THE DEREGULATION DEBATE 19–53 (1992); NEIL GUNNINGHAM & PETER N. GRABOSKY, SMART REGULATION: DESIGNING ENVIRONMENTAL POLICY 401–2 (1998) (considering sanctions of private and public enforcers of environmental law).

- 6 *See infra* Part II(C). Legal scholars have focused on the implications of the Fight for Fifteen movement on labor law. *See* Kate Andrias, *The New Labor Law*, 126 YALE L.J. 2, 8 (2016), but have not yet explored its implications for regulation. While beyond the scope of this Article, the benefits and costs of social movement actor collaboration with public agencies deserves a fuller examination. Law and social movement scholarship primarily explores the role of private, public interest attorneys (often in PIOs) reacting to political opportunities to shape the aims of social movements. *See* Scott Cummings, *Movement Lawyering*, 2017 U. ILL. L. REV. 1645, 1669–89 (2017) (describing history of movement lawyering in U.S. from the civil rights era to present). As Douglas NeJaime argues, this focus on private attorneys can obscure the role of government attorneys seeking to advance social movement goals and channeling social movement activities into state-centered tactics. *See* Douglas NeJaime, *Cause Lawyers Inside the State*, 81 FORDHAM L. REV. 649, 654 (2012). Examination of the choice by the Fight for Fifteen movement to seek public agency enforcement to obtain and protect its law reform goals, and by public agency attorneys to prioritize PIO referrals and law reform goals, would provide a fuller understanding of the benefits and costs of collaborative enforcement to the social movements that make such collaboration possible.
- 7 CYNTHIA ESLUND, REGOVERNING THE WORKPLACE: FROM SELF-REGULATION TO CO-REGULATION 15 (2010); JANICE FINE, CO-PRODUCTION: BRINGING TOGETHER THE UNIQUE CAPABILITIES OF GOVERNMENT AND SOCIETY FOR STRONGER LABOR STANDARDS ENFORCEMENT 18–29 (2015), http://theliftfund.org/wp-content/uploads/2015/09/LIFTReportCoproductionOct_ExecSumm-rf_4.pdf [hereinafter CO-PRODUCTION]; Matthew T. Bodie, *The Potential for State Labor Law: The New York Greengrocer Code of Conduct*, 21 HOFSTRA LAB. & EMP. L.J. 183, 194 (2003); Fine & Gordon, *supra* note 4, at 575–76.
- 8 *See* JODY FREEMAN & MARTHA MINOW, *Introduction: Reframing the Outsourcing Debates*, in GOVERNMENT BY CONTRACT: OUTSOURCING AND THE

indeed create incentives for PIOs to misuse delegations and for public agencies to coopt PIOs that might otherwise act as public enforcement watchdogs, requiring political and administrative controls to prevent abuse that will be discussed in Part IV. But a detailed analysis of emergent enforcement collaborations between state and local public agencies and PIOs reveals that restricting delegations to PIOs, and the limited delegation and deeply intertwined nature of collaborative enforcement, make abuse less likely than other private delegations. PIOs have less incentive to abuse a delegation than for-profit firms, particularly PIOs accountable to communities in which public agencies channel enforcement, and which have a shared interest in improving enforcement outcomes in those communities.⁹ The limited delegation of collaborative enforcement, which maintains a bright line separating public and private enforcers, affords public agencies the independence to make value-laden, contestable policy judgments about enforcement priorities without undue interference. Its intertwined nature permits public agencies to exercise meaningful oversight over private delegations to PIOs, and PIOs to act as an important counterweight to prevent capture by the regulated entities by holding public agencies politically accountable for shifts in enforcement priorities that might otherwise be difficult to detect. Collaborative enforcement can also serve the democratic value of facilitating tripartite forms of regulation by encouraging PIOs to channel worker voice into public regulation. These benefits set collaborative enforcement apart from, and suggest a lower threat to democratic government as, other forms of private delegation.

While there are many examples of informal, ad hoc collaboration between PIOs and public agencies, this Article focuses on two formal, emergent collaborative enforcement techniques. In the first, which this Article calls *remedial enforcement*, public agencies coordinate with PIOs to intensively focus their public and private enforcement tools on a particular sector in which current remedies are ineffective or insufficient. In the second collaborative enforcement approach, which this Article calls *grant-based enforcement*, state and local legislatures and agencies fund PIOs to provide capacity and expertise to assist agencies in sectoral or regional enforcement.

AMERICAN DEMOCRACY 9–13 (2009).

9 See *infra* Part IV(A). The greatest risk of abuse in privatizing public litigation comes from the complete delegation of public enforcement to private, for-profit firms. See Margaret H. Lemos, *Privatizing Public Litigation*, 104 GEO. L.J. 515, 537–56 (2016).

Examination of these techniques provides texture to empirical studies showing that the presence of PIOs in the workplace can increase “the scope and rigor of regulatory oversight.”¹⁰ In low-wage sectors, where firms are often small and undercapitalized, collaborative enforcement can more effectively and efficiently resolve legal violations and provide restitution to victims than traditional investigations or litigation. Intensively focusing enforcement in previously underregulated sectors can reveal important gaps in remedies that can be addressed through law reform. Grant-based enforcement can also improve the durability of collaborative enforcement by funding the private enforcer’s participation. Both forms of collaboration can foster tripartite bargaining between employers and PIOs to strengthen and raise labor standards through lawmaking, unionization, and codes of conduct.

A detailed examination of remedial and grant-based enforcement also reveals how public and private enforcers configure their collaborations to account for their different enforcement tools, motivations, and legal restrictions. The National Labor Relations Act¹¹ (“NLRA”), for instance, may limit union collaboration in public agency enforcement in ways that pose no obstacle to PIOs outside NLRA regulation. For PIOs that cannot lobby because of their tax-exempt status, political campaigns for stronger labor protections are out of reach without non-exempt PIOs, such as unions, playing a lead role.¹² Non-union PIOs, which are reliant on government funding because they lack the membership funding base of unions, are a more natural fit for collaboration funded through public grants.¹³ Unions,

10 Alison D. Morantz, *What Unions Do for Regulation*, 13 ANN. REV. LAW & SOC. SCI. 515, 523 (2017) (finding that studies of the regulation show more frequent and intensive inspection of unionized establishments compared with non-union establishments).

11 29 U.S.C. §§ 151–169 (2012).

12 26 U.S.C. § 501(c)(3) (2015) (stating that, while 501(c)(5) organizations may engage in a wide variety of lobbying activities, other PIOs that have 501(c)(3) tax status may risk their tax-exempt status if lobbying is a substantial part of their activities).

13 See Leslie C. Gates et al., *Who Funds Worker Centers?: An Analysis of Publicly-available Data and Directions for Future Research* (2008-2014), in NO ONE SIZE FITS ALL: WORKER ORGANIZATION, POLICY AND MOVEMENT FOR A NEW ECONOMIC AGE (LERA Edited Volume) (forthcoming 2018) (finding that over half of surveyed worker centers nationally receive funding from government); ELS DE GRAAUW, *MAKING IMMIGRANT RIGHTS REAL: NONPROFITS AND THE POLITICS OF INTEGRATION IN SAN FRANCISCO* 53 (2016) (In survey of 100 immigrant-serving PIOs in San Francisco, “sixty-

which are not restricted in their lobbying and can access funding from their membership base, but which face shifting membership demands, are better suited to the short-term and lobbying-focused collaborations of remedial enforcement.

This Article offers collaborative enforcement as a plausible pathway to counteract entrenched norms of noncompliance in low-wage sectors of the economy. Remedial and grant-based enforcement can be deployed as temporary means to create new legal remedies or elaborate a new labor protection, or in an ongoing, two-tier enforcement approach to address persistent enforcement gaps. While the mini-case studies offered here draw from better-known examples in states with a long history of labor standards enforcement, they have important implications for other jurisdictions in need of meaningful remedies and a sustained engagement between public agencies and PIOs to enforce them. Even in jurisdictions untouched by local labor lawmaking, formal coordination with PIOs to identify and assist complainants can enable state and local agencies to channel enforcement resources to sectors characterized by widespread noncompliance. The availability of meaningful remedies and the sustained, coordinated engagement of public agencies and PIOs that collaborative enforcement can offer may also set the stage for enforced self-regulation as a next-generation regulatory strategy.

Collaborative enforcement, while an important, emergent model, also poses challenges to the resulting enforcement design. The use of public resources to fund private enforcement raises legitimate concerns about private delegation undermining public agency and PIO independence.¹⁴ This may require measures to prevent PIO misuse of public funds and to ensure that PIO grants are not simply substituting for capacity and expertise that public agencies could provide alone.¹⁵ The possibility that grant-based enforcement will undermine PIO independence in monitoring public agency enforcement will require strong whistleblower protections as well. Where these risks cannot be controlled through collaborative enforcement's administrative design, it may be necessary to identify alternative means to fund PIO enforcement. While doctrinal

two reported receipt of government funding, with an average of 55 percent of their operating budget coming from government funding in 2006.”).

14 See Jon D. Michaels, *An Enduring, Evolving Separation of Powers*, 115 COLUM. L. REV. 515, 583–85 (2015).

15 See Nester M. Davidson, *Localist Administrative Law*, 126 YALE L.J. 564, 617–19 (2017).

limitations do not threaten the core requirements of collaborative enforcement,¹⁶ NLRA preemption and the nondelegation doctrine limit collaborative enforcement's legislative aims and scope of delegation.

While the focus of this Article is the civil enforcement of employment laws in the low-wage workplace, its conceptual framework has implications for the public and private enforcement of other laws that call for public and private civil enforcement, such as consumer protection, fair housing and civil rights, in which private enforcers cannot effectively vindicate private rights without access to public enforcement tools. In contrast, it has limited applicability to enforcement regimes in which there is no private right to vindicate, such as the enforcement of criminal and immigration laws. In these areas, private delegations often empower private stakeholders to make public enforcement decisions, which can create unmanageable risks of fraud and abuse.¹⁷ This Article will refer to the wholesale

16 See *infra* Part IV.

17 Collaborative enforcement has direct applicability to the underenforcement of consumer protection laws because of mandatory arbitration of consumer disputes since *AT&T Mobility LLC v. Concepcion*, 563 U.S. 333 (2011), which has displaced the consumer protection class action bar from its traditional role as private attorneys general. Myriam Gilles and Gary Friedman, *After Class: Aggregate Litigation in the Wake of AT&T Mobility v. Concepcion*, 79 U. CHI. L. REV. 623, 627 (2012). The formal coordination of PIOs and public agencies may, for example, assist in the enforcement of consumer protection statutes by identifying complainants with pattern-or-practice complaints, just as public and private enforcers have historically coordinated in the development of matched-pair testing programs to enforce fair housing and civil rights laws. See JORGES ANDRES SOTO & DEIDRE SWESNIK, THE PROMISE OF THE FAIR HOUSING ACT AND THE ROLE OF FAIR HOUSING ORGANIZATIONS 5–7 (2012) (FHA authorizes the HUD to provide grants to PIOs, which often use testing as an enforcement tool, as evidence in civil litigation or in complaints to federal enforcement agencies.). Delegation of immigration and criminal enforcement powers, in contrast, has been criticized for failing to account for ways in which private interests can undermine public enforcement goals. See, e.g., Roger A. Fairfax, Jr., *Delegation of the Criminal Prosecution Function to Private Actors*, 43 U.C. DAVIS L. REV. 411, 427–45 (2009) (arguing that privatizing criminal prosecutions undermines the public goals of public enforcement and creates conflicts of interest and risks of abuse). The delegation to private employers, for example, of the responsibility to determine their employees' authorization to work in the U.S. under immigration law, has been widely criticized for failing to deter illegal immigration or protect U.S. labor markets. See, e.g., Michael J. Wishnie, *Prohibiting the Employment of Unauthorized Immigrants: The Experiment Fails*, 2007 U. CHI. LEGAL F. 193, 195 (2007) (arguing that employer sanctions "has led to increased workplace exploitation of undocumented immigrants, strengthened the 'jobs magnet' that sanctions

delegation of public enforcement responsibilities, including those in which there is no private right to vindicate, such as deputization, to distinguish it from the private enforcer's role in collaborative enforcement to vindicate private rights.

This Article proceeds as follows. Part II will describe the enforcement gaps that have led to the underdeterrence of labor standards and explore the ways in which public agencies and PIOs have addressed these pathologies along a spectrum of public and private enforcement. Part III will explain and evaluate collaborative enforcement through examples of remedial and grant-based enforcement in different jurisdictions and offer them together as a conceptual framework for enforcing labor standards, particularly in the informal economy and to elaborate new labor standards. Part IV will examine the risk of the erosion of public and private enforcer independence in collaborative enforcement, and the doctrinal limitations under NLRA preemption and state nondelegation doctrine. The Article concludes that collaborative enforcement is a plausible pathway to improve state and local employment law compliance, increase the political accountability of public enforcement, and foster tripartite regulation of the workplace.

II. Public and Private Enforcement Pathologies

Workers in the informal economy report routine violations of basic wage and hour and health and safety laws.¹⁸ At the heart of this failure to protect vulnerable workers lies a frayed system of public and private enforcement.¹⁹ Private, for-profit attorneys have little incentive to represent plaintiffs where labor standards are

aimed to weaken, encouraged illegal immigration, and eroded wages and working conditions for U.S. workers").

18 See EASTERN RESEARCH GROUP, THE SOCIAL AND ECONOMIC EFFECTS OF WAGE VIOLATIONS: ESTIMATES FOR CALIFORNIA AND NEW YORK 2–3, 26 (2014), <https://www.dol.gov/asp/evaluation/completed-studies/wageviolationsreportdecember2014.pdf> [hereinafter EASTERN RESEARCH GROUP] (Study commissioned by USDOL estimates that over 10% of low-wage employees in California and nearly 20% of low-wage employees in New York were paid below the minimum wage in the previous month); ANNETTE BERNHARDT ET AL., BROKEN LAWS, UNPROTECTED WORKERS: VIOLATIONS OF EMPLOYMENT IN LABOR LAWS IN AMERICA'S CITIES (2009) [hereinafter BROKEN LAWS, UNPROTECTED WORKERS] (2009 survey of low-wage workers in several states finds that three-quarters of them were not paid owed overtime and reported not being paid at all for some portion of their shift.).

19 GUY DAVIDOV, A PURPOSIVE APPROACH TO LABOUR LAW 224–29 (2016).

the worst because these employers are often small and judgment proof.²⁰ Public agencies tasked with enforcing labor standards lack the resources to inspect all workplaces effectively, the expertise to detect labor and safety and health violations, and the enforcement tools to remedy them.²¹ Public enforcers have adopted different strategic enforcement approaches to prioritize the worst violations and the least compliant industries.²² Yet, and particularly with the retreat of federal agency labor standards enforcement, state and local agencies face strong headwinds in changing employer norms and ensuring compliance with new statutory protections.

PIOs such as unions, worker centers, and community-based organizations have historically played a significant role in bridging these enforcement gaps through private enforcement the informal economy and by holding public agencies accountable for public enforcement outcomes.²³ But these PIOs also face limitations in their capacity, expertise and remedies.

A. *The Enforcement Pathologies of Independent Enforcement*

Scholarship examining enforcement gaps undermining enforcement by public agencies and PIOs traditionally treats public and private enforcement as independent approaches with separate regulatory pathways. In this binary framework, public agencies are traditionally the enforcers and PIOs either privately enforce the law or hold public agencies accountable for public enforcement outcomes.²⁴ This is, of course, in part because statutes either contemplate or

20 See generally David Freeman Engstrom, *Agencies as Litigation Gatekeepers*, 123 YALE L.J. 616, 662 & n.48–49 (2013) (“High-harm misconduct may . . . attract suboptimal private enforcement efforts . . . because regulatory targets are judgment-proof . . .”).

21 David Weil & Amanda Pyles, *Why Complain? Compliance, and the Problem of Enforcement in the U.S. Workplace*, 27 COMP. LAB. L. & POL’Y J. 59, 59 (2005); see also Linda Delp & Kevin Riley, *Worker Engagement in the Health and Safety Regulatory Arena Under Changing Models of Worker Representation*, 40 LAB. STUD. J. 54, 56–57 (2015).

22 See generally DAVID WEIL, IMPROVING WORKPLACE CONDITIONS THROUGH STRATEGIC ENFORCEMENT: A REPORT TO THE WAGE AND HOUR DIVISION (2010).

23 See Scott L. Cummings, *The Internationalization of Public Interest Law*, 57 DUKE L.J. 891, 912–23 (2008) (describing community-based efforts by worker centers and social services agencies, and traditional legal services provided by public interest lawyers as two distinct types of PIOs enforcing labor standards on behalf of immigrant workers).

24 See, e.g., AYRES & BRAITHWAITE, *supra* note 5, at 56.

require independent enforcement. Many statutes, such as the Occupational Safety and Health Act of 1970 (OSHA), the NLRA, and state unemployment insurance and workers' compensation anti-fraud provisions, only call for enforcement by public agencies.²⁵ Others provide for public and private enforcement, such as the Fair Labor Standards Act of 1938 (FLSA), as alternative forms of redress.²⁶ Public agencies, PIOs and legislatures have, accordingly, traditionally treated particular public or private enforcement gaps independently rather than as broader, systemic pathologies. This view can elide ways in which enforcement pathologies reach across the public-private divide.

This Part will first show the enforcement gaps in public and private enforcement of employment laws in order to demonstrate how collaboration can bridge these gaps in the next Part. To visually demonstrate how enforcement gaps impact independent enforcement, this Part adapts the familiar enforcement pyramid first introduced by Ian Ayres and John Braithwaite in *Responsive Regulation*. In an enforcement pyramid, Ayres and Braithwaite depict an escalating series of sanctions designed to deter rational actors from violating the law and to incapacitate bad actors. The purpose of this depiction is to explain how a public agency's access to an effective enforcement pyramid can encourage the regulated entities to self-regulate.²⁷ The purpose here, however, is different. It is to show gaps in public and private enforcement pyramids that undermine the effectiveness of independent enforcement, and suggest their resolution through the integration of public and private enforcement tools.

1. *Public Agency Enforcement*

Many state and local public agencies regulating the workplace begin their enforcement with a workplace inspection or a subpoena to commence an investigation, while others primarily resolve workplace complaints through administrative hearings. Where civil inspections, investigations and adjudications are unsuccessful, public agencies may commence civil litigation. A number of state and municipal agencies also have the power to suspend or terminate business licenses for labor or safety and health-related violations.²⁸

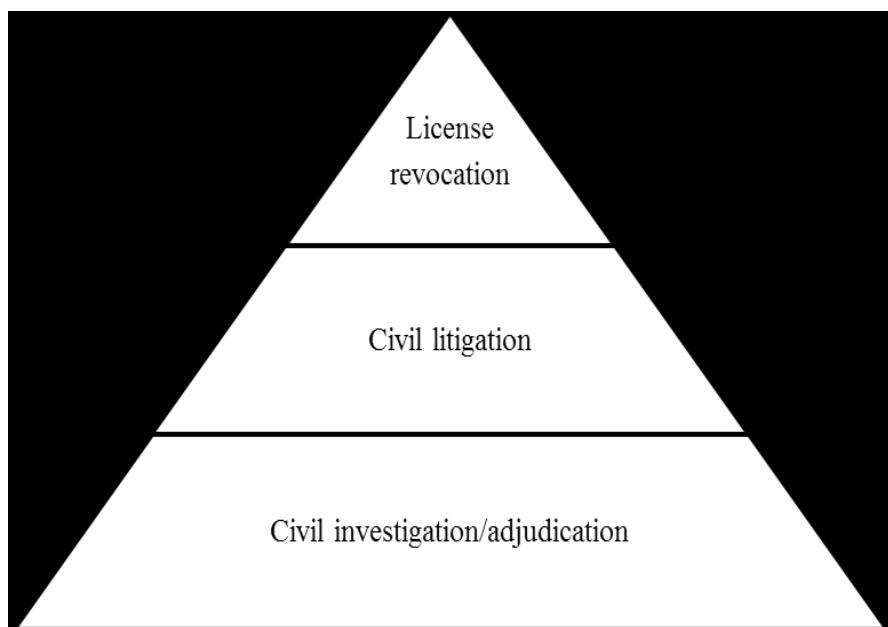
25 See 29 U.S.C. § 160(c) (2012); 29 U.S.C. § 653(b)(4) (2012)

26 See 29 U.S.C. § 216 (2012).

27 AYRES & BRAITHWAITE, *supra* note 5, at 35–36.

28 See Andrew Elmore and Muzaffar Chishti, *Strategic Leverage: Use of State and Local Laws to Enforce Labor Standards in Immigrant-Dense Occupations*, Migration Policy

Together, a simplified public agency workplace enforcement pyramid is thus:



Public Agency Enforcement

Public agencies with an enforcement pyramid with access to several remedies at different levels of severity may develop an “active escalation” regulatory approach. In this approach, most enforcement interactions begin at the base of the pyramid. Public agencies may escalate cases that employers refuse to informally resolve with intermediate sanctions.²⁹ By pushing most regulatory interventions to low and intermediate steps, public agencies reserve severe, resource-intensive interventions at the top of the pyramid as a final threat to incapacitate bad actors that will not otherwise comply with the law.³⁰

Institute 22–30 (2018), <https://www.migrationpolicy.org/research/strategic-leverage-use-state-and-local-laws-enforce-labor-standards-immigrant>.

29 AYRES & BRAITHWAITE, *supra* note 5, at 39.

30 *Id.* at 35–36. While public agencies may seek to address an inadequate sanctions regime by seeking to persuade legislators to provide additional resources and enforcement tools, they may not engage in grass-roots lobbying campaigns. *See* Anti-Lobbying Act, 18 U.S.C. § 1913 (2012) (prohibiting federal agencies from substantial participation in grass-roots lobbying campaigns or spending money on lobbying, but permitting public and private

However, the effectiveness of the active escalation approach is contingent on the public agency's ability to detect violations.³¹ First, in most instances this requires access to worker complaints.³² But workers do not complain, especially in workplaces where conditions are the worst.³³ Many employees are unaware that they are paid unlawfully, possibly because they are misclassified as independent contractors, or because of language differences, low literacy levels, or industry norms that—while out of sync with legal obligations—set low worker expectations of permissible workplace standards.³⁴ Even for workers who understand their legal rights, few know where to file a complaint with a public agency.³⁵

communications to promote agency positions). State and local public agencies and officials are subject to different lobbying restrictions. *Compare CAL. GOV'T CODE §§ 50023, 53060.6, 86300* (prohibiting state and local officials from providing gifts of value to elected officials but permitting state and local agencies and officers to engage in public and private communications to support their agency positions and permitting local agencies and officers to join associations for lobbying purposes) *with WASH. REV. CODE § 42.17A.635* (prohibiting expenditure of taxpayer funds or facilities to lobby, and limiting communication to legislature to communications invited by legislators and requests for legislative action “through the proper official channels”).

31 See ROBERT BALDWIN, MARTIN CAVE & MARTIN LODGE: UNDERSTANDING REGULATION: THEORY, STRATEGY, AND PRACTICE 272–74 (Oxford Univ. Press, 2n ed. 2012).

32 Fine & Gordon, *supra* note 4, at 556.

33 Weil & Pyles, *supra* note 21, at 90–92; WEIL, *supra* note 1, at 245–48 (Wage and hour complaints to USDOL have declined 26% over the past decade, even as violations have increased during the same period.); Charlotte S. Alexander & Arthi Prasad, *Bottom-Up Workplace Law Enforcement: An Empirical Analysis*, 89 IND. L.J. 1069, 1089 (2014) (Almost all employees who do complain about rights violations complain directly to their employer, and almost none report illegal conduct to public agencies.); Rebecca Smith, *Immigrant Workers and Workers' Compensation: The Need for Reform*, 55 AM. J. INDUS. MED. 537, 537–44 (2012) (Few injured immigrant workers apply for workers' compensation benefits.); OSHA, ADDING INEQUALITY TO INJURY 6–7 (2015), <https://www.dol.gov/osta/report/20150304-inequality.pdf>, (summarizing studies of injured workers, which “found that fewer than 40 percent of eligible workers apply for any workers' compensation benefits at all”).

34 WEIL, *supra* note 1, at 245; Alexander & Prasad, *supra* note 33, at 1085, 1097 (analyzing the results of BROKEN LAWS, UNPROTECTED WORKERS, *supra* note 18, finding that “about 59% of workers did not know their minimum wage and overtime rights”); Ruckelshaus, *supra* note 4, at 380–81. Awareness may be particularly low of subnational protections that do not extend to the federal level.

35 Alexander & Prasad, *supra* note 33, at 1095 (Of surveyed workers, “77% did not know where to file a workplace complaint with the government.”); DE GRAAUW, *supra* note 13, at 8–9.

Public agency efforts to encourage complaints through education is hindered by the justifiable complainant fear that a complaint will result in retaliation, particularly in immigrant-dense industries, where workers may fear that a complaint to a public agency will lead to arrest and deportation.³⁶

Second, even with complainants, public agencies often lack the enforcement capacity to investigate, resolve and monitor cases.³⁷ Capacity deficits are acute and unevenly distributed at the state and local level. Most states have fewer than ten investigators to enforce a wide range of workplace laws,³⁸ and five states do not engage in any labor or safety and health regulation at all.³⁹ Local governments, the site of recent historic increases in regional and sectoral minimum wages and other labor protections, often have no workforce agency to enforce them.⁴⁰

-
- 36 Alexander & Prasad, *supra* note 33 at 1091–92 (explaining that nearly half of workers who reported a complaint to their employer in the previous year experienced some form of retaliation, and only 15% of “employers addressed or promised to address the workers’ claims”); Jayesh M. Rathod, *Beyond the “Chilling Effect”: Immigrant Worker Behavior and the Regulation of Occupational Safety & Health*, 14 EMP. RTS. & EMP. POL’Y J. 267, 280–91 (2010) (citing economic insecurity, language differences, low literacy levels, social and cultural expectations of permissible workplace conditions, lack of experience, as well as age and gender as possible factors driving low OSHA complaint rates in dangerous industries); *see also* Michael J. Wishnie, *Immigrants and the Right to Petition*, 78 N.Y.U. L. REV. 667, 673–80 (2003).
- 37 Despite a rising total number of complaints, staffing at the USDOL and OSHA has declined since the 1980’s in both absolute terms and relative to the rising numbers of low-wage workers in the U.S. Devah Pager, Bruce Western & David Pedulla, *Employment Discrimination and the Changing Landscape of Low-Wage Labor Markets*, 2009 U. CHI. LEGAL F. 317, 325–26 (USDOL staffing decreased between 1975 and 2004, and OSHA staffing and annual federal and state OSHA inspections declined by half between 1980 and 2006.).
- 38 ZACH SCHILLER & SARAH DECARLO, *INVESTIGATING WAGE THEFT: A SURVEY OF THE STATES* 6 (2010), <http://www.fairwarning.org/wp-content/uploads/2014/04/Link23.pdf>.
- 39 JACOB MEYER & ROBERT GREENLEAF, NAT’L STATE ATTORNEYS GEN. PROGRAM AT COLUMBIA LAW SCH., *ENFORCEMENT OF STATE WAGE AND HOUR LAWS: A SURVEY OF STATE REGULATORS* 13 (2011) (Alabama, Florida, Georgia, Louisiana and Mississippi).
- 40 For example, most of the thirteen municipalities in New Jersey that recently passed mandatory sick pay laws have vested enforcement powers in agencies with no expertise in workplace enforcement. *See, e.g.*, BLOOMFIELD, N.J., CODE § 160–2(1) (2015) (delegating enforcement power to local department of health) EAST ORANGE, N.J., CODE § 140–10, § (2016) (same); IRVINGTON, N.J., CODE MC 3513, § 1(1) (2014) (vesting enforcement powers in Department of Neighborhood Services).

Third, an escalation strategy depends on the availability of intermediate steps to secure the immediate compliance of reasonable employers.⁴¹ But many agencies possess only a single remedy and are limited to threatening or using its sole sanction.⁴² Public agencies also often lack meaningful top-pyramid sanctions.⁴³ Some violations of workplace laws carry penalties that are lower than the cost of compliance.⁴⁴ And even with access to a single, top-pyramid deterrence option, such as license revocation, agencies cannot effectively regulate most employers without access to intermediate steps because “it is politically impossible and morally unacceptable to use it with any but the most extraordinary offenses.”⁴⁵

2. Private Enforcement

PIOs often seek to resolve the same types of workplace problems as public agencies. Some PIOs, such as worker centers and community organizations, can informally resolve myriad workplace cases, often on the spot, with a call to the employer. Those that PIOs cannot informally resolve may be referred to public agencies for enforcement that provide case-related assistance, while some PIOs may engage in civil litigation.⁴⁶ Some PIOs may organize a protest

41 See WEIL, *supra* note 1, at 237 (finding that complaint-driven inspections take up more than 70% of USDOL’s wage-and-hour law investigations).

42 This describes USDOL’s Wage and Hour Division, which may only either negotiate an investigative finding or refer the matter to USDOL’s Solicitor’s Office to litigate. See Elmore and Chishti, *supra* note 28, at 17.

43 USDOL recovery for failure to pay minimum wages and overtime is often limited to one to two times the owed wages. 29 U.S.C. § 216(b) (2012).

44 Despite recent increases in maximum penalties, OSHA often imposes penalty amounts less than the cost of compliance that are unlikely to deter violations in industries where complaints are rare. Weil & Pyles, *supra* note 21, at 62; 29 U.S.C. § 666(b) (2012). Employment law violations, even if serious and willful, often only carry a low-level criminal penalty unlikely to garner the attention of criminal prosecutors. See 29 U.S.C. § 666(e) (2012) (Willful OSHA violations is a misdemeanor only if OSHA violation causes the death of an employee.); 29 U.S.C. § 216(a) (2012) (Willful FLSA violation is a misdemeanor, and cannot result in imprisonment until after first offense.).

45 AYRES & BRAITHWAITE, *supra* note 5, at 36. This may help explain why, despite the recent proliferation of “wage theft” laws in municipalities across the country, they have resulted in few criminal prosecutions. MEYER & GREENLEAF, *supra* note 39, at 16.

46 JANICE FINE, WORKER CENTERS: ORGANIZING COMMUNITIES AT THE EDGE OF THE DREAM 72–91, 73, 78–83 (2006) (“The most common complaint [received by worker centers] by far is unpaid wages, which includes paying below the minimum wage and nonpayment or underpayment of overtime.”).

to resolve a specific case, while unions often have the ultimate end of unionizing the workplace. Unions have also traditionally⁴⁷—and increasingly⁴⁸—lobbied state and local government for higher and more enforceable workplace standards,⁴⁹ a shift from firm-based bargaining to what Kate Andrias calls “social bargaining”—i.e., bargaining that occurs in the public arena on a sectoral and regional basis.”⁵⁰

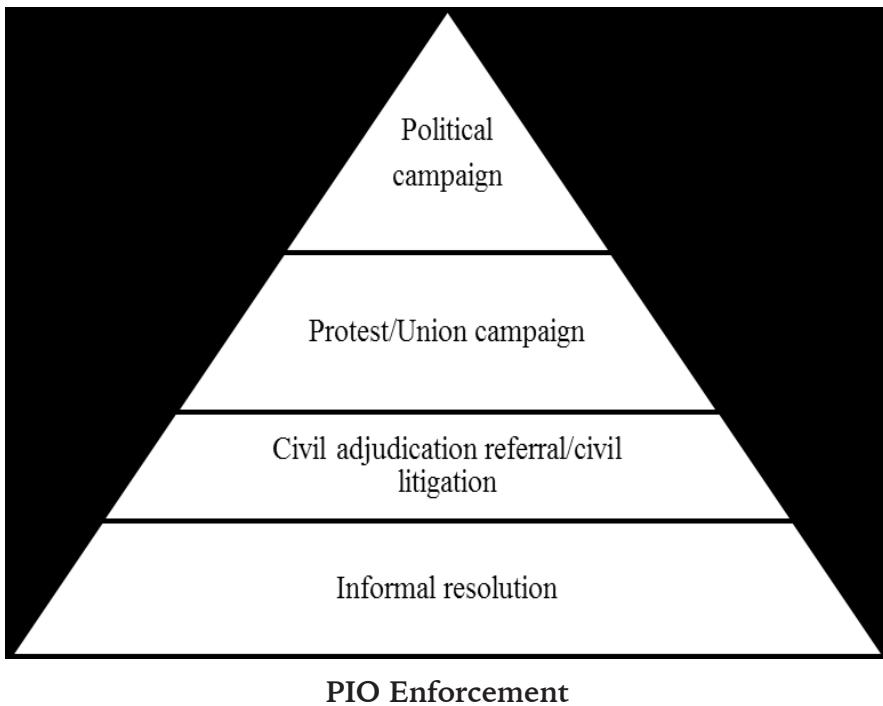
47 Unions have traditionally influenced politics through contributions and membership mobilization. *See* RICHARD B. FREEMAN & JAMES L. MEDOFF, WHAT DO UNIONS Do? 191–206 (1984); JAKE ROSENFIELD, WHAT UNIONS No LONGER Do 1–2, 31–34 (2014) (describing, for example, Service Employees International Union’s successful lobbying efforts in 1990s and 2000s for states to create public agencies to assume the role of employer to bargain collectively with over 100,000 home-based child care workers in Illinois, California and other states).

48 *See generally* DAVID ROLF, THE FIGHT FOR \$15: THE RIGHT WAGE FOR A WORKING AMERICA 122–64 (2016) (describing union-led lobbying in Fight for \$15 movement and its legislative victories in Washington State, including \$15 minimum wage in Seattle).

49 Unions are incorporated under a provision of the tax code that has no lobbying restrictions. 26 U.S.C. § 501(c)(5) (2015). In contrast, and as with public agencies and officials, most non-union PIOs may not engage in electioneering because of their exempt tax status. 26 U.S.C. § 501(c)(3). This restriction provides ample opportunity for PIOs to engage in administrative advocacy and other forms of lobbying, either by ensuring that their lobbying activities are not “substantial,” or by separately funding lobbying activities through an alter ego 501(c)(4) social welfare organization. *See id.*; *Regan v. Taxation with Representation*, 461 U.S. 540, 551 (1983) (upholding 501(c)(3) lobbying restrictions against first amendment challenge on ground that PIOs can exercise free speech through separately incorporated 501(c)(4)).

50 Andrias, *supra* note 6, at 8.

Altogether, the PIO enforcement pyramid is:



As with public enforcement, the private enforcement pyramid reveals significant enforcement gaps. PIOs are also unable to access complainants at times. In many low-wage sectors, unions, which historically have assisted workers in understanding the laws and how to enforce them, are virtually absent.⁵¹ In their wake, non-union PIOs such as worker centers have increasingly sought firm-level compliance in the low-wage workplace.⁵² But, lacking significant

51 WEIL, *supra* note 1, at 41, 245–46, 254 (Union density was about 7% of total employment when NLRA passed, rose to about 35% in 1954, and has fallen once again to 11.3% by 2012, and only 6.6% of the private sector workforce. Meanwhile, in many sectors in the informal economy, such as commercial cleaning and restaurants, union density is effectively zero.).

52 ESTLUND, *supra* note 7, at 17. These PIOs often have regular contact with the low-wage, disproportionately foreign-born workers they serve, and as a result have specialized expertise about the particular labor and safety problems in the industry or worker community that the PIO serves. FINE, *supra* note 46, at 73 (“Most workers first come into contact with a [worker] center because they are seeking help with an employment-related problem.”); JENNIFER

membership dues or legal fees, these PIOs often depend at least in part on unstable philanthropic funding sources that are cyclical and subject to shifting priorities.⁵³ Those PIOs that only informally resolve or refer cases face the same enforcement problem as single-remedy public agencies.

As with the private bar, civil litigation barriers also undermine PIOs' access to intermediate and top-pyramid enforcement remedies. The ability of PIOs to represent workers in private litigation is limited by its considerable cost,⁵⁴ and the inability of PIOs that receive federal funding for pro bono legal services to represent clients who lack authorization to work.⁵⁵ Even assuming PIO capacity, there is no private right of action to enforce NLRA or OSHA,⁵⁶ and many back wage claims are too small to justify the hassle and expense of a FLSA lawsuit.⁵⁷ In contrast to the ease with which a small employer can evade legal judgments for labor standards violations, unions seeking to protest the same practices could run afoul of the NLRA's labor picketing restrictions.⁵⁸ For these reasons, even

GORDON, *SUBURBAN SWEATSHOPS: THE FIGHT FOR IMMIGRANT RIGHTS* 67–111, 185–218, 299 (2005) (describing worker center's legal clinic and organizing work around immigrant worker labor standards complaints, and legal and policy organizations that assist worker centers).

- 53 See Gates et al., *supra* note 13, at 8–9, 17–18 (finding that worker center funding is unstable, and that over 20% of funding comes from foundations).
- 54 Fee-shifting statutes, which generally require a collectable judgment, is just as uncertain a funding stream as philanthropy, contingent on becoming a prevailing party in a suit against an employer with an ability to pay. DEBORAH L. RHODE, *ACCESS TO JUSTICE* 115 (2004).
- 55 *Id.* at 116; 45 C.F.R. § 1626.3 (2014) (ordering that, except in limited circumstances, Legal Services Corporation “recipients may not provide legal assistance for or on behalf of an ineligible alien”); *see* Cummings, *supra* note 23, at 914 (describing impact of LSC restriction on immigrant workers’ rights infrastructure).
- 56 See 29 U.S.C. § 653(b)(4) (2012) (stating that there is no private right of action under OSHA); Cynthia L. Estlund, *The Ossification of American Labor Law*, 102 COLUM. L. REV. 1527, 1551–58 (2002) (arguing in favor of a private right of action to enforce NLRA).
- 57 Although the FLSA has an attorney’s fee-shifting provision under § 216, the private bar may be justifiably wary to take small FLSA cases in jurisdictions where settlement offers may be conditioned on fee waivers, or judgment collection is uncertain.
- 58 29 U.S.C. § 158(b)(7)(C) (2012). *But see* Catherine Fisk & Jessica Rutter, *Labor Protest Under the New First Amendment*, 36 BERKELEY J. EMP. & LAB. L. 277, 293–321 (2015) (arguing that an interpretation of the “blackmail picketing” statute to prohibit peaceful picketing of employers because of substandard labor conditions violates the First Amendment).

PIOs with the resources to litigate must triage carefully to conserve resources, and often lack intermediate steps for cases that do not merit resource-intensive interventions. PIOS, moreover, lack a top-pyramid enforcement tool to incapacitate bad actors that can evade a legal judgment.

At the center of the remedies gap for public and private enforcement in the low-wage workplace is the lack of meaningful remedies for retaliation. The NLRB may only remedy retaliation for concerted activities such as complaints about labor standards with the “paltry financial threat” of placing the aggrieved employee in the ex-ante position, but only after a long administrative hearing, and without any penalty or other compensatory damages for chilling speech.⁵⁹ The Supreme Court restricted even that remedy in *Hoffman Compounds, Inc. v. NLRB*⁶⁰ for the eight million workers in the U.S. workforce who lack work authorization under immigration law, most of whom work in low-wage workplaces.⁶¹ Although unauthorized workers are employees under the NLRA,⁶² they cannot receive reinstatement or post-termination backpay as a remedy for an NLRA violation.⁶³ For workplaces with high concentrations of unauthorized workers, the most likely remedy for an NLRA violation is a nearly costless cease and desist order and posting.⁶⁴ While courts generally permit recovery of owed wages notwithstanding immigration status,⁶⁵ after *Hoffman* the availability of backpay after the employer’s discovery of unauthorized status for violation of other anti-retaliation laws remains an open question as well.⁶⁶

59 Estlund, *supra* note 56, at 1553–54. See 29 U.S.C. § 160(c) (2012).

60 *Hoffman Plastic Compounds, Inc. v. NLRB*, 535 U.S. 137 (2002).

61 Jeffrey S. Passel, D’Vera Cohn & Molly Rohal, *Share of Unauthorized Immigrant Workers in Production, Construction Jobs Falls Since 2007*, PEW RESEARCH CTR., (Mar. 26, 2015) <http://www.pewhispanic.org/2015/03/26/share-of-unauthorized-immigrant-workers-in-production-construction-jobs-falls-since-2007/>.

62 *Agri Processor Co. v. NLRB*, 514 F.3d 1, 4–5 (D.C. Cir. 2008).

63 *Hoffman*, 535 U.S. at 151–52.

64 *Id.*

65 See *Lucas v. Jerusalem Café*, 721 F.3d 927, 930 (8th Cir. 2013) (holding unauthorized immigrant may recover for employer’s failure to pay minimum wage and overtime pay); *Patel v. Quality Inn So.*, 846 F.2d 700, 704 (11th Cir. 1988) (“FLSA’s coverage of undocumented aliens goes hand in hand with the policies behind the IRCA. . . . If the FLSA did not cover undocumented aliens, employers would have an incentive to hire them.”).

66 Compare *Ambrosi v. 1085 Park Ave. LLC*, No. 06-CV-8163(BSJ), 2008 WL 4386751, at *12 (S.D.N.Y. Sept. 25, 2008) (relying on *Hoffman* in declining to award future lost wages to an unauthorized immigrant in a personal injury

Gaps in the capacity, expertise and remedies of public and private enforcers have undermined the effectiveness of their enforcement tools. The resulting underenforcement of the law has degraded labor and safety and health standards in many low-wage workplaces.⁶⁷ To address these enforcement pathologies, PIOs have lobbied for, legislatures have enacted, and public agencies have implemented reforms to expand private and public enforcement. But, as the next section will show, these reform efforts have been hampered by their exclusive focus on public and private enforcement as independent regulatory pathways.

B. The Limitations of Independently Addressing PIO and Public Agency Enforcement Pathologies

Public agencies have addressed enforcement pathologies by lobbying for increased public enforcement resources, building internal expertise to investigate the informal economy, and expanding public and private enforcement remedies. State and local agencies have developed internal expertise through agency units to focus on industries where noncompliance rates are high and to elaborate new statutory requirements, and have conditioned business licenses on compliance with labor and safety and health standards.⁶⁸ These measures have improved public enforcement

suit); *Wielgus v. Ryobi Technologies*, 875 F. Supp. 2d 854, 862–63 (N.D. Ill. 2007) (holding unauthorized immigrant bringing workers compensation suit could not recover for lost U.S. wages, but could seek recovery for lost wages in his home country) *with Bollinger Shipyards v. Office of Workers Compensation Programs*, 604 F.3d 864, 877 (5th Cir. 2010) (finding unauthorized workers eligible for workers compensation benefits under Longshore and Harbor Workers Compensation Act); *Salas v. Sierra Chemical Co.*, 327 P.3d 797, 809 (Cal. 2014) (finding that immigration law does not preempt state law that permits post-termination backpay up until employer's knowledge of employee's unauthorized status).

67 See Zatz, *supra* note 4, at 42–47; see EASTERN RESEARCH GROUP, *supra* note 18, at 48 (Over 110,000 people in New York State and California fell beneath the poverty level because of minimum wage violations.); Occupational Safety & Health Admin., *supra* note 31, at 6 (50% of the cost of workplace injuries are paid out of pocket by workers and their families.).

68 See Elmore & Chishti, *supra* note 28, at 24–28; JULIE A. SU, REPORT ON THE STATE OF THE DIVISION OF LABOR STANDARDS ENFORCEMENT 1–3 (2013), http://www.dir.ca.gov/dlse/Publications/DLSE_Report2013.pdf (reporting quintupling of assessment of minimum and overtime wages, 80% increase in penalties, creation of criminal enforcement unit to file felony charges for egregious labor complaints, and reducing time required for wage hearings).

outcomes, but violation rates have remained stubbornly high, which former USDOL Wage and Hour Administrator David Weil attributes to a lack of channeling of public enforcement capacity to where it is most needed.⁶⁹

PIOs have also sought to address enforcement pathologies by lobbying to expand public and private remedies. One PIO successfully lobbied in the 1990s for increased civil and criminal penalties that the New York State Department of Labor (NYDOL) could impose,⁷⁰ and in California, PIOS successfully lobbied in 2003 to require employers in the car wash industry to obtain a license and post a bond to operate in the state.⁷¹ The effectiveness of public enforcement reforms, however, is contingent on the public agency's ability to identify complainants and to resist capture by the regulated entities. PIOS have addressed this latter shortcoming by acting as watchdogs for public enforcement, making public agencies politically accountable for their enforcement outcomes.⁷²

PIOs have also expanded their lobbying efforts to focus on expanding private enforcement. For example, in 2004, California amended its labor code to provide a Private Attorney General Act,⁷³ and in 2011, PIOS in New York successfully lobbied for the

69 WEIL, *supra* note 1, at 222–56 (emphasizing the need for public agencies to focus on the higher level of industry structures, bad actors who repeatedly and willfully violate the law, and to encourage complaints).

70 See JENNIFER GORDON, CARNEGIE ENDOWMENT FOR INT'L PEACE, THE CAMPAIGN FOR THE UNPAID WAGES PROHIBITION ACT: LATINO IMMIGRANTS CHANGE NEW YORK WAGE LAW 39 n.8 (Sept. 1999), http://carnegieendowment.org/files/imp_wp4gordon.pdf [hereinafter GORDON, UWPA].

71 Cal. Car Wash Worker Law, CAL. LAB. CODE §§ 2050–118 (West 2016).

72 See GORDON, UWPA, *supra* note 70; KEVIN BARRY, MARCY KOUKHBAB & CHLOE OSMER, REGULATING THE CAR WASH INDUSTRY: AN ANALYSIS OF CALIFORNIA'S CAR WASH WORKER LAW (Apr. 2009); CAMPAIGN TO END WAGE THEFT, PROTECTING NEW YORK'S WORKERS: HOW THE STATE DEPARTMENT OF LABOR CAN IMPROVE WAGE-AND-HOUR ENFORCEMENT 16 (Dec. 2006), <http://www.mfy.org/wp-content/uploads/reports/Protecting-Workers-Dept-of-Labor.pdf> (recommending measures to improve NYSDOL enforcement); MARCI SEVILLE, REINFORCING THE SEAMS: GUARANTEEING THE PROMISE OF CALIFORNIA'S LANDMARK ANTI-SWEATSHOP LAW — AN EVALUATION OF ASSEMBLY BILL 633 SIX YEARS LATER, WOMEN'S EMPLOYMENT RIGHTS CLINIC (2005), <http://digitalcommons.law.ggu.edu/cgi/viewcontent.cgi?article=1003&context=werc> (evaluating California Department of Labor Standards Enforcement implementation of law to hold manufacturers liable for violations of labor standards by garment contractors).

73 CAL. LAB. CODE §§ 2698, et seq. (2004).

Wage Theft Prevention Act to increase the civil damages for wage and hour law violations and retaliation that plaintiffs could receive under state law.⁷⁴ Local government private enforcement reforms have also channeled private enforcement to small, pro se claims. In Florida, several counties and one municipality have passed “wage theft” laws to provide low-wage workers with a pro se forum to recover small amounts of owed wages,⁷⁵ and some municipalities have experimented with government-funded, day-laborer hiring hall sites to regulate contingent work in residential construction sites.⁷⁶

There is evidence that expanding private capacity and remedies can improve compliance with labor and safety and health laws.⁷⁷ But law scholars have criticized privatizing public enforcement for failing to consider how to guide private enforcement to improve compliance.⁷⁸ For Margaret Lemos, this concern is most problematic when public agencies delegate a public litigation role to private attorneys.⁷⁹ David Freeman Engstrom’s empirical analysis of private litigants who enforce public law in *qui tam* whistleblower cases, for example, while finding no support for the claim that delegating public enforcement leads to outright abuse, did find that private attorneys have used the False Claims Act to exploit regulatory ambiguities rather than to reveal enforcement gaps.⁸⁰ Particularly where private attorney decisions to file wage and hour lawsuits are motivated in part by the ability to enforce judgments that generate attorney’s fees, private remedies are most likely to be used against larger,

74 N.Y. LAB. LAW §§ 198, 215 (McKinney 2011).

75 These laws target small claims with a small minimum claim amount and individual, pro se hearings, and treble damages to encourage private settlement. Tracey McManus, *Pinellas County Cracks Down on Wage Theft*, TAMPA BAY TIMES (Nov. 10, 2015), <http://www.tampabay.com/news/politics/local/pinellas-county-cracks-down-on-wage-theft/2253390>.

76 See VALENZUELA ET AL., ON THE CORNER: DAY LABOR IN THE UNITED STATES II 6 (2006), <http://www.coshnetwork.org/sites/default/files/Day%20Labor%20study%202006.pdf> (identifying “63 day-labor worker centers created by community organizations, municipal governments, faith-based organizations and other local stakeholders”).

77 Government-funded hiring halls, for example, reduce the frequency of labor standards violations and workplace injuries experienced by day laborers. Abel Valenzuela, *A Better Day Day-Labor Market*, L.A. TIMES (Aug. 13, 2008), <http://www.latimes.com/opinion/la-oew-valenzuela13-2008aug13-story.html>.

78 Lemos, *supra* note 9, at 529, 578–82.

79 *Id.* For a discussion of this critique of private enforcement of public law, see David Freeman Engstrom, *Private Enforcement’s Pathways: Lessons from Qui Tam Litigation*, 114 COLUM. L. REV. 1913, 1925–30 (2014).

80 Engstrom, *supra* note 79, at 1964, 2000–01.

solvent employers. As a result, while there has been a substantial uptick for FLSA suits since 2000,⁸¹ it is not clear that these efforts have substantially improved enforcement outcomes either to fill gaps in public enforcement or to address the worst offenses. Shifting private remedies to pro se fora also may not increase complaints from the informal economy without PIO assistance. It seems equally plausible that the workers most likely to take advantage of private enforcement without PIO assistance are those who are most likely to complain to a government agency.

In short, independent public and PIO enforcement reform efforts have shown that addressing noncompliance with labor standards must attend to enforcement gaps as pathologies that extend across the public-private divide. Public and private enforcement require complaints to activate their enforcement pyramids, remedies for complainants and sanctions to deter rational actors and incapacitate bad actors. Public enforcement alone is unlikely to improve compliance without the political accountability that PIOs provide as public enforcement watchdogs, while the effectiveness of private enforcement is contingent on whether it can be channeled to sectors in which it is most needed. The next section will explore how early examples of public-private regulatory experimentation have sought to amplify the deterrent effect of enforcement, with mixed results.

C. Tripartism to Deter Rather Than to Encourage Self-Regulation

The previous section has shown how public and private enforcement pathologies can drive underdeterrence in low-wage workplaces. This section will show how PIOs and state and local agencies have increasingly sought to collaborate in order to address enforcement pathologies that limit the effectiveness of independent public and private enforcement. It will conclude that, unlike previous public-private regulatory experimentation that has sought to shift responsibility for legal compliance to the regulated entities, emergent public-private collaboration in the low-wage workplace seeks to improve compliance by amplifying enforcement's deterrent effect.

⁸¹ Lisa Nagele-Piazza, *Uptick in FLSA Litigation Expected to Continue in 2016*, BLOOMBERG BNA (Nov. 27, 2015), <https://www.bna.com/uptick-flsa-litigation-n57982064020/> (noting that FLSA filings have increased 450% from 2000 to 2015).

Scholars of public regulation in law and related disciplines often examine public-private enforcement experimentation through the lens of “tripartism.” Tripartism describes regulation that creates incentives, often in the form of flexible standards and addressing information asymmetries, to facilitate private bargaining between regulated entities and PIOS for firm-specific compliance.⁸² Tripartism can encourage self-regulation by the regulated entities through public enforcement and monitoring by PIOS.⁸³ However, self-regulation, even when backed by public and private enforcement, can have limited potential in low-wage workplaces in which employers are undercapitalized, competition is fierce and labor is a substantial proportion of the business cost. In many low-wage sectors, firms have no personnel department, and owners are often directly responsible for employment law violations. In these sectors, there is a powerful norm of noncompliance with basic employment laws.⁸⁴

Facing entrenched norms of noncompliance and a lack of employer expertise in compliance and motivation to comply, public-private regulatory experimentation to shift employment law compliance to employers through tripartism can have limited impact. OSHA’s Voluntary Protection Program (VPP) is a case in point. Under VPP, participating employers may opt out of OSHA inspections in return for their adoption of a safety program, executed by trained employees with an established protocol to notify managers of and respond to hazards.⁸⁵ Early evidence suggested that employers that

82 Guided by New Governance scholarship, tripartism can describe any regulatory approach that seeks to enlist private participation in regulatory experimentation through improved governance. See Michael Waterstone, *A New Vision of Public Enforcement*, 92 MINN. L. REV. 434, 479–97 (2007); Orly Lobel, *The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought*, 89 MINN. L. REV. 342, 372 (2004). For a description of New Governance literature, and a critical evaluation of its application to workplace regulation, see ESTLUND, *supra* note 7, at 136–39.

83 AYRES & BRAITHWAITE, *supra* note 5, at 57–58.

84 In the car wash sector, for example, a 2008 state investigation found that nearly 80% of New York City car wash establishments violated wage-and-hour law, Steven Greenhouse, *Carwashes Violating Wage Laws, State Finds*, N.Y. TIMES (Aug. 15, 2008), and a California state investigator estimated that over half of car washes in Southern California failed to pay the minimum wage during the same period. Sonia Nazario & Doug Smith, *Inspectors Find Dirt on Books at Southern Calif. Carwashes*, L.A. TIMES (Mar. 23, 2008).

85 Orly Lobel, *Interlocking Regulatory and Industrial Relations: The Governance of Workplace Safety*, 57 ADMIN. L. REV. 1071, 1104–06, 1133 (2005) (“Since 2000, these programs have expanded to thousands of worksites, and the agency has increased the proportion of its resources dedicated to these

participated in VPP improved their safety and health standards.⁸⁶ But more recently, VPP-participating employers have been found to have ignored safety and health standards, which government investigations attribute to OSHA's failure to provide external accountability.⁸⁷ The VPP's devolution from a paradigmatic form of tripartite regulation to a form of deregulation calls into question the usefulness of tripartism in sectors characterized by widespread noncompliance. At a minimum, it suggests that tripartism requires strong, durable roles for public agencies and PIOs to make self-regulation enforceable and accountable.⁸⁸

Law scholars have theorized that tripartism can improve enforced self-regulation by delegating public enforcement responsibilities to PIOs.⁸⁹ But even tripartism backed by delegated private enforcement can be undermined by the divergent interests of public agencies and PIOs in enforcement and the possibility of backlash from the regulated entities.⁹⁰ As Estlund observes, as the

activities to almost 30 percent of its entire budget.”).

86 *Id.*

87 Susan Bisom-Rapp, *Puzzling Evidence from a Troubled Time: Rethinking State Promotion of Safe Work During the Bush Administration*, 14 EMP. RTS. & EMP. POL'Y J. 295, 308–10 (2010) (citing U.S. GOV'T ACCOUNTABILITY OFFICE, REPORT TO CONGRESSIONAL REQUESTERS: OSHA's VOLUNTARY PROTECTION PROGRAMS: IMPROVED OVERSIGHT AND CONTROLS WOULD BETTER ENSURE PROGRAM QUALITY 7–10 (2009)). A 2011 internal OSHA report recommended restructuring the VPP to maintain oversight over VPP participants, but by 2013 a USDOL Office of Inspector General report found that VPP controls were still insufficient to ensure that only employers with adequate safety and health systems participated. Nancy Smith et al., *Voluntary Protection Programs Review* (2011), https://www.osha.gov/dcsp/vpp/vpp_report_nov_2011_rev_7-11-12.pdf; USDOL, *Voluntary Protection Program: Controls are not Sufficient to Ensure Only Worksites with Exemplary Safety and Health Systems Remain in the Program*, 4 (2013), <http://www.oig.dol.gov/public/reports/oa/2014/02-14-201-10-105.pdf>.

88 See Cynthia Estlund, *Employment Rights and Workplace Conflict: A Governance Perspective*, in THE OXFORD HANDBOOK OF CONFLICT MANAGEMENT IN ORGANIZATIONS 54, 64 (2014) (arguing that voluntary self-regulation without accountability can “become a thinly disguised form of deregulation”); Fine & Gordon, *supra* note 4, at 558 (Unchecked self-regulation could “contribute to the further degeneration of standards in low-wage sectors.”); Leah F. Vosko, John Grundy & Mark P. Thomas, *Challenging New Governance: Evaluating New Approaches to Employment Standards Enforcement in Common Law Jurisdictions*, 26 ECON. & IND. DEMOCRACY 373, 391–93 (2014). Lobel also warns that voluntary self-regulation is vulnerable to deregulation. See Lobel, *supra* note 85, at 1111–14.

89 ESTLUND, *supra* note 7, at 23.

90 Cynthia Estlund identifies as an example of PIO-monitored tripartism a

interests of public agencies and PIOs diverge, tripartism backed by public enforcement faces the risk of PIO and public agency exit from the collaborative relationship.⁹¹ While enforced self-regulation backed by public agency and PIO enforcement is a promising example of tripartism to regulate the informal economy, it suggests the need to build durable collaborative relationships beyond a single case outcome given unstable PIO funding sources and the shifting priorities of PIOs and public agencies.

To address this durability question, political scientist Janice Fine and law scholar Jennifer Gordon have theorized a form of tripartism in which PIOs and public agencies extensively collaborate beyond the outcome of a particular case.⁹² But formal public-private collaborations to amplify the deterrent effect of enforcement has not yet accounted for the potential risks of abuse and cooptation that collaboration may invite, and doctrinal limitations to the scope of its delegation and its legislative aims.⁹³ Avoiding the employer

code of conduct developed by the New York Office of Attorney General (“NYAG”) to resolve a labor dispute between a union and immigrant service organization seeking to organize employees of small groceries, or “greengrocers,” and thousands of greengrocer owners in New York City. *Id.* at 112–16. After a union organizing drive among greengrocery workers, which identified persistent labor violations, the union and a greengrocer association approached the NYAG, which negotiated the Greengrocer Code of Conduct. *Id.* See also Matthew T. Bodie, *The Potential for State Labor Law: The New York Greengrocer Code of Conduct*, 21 HOFSTRA LAB. & EMP. L.J. 183, 194 (2003). The code of conduct required the owners to set workplace conditions at or above the legal requirements and agree to monitoring by PIOs, in return for NYAG’s using its prosecutorial discretion not to target them for enforcement. ESTLUND, *supra* note 7, at 112–113.

- 91 In the greengrocer code of conduct example, after the unionization campaign fizzled, the immigrant service organization and public agency lacked the resources to sustain firm-level monitoring of code compliance. With the PIOs’ withdrawal from the industry, and without meaningful monitoring, the employer had little incentive to extend it. ESTLUND, *supra* note 7, at 114.
- 92 Fine & Gordon, *supra* note 4, at 560. Fine and Gordon point to the New York State Department of Labor’s (NYSDOL) formation in 2009 of the New York Wage and Hour Watch (W&HW), modeled on neighborhood watch associations, in which NYSDOL designated specific PIOs, including unions, the task of educating workers in specific neighborhoods and referring complaints to NYSDOL. *Id.* at 569–70.
- 93 Unfortunately, the W&HW was short-lived. By 2010, NYDOL disbanded the W&HW to avoid politicizing the nomination of then-NYDOL Commissioner Patricia Smith as Solicitor of Labor for USDOL. Janice Fine, *Solving the Problem from Hell: Tripartism as a Strategy for Addressing Labour Standards Non-Compliance in the United States*, 50 OSGOODE HALL L.J. 813, 841 (2013). During those confirmation hearings, W&HW became a major focus, with one employer-

backlash that has undermined previous public-private regulatory experimentation will require exploration of these risks and doctrinal threats, which this Article will address in Part IV.

Recently, forms of public-private regulatory experimentation have emerged as durable models to amplify the deterrent effect of enforcement.⁹⁴ The New York Attorney General (NYAG) has partnered intensively with PIOs in low-wage sectors,⁹⁵ and since 2007 San Francisco has required its Office of Labor Standards Enforcement (SF-OLSE) to fund non-union PIOs with renewable grants to channel worker complaints to SF-OLSE.⁹⁶ These successes with PIO collaboration have inspired other state and local governments to experiment with formal collaboration to regulate workplace law.⁹⁷

To date there has been insufficient attention paid to why collaboration between PIOs and these agencies have endured where these other forms of tripartism could not.⁹⁸ To evaluate this emergent

backed advocacy group claiming that W&HW “could very likely be a model used by Smith . . . on a national level [I]t could turn tens of thousands of ‘community organizers’ into raving vigilantes nationwide.” *Id.* at 831. While Smith was eventually confirmed, W&HW was shelved. *Id.*

- 94 Janice Fine has provided detailed case studies of many examples of collaboration between PIOs and public agencies, ranging from informal cross-referral systems to more sophisticated systems to integrate public and private enforcement strategies to enforce labor and safety and health standards. FINE, CO-PRODUCTION, *supra* note 7, at 18–29.
- 95 See generally Jennifer Brand, *The Role of State Attorneys General in the Enforcement of Labor Laws*, NATIONAL STATE ATTORNEYS GENERAL PROGRAM AT COLUMBIA LAW SCHOOL (2007), available in <http://www.stateag.org/policy-areas/labor/labor-resources/2016/11/22/the-role-of-state-attorneys-general-in-the-enforcement-of-labor-laws>.
- 96 FINE, CO-PRODUCTION, *supra* note 7, at 23–25; see San Francisco, S.F., Cal., Ord. No. 140687, Amending S.F. Mun. Code § 12R.25 (July 17, 2014) (“The Office of Labor Standards Enforcement shall establish a community-based outreach program to conduct education and outreach to employees. In partnership with organizations involved in the community-based outreach program, the Office of Labor Standards shall create outreach materials that are designed for workers in particular industries.”).
- 97 Since SF-OLSE began funding PIOs to collaborate, three other cities and counties, including Seattle, Washington, and Berkeley and Los Angeles, California, have adopted a similar approach. Seattle, Wash., CODE, § 3.14.934 (2014); L.A. County, Cal., Wage Enforcement Order § 8.101.090(G) (2016); Berkeley, Berkeley, Cal., Mun. Code Ch. 13.99, § 13.99.080 (2016); Los Angeles Cnty., Cal., Ord. No. 102703 amending County Code, Title 8 - Consumer Protection, Business and Wage Regulations relating to the enforcement of the County Minimum Wage Ordinance, § 8.101.090(G) (Apr. 26, 2016).
- 98 Matthew Amengual and Janice Fine propose building internal PIO and public agency support for collaborations, which they call “co-enforcement,”

enforcement approach, and its implications for public regulation, a reevaluation of the benefits, limitations and risks of collaborative enforcement is in order. The next Part will offer two different collaborative techniques to bridge enforcement gaps rather than to shift the responsibility for compliance to the regulated entities. Examining these forms of emergent public-private regulatory experimentation will show how collaboration can effectively and efficiently counteract entrenched norms of noncompliance and elaborate new standards.

To understand why this is the case, the next Part will first describe the collaborative enforcement approaches in detail through mini-case studies. It will visually depict the integration of public and private enforcement tools in a single, hybrid enforcement pyramid. It will then evaluate two different techniques for collaborative enforcement, remedial and grant-based, and offer them together as a conceptual framework to amplify the deterrent effect of enforcement by integrating enforcement tools, often with the ultimate aim of strengthening labor standards through law reform. Part IV will then examine the normative risks of and doctrinal threats to collaborative enforcement.

III. Public-Private Regulatory Experimentation to Improve the Deterrent Effect of Enforcement

This Part first introduces the concept of collaborative enforcement through case studies, showing how public agencies and PIOs have increasingly collaborated through two different techniques: (1) *remedial enforcement*, in PIO-led campaigns to intensively deploy public and private enforcement tools to change the structure of the law in regions and sectors where existing remedies are inadequate; and (2) *grant-based enforcement*, by funding PIOs to integrate their private capacity and expertise into public enforcement to channel enforcement into underregulated workplaces or to elaborate a new

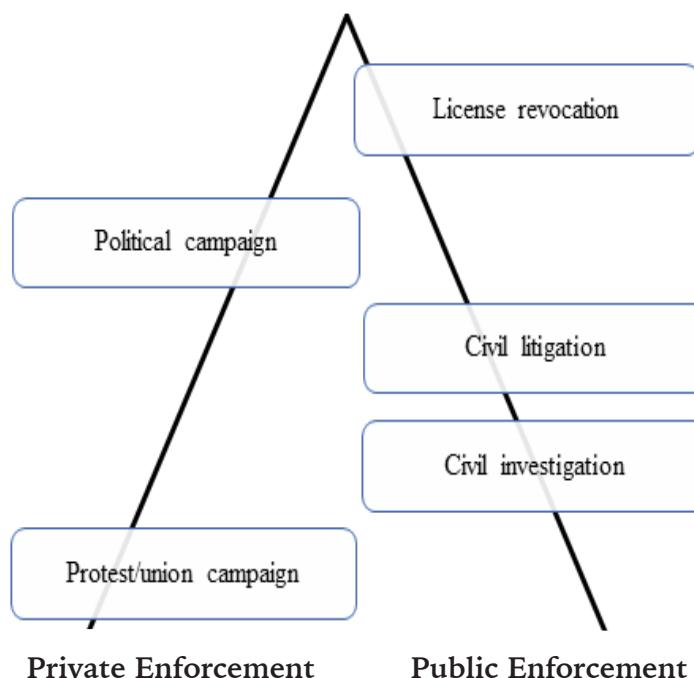
by sustaining partnerships between public agency and PIO staff. Matthew Amengual & Janice Fine, *Co-enforcing Labor Standards: The Unique Contributions of State and Worker Organizations in Argentina and the United States*, 11 REG’N & GOVERNANCE 129, 129–30 (2017). This Article agrees that to expand public-private regulatory experimentation, it will be necessary for public agency and PIO staff to recognize the benefits of collaboration and overcome mutual mistrust while negotiating conflict. But it will also be necessary to understand why some collaborations succeed while others fail, and the legal limitations and incentives for abuse that can constrain or undermine collaborative techniques.

standard. It uses the same visual depiction of an enforcement pyramid from the previous Part, but now public and private enforcement tools are integrated into a single pyramid to show how collaborative enforcement can effectively address enforcement gaps identified in the previous Part.

A. Collaborative Enforcement

1. Remedial Enforcement

In the remedial enforcement approach, PIOs and public agencies coordinate their use of top-tier private and public enforcement tools in regional and sectoral initiatives. The goal of remedial enforcement is often to improve and create new remedies in a low-wage sector through law reform.



In a remedial enforcement approach, PIOs and public agencies temporarily and intensively focus public and private remedies on a particular problem or industry. Public agencies and PIOs access top-pyramid tools, such as public agency civil investigation, litigation, and license revocation powers, and PIO public protests, unionization drives and political campaigns to improve legal remedies through

law reform. Public agencies and PIOs that engage in remedial enforcement often expressly affiliate with each other's goals in joint media statements and jointly support law reform to improve remedies.

A remedial enforcement approach permits public agencies and PIOs to access uniquely private and public enforcement tools to resolve cases and to improve and create new remedies without formal delegation. In a remedial enforcement model, PIOs and public agencies engage in only limited coordination of priorities and referrals, and there is no delegation of case detection, selection or resolution to PIOs. The fast food and car wash industries are recent sites of remedial enforcement by PIOs and public agencies.

Fast Food Establishments in New York State

Beginning in 2012, Service Employees International Union ("SEIU") and PIOs in New York began organizing fast food workers in a "Fast Food Forward" campaign with the aspirational goal of a \$15 an hour wage and a collective bargaining agreement in the fast food industry.⁹⁹ By November 2012, 200 New York City fast food workers walked off their jobs, and protests swelled to 200 cities by 2015, the largest protest of low-wage workers in U.S. history.¹⁰⁰

In New York beginning in 2012, Fast Food Forward identified witnesses with similar complaints of violations of wage and hour law and referred them to NYAG to fuel a state-wide investigation into violations by wage and hour law by franchise stores of many of the nation's largest franchisors.¹⁰¹ By 2015 NYAG announced over a dozen settlements with Domino's Pizza, Papa John's, McDonald's, and KFC franchisees, amounting to nearly \$3 million in restitution for franchise store employees throughout New York State.¹⁰²

99 Steven Greenhouse & Jana Kasperkevic, *Fight for \$15 Swells into Largest Protest by Low-Wage Workers in US History*, GUARDIAN (Apr. 15, 2015), <http://www.theguardian.com/us-news/2015/apr/15/fight-for-15-minimum-wage-protests-new-york-los-angeles-atlanta-boston>; William Finnegan, *Dignity*, NEW YORKER, Sept. 15, 2014, at 74; see also ROLF, *supra* note 48, at 91–96 (describing formation for Fast Food Forward movement).

100 Greenhouse & Kasperkevic, *supra* note 99.

101 Paul Harris, *New York Attorney General Investigating State's Fast Food Industry*, GUARDIAN (May 16, 2013), <https://www.theguardian.com/world/2013/may/16/new-york-fast-food-attorney-general-investigation>.

102 Press Release, N.Y. Office of the Attorney General, A.G. Schneiderman and U.S. Dep't of Labor Announce Criminal Charges Against, and Civil Settlement With, Papa John's Franchisee for Wage Theft (July 15, 2015), <https://ag.ny.gov/press-release/ag-schneiderman-and-us-department-labor-announce>.

In one instance, after Fast Food Forward referred labor complaints about these establishments to NYAG, NYAG received allegations that a franchisee store manager had fired staff who complained about performing extra work after clocking out for no pay.¹⁰³ By the following day, NYAG reached an agreement reinstating the complainants immediately.¹⁰⁴

NYAG's coordination with Fast Food Forward and subsequent finding of widespread employment law violations in fast food stores also led the public agency to support the PIOs' demand for higher state labor standards. The New York Attorney General appeared in Fast Food Forward rallies,¹⁰⁵ and on April 15, 2015 he called on the New York Governor to unilaterally raise the state minimum wage for fast food workers.¹⁰⁶ The following month New York Governor Cuomo announced a plan to convene a wage board specifically for the fast food industry with representation from the fast food employers, unions, and the state.¹⁰⁷ The wage board recommended a phased-in \$15 minimum wage for fast food workers in New York State, which the New York Department of Labor issued as an order in September 2015.¹⁰⁸ The remedial enforcement effort also resulted in a lawsuit

criminal-charges-against-and-civil; *see also* Will Bredderman, *NY Attorney General Declares Fast Food Wage Theft a 'Crime Wave'*, *OBSERVER* (Mar. 5, 2015), <http://observer.com/2015/03/ny-attorney-general-declares-fast-food-wage-theft-a-crime-wave>.

- 103 See Jan Ransom, *Fired Washington Hts. Dominos Workers to be Reinstated After Walkout, Sez Attorney General*, N.Y. DAILY NEWS (Dec. 12, 2013), <http://www.nydailynews.com/new-york/manhattan/ag-delivers-dominos-workers-article-1.1546342>.
- 104 3683 Washington Heights Pizza, LLC. et al., AOD No. 13-491 (N.Y. Att'y Gen., Labor Bureau Dec. 12, 2013) (assurance of discontinuance).
- 105 Ginger Adams Otis, *AG Eric Schneiderman Will Join Fight for \$15 Rally Alongside Health Care Workers*, DAILY NEWS (Apr. 14, 2015), <http://www.nydailynews.com/new-york/ag-schneiderman-join-fight-15-rally-article-1.2185186>.
- 106 Eric Schneiderman, *Gov. Cuomo Can Unilaterally Hike N.Y. Wages: How His Labor Department Can Use Its Existing Legal Authority to Raise Suffering Workers' Pay*, DAILY NEWS, (Apr. 16, 2015), <http://www.nydailynews.com/opinion/schneiderman-cuomo-unilaterally-hike-n-y-wages-article-1.2186808>.
- 107 Andrias, *supra* note 6, at 64–65. New York Labor Law permitted the creation of wage boards, with representation by employers, employees and the public, to recommend wage rates for occupations in New York State. N.Y. Lab. L. § 655. *See Nat'l Rest. Ass'n v. Comm'r Labor*, No. 15-001, Resolution of Decision and Order, at 1–4 (N.Y. Ind. Bd. App. Dec. 9, 2015), https://www.governor.ny.gov/sites/governor.ny.gov/files/atoms/files/WB_5-001_Decision.pdf. The wage board's powers were suspended after the minimum wage increase. *See* Andrias, *supra* note 6, at n.344.
- 108 Patrick McGeehan, *New York Plans \$15-an-Hour Minimum Wage for Fast Food*

by the New York Attorney General against a franchisor, Domino's Pizza, for wage-and-hour-law violations by its franchisees in New York franchise stores.¹⁰⁹ That case is pending.

In response to Fast Food Forward, New York City has also recently facilitated the self-funding of PIOs in the fast food sector, passing a law in May 2017 permitting fast food employees to make voluntary contributions to not-for-profit organizations of their choice through payroll contributions.¹¹⁰

Car Wash Establishments in California and New York

Remedial enforcement in the car wash sector began in Los Angeles in 2008, where PIOs formed a coalition called the "CLEAN Carwash Campaign," which assisted workers in bringing their complaints of wage theft and health and safety violations to state and local enforcement agencies.¹¹¹ Through these referrals, over five years the California Labor Commissioner issued 1,423 citations to car wash establishments in the state for wage and hour law violations and registration violations and has assessed \$11.6 million in penalties and \$4.2 million in back wages.¹¹² The California Attorney General's office brought a series of civil lawsuits against other operators and in 2012 announced a \$1 million settlement with an owner of eight car wash establishments in Southern California for underpayments.¹¹³

This public and private enforcement provided support for union-PIO lobbying for improved remedies and a unionization campaign.¹¹⁴ To address the persistent problem of judgment collection,¹¹⁵ the CLEAN campaign successfully lobbied for a state restitution fund for workers to recover unpaid wages, which

Workers, N.Y. TIMES, (July 22, 2015), <https://www.nytimes.com/2015/07/23/nyregion/new-york-minimum-wage-fast-food-workers.html>.

109 New York v. Domino's Pizza, Inc., No. 450627/2016 (N.Y. Sup. Ct. May 23, 2016) (alleging that the franchisor, Domino's Pizza, jointly employed franchise store employees and violated state franchise law in failing to correct or disclose flaws in its payroll system that underreported wages owed by franchisees to their employees).

110 N.Y.C. Mun. L. 2017/098 (2017).

111 FINE, CO-PRODUCTION, *supra* note 7, at 18–19.

112 Margot Roosevelt, *Union-Backed Car Wash Workers Fight for More Pay*, *Shade, ORANGE COUNTY REGISTER* (Sept. 1, 2014, 4:04 PM), <http://www.ocregister.com/articles/car-633355-workers-wash.html>.

113 FINE, CO-PRODUCTION, *supra* note 7, at 25.

114 Sonia Nazaro, *Unions Join to Organize Carwash Workers*, *L.A. TIMES* (Mar. 28, 2008), <http://www.latimes.com/local/la-me-carwash-280308-story.html>.

115 FINE, CO-PRODUCTION, *supra* note 7, at 25.

is funded by car wash employer registration and annual fees and civil penalties for failure to register.¹¹⁶ The CLEAN campaign also successfully lobbied to increase the labor bond requirement for the car wash industry.¹¹⁷ Since that time, 25 car wash establishments have entered into collective bargaining agreements with the union.¹¹⁸

In New York City, PIOs used a similar remedial enforcement approach to improve labor standards in New York car wash establishments. The PIOs formed a coalition called “Wash New York” to organize car wash workers to improve labor standards in the car wash industry.¹¹⁹ It targeted the owners of a large network of car wash establishments, one of whom had previously been found to have violated wage-and-hour law by USDOL. By 2014 NYAG announced an agreement for these employers to pay nearly \$4 million in restitution for unpaid wages and penalties for failing to pay unemployment insurance contributions and workers compensation premiums.¹²⁰ During this investigation, after the owner sold one of his car wash establishments under investigation, the NYAG obtained an agreement to transfer the workers to other work sites.¹²¹ Since then, workers in eleven car wash establishments in New York won union recognition.¹²²

The New York City Council, similar to California’s legislature, responded with legislation requiring car wash operators to be licensed with the city and to obtain a bond to pay for unpaid wages, with an exemption (or “opt-out”) for establishments with a

116 CAL. LAB. CODE §§ 2055 (describing registration), 2065 (describing restitution fund) (West 2015).

117 CAL. LAB. CODE § 2055(b)(4) (West 2015).

118 Roosevelt, *supra* note 112; FINE, CO-PRODUCTION, *supra* note 7, at 19.

119 WASH NEW YORK, THE DIRTY BUSINESS OF CLEANING NYC’S CARS: CAR WASH WORKERS FACE LOW PAY, OFFENSIVE CONDITIONS, AND POOR TREATMENT (2012), <http://op.bna.com.s3.amazonaws.com/dlrcases.nsf/r%3FOpen%3Dkpin-8s5ssg>.

120 Erica Pearson, *Carwash Kingpin John Lage and Associates to Pay \$3.9M to Settle Labor Violations, Says N.Y. Attorney General Eric Schneiderman*, DAILY NEWS (Mar. 6, 2014), <http://www.nydailynews.com/new-york/carwash-kingpin-pay-3-9m-labor-violations-ag-article-1.1712206>.

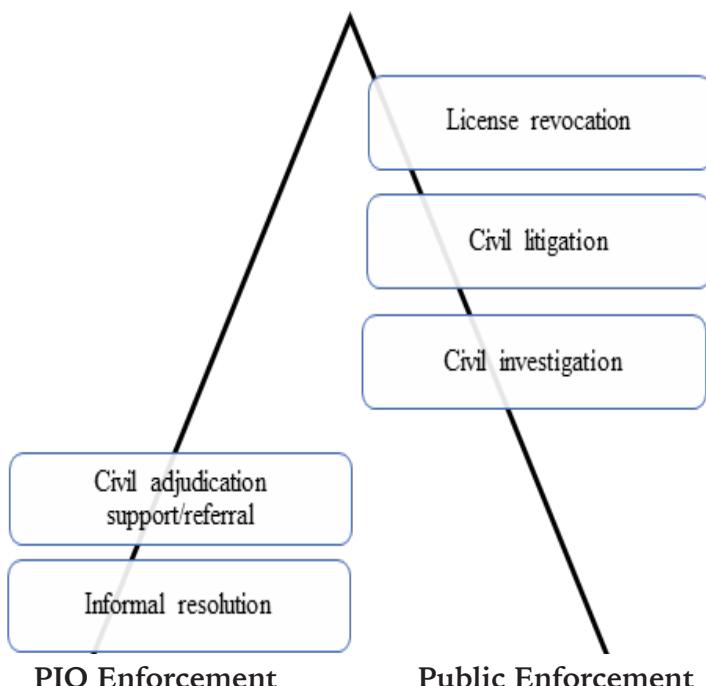
121 Julie Turkewitz, *Workers Who Protested Sale of Carwash Will Get New Jobs*, N.Y. TIMES (Feb. 20, 2013), <http://www.nytimes.com/2013/02/21/nyregion/after-sale-of-carwash-owners-agree-to-give-workers-new-jobs.html>.

122 Stephen Rex Brown, *Brooklyn ‘Carwasheros’ Vote to Join Union to Secure Minimum Wage Plus Tips, Overtime*, N.Y. DAILY NEWS (Feb. 16, 2016), <http://www.nydailynews.com/new-york/brooklyn-carwash-workers-vote-join-union-article-1.2533762>.

collective bargaining agreement.¹²³ This opt-out provision of the law, however, has been enjoined by a trial court, which found that it is preempted by the NLRA.¹²⁴

2. *Grant-based Enforcement*

In a grant-based enforcement approach, legislatures direct public agencies to provide grants to PIOs to identify and triage complaints and channel them into public enforcement. While a grant-based enforcement model requires PIOs to deliver specific services—typically education, triage and referral—it does not deputize PIOs as public enforcement officers or delegate policymaking decisions, such as whether and how to use particular public enforcement tools.



The most prominent example of grant-based enforcement is administered by San Francisco's Office of Labor Standards

123 Erin Durkin, *NYC Car Washes Must Get Licenses Under Bill Passed by City Council in Effort to Stem Labor Violations*, N.Y. DAILY NEWS (June 10, 2015), <http://www.nydailynews.com/new-york/nyc-car-washes-licenses-bill-article-1.2253812>.

124 Ass'n of Car Wash Owners Inc. v. New York, 15-CV-08157 (AKH), Order Granting Pl's S.J. Mot. 5–10 (S.D.N.Y. May 26, 2017).

Enforcement (SF-OLSE).¹²⁵ The city and county legislature grants a million dollars annually to fund seven PIOs to provide enforcement services, overseen by SF-OLSE.¹²⁶ The grant requires the PIOs to counsel and refer workers with allegations of labor standards violations to SF-OLSE, the California Department of Labor Standards Enforcement, or USDOL, or to resolve them directly with the employer.¹²⁷ SF-OLSE also requires PIOs to issue joint media statements about resolved cases and to provide trainings with SF-OLSE investigators in the communities targeted for enforcement. SF-OLSE investigators may accept documentation about cases from the PIOs and jointly work on cases with PIOs, and must participate in quarterly meetings with PIOs.¹²⁸

The legislature delegates to SF-OLSE the selection, training and monitoring of PIOs,¹²⁹ while SF-OLSE delegates to a single lead PIO, the Chinese Progressive Association, the task of coordinating with other selected PIOs to meet grant obligations.¹³⁰ The legislature in 2012 also created a “Wage Theft Task Force,” to make recommendations and issue reports about how public agencies may improve their public enforcement, and appointed representatives, including SF-OLSE and other public agencies responsible for enforcing labor standards and other PIOs and employer groups.¹³¹

San Francisco’s grant process has resulted in a large inflow of complaints from the underregulated, low-wage workplaces. Nearly half of the complaints SF-OLSE receives are about restaurants, many of which come from referrals by the Chinese Progressive Association, which focuses on the Chinatown restaurant sector.¹³² After referral,

125 For a detailed description of the SF-OLSE’s partnerships with PIOs, see FINE, CO-PRODUCTION, *supra* note 7, at 23–25.

126 See FINE, CO-PRODUCTION, *supra* note 7, at 23–25.

127 San Francisco Office of Labor Standards Enforcement Wage Theft Prevention Education & Outreach Services Agreement, A-3(1)(A) (on file with author).

128 See FINE, CO-PRODUCTION, *supra* note 7, at 24.

129 San Francisco Office of Labor Standards Enforcement Wage Theft Prevention Education & Outreach Services Agreement, A-3(1)(A) (on file with author).

130 *Id.* at A-1.

131 S.F., CAL., Ordinance No. 102–12 (2012); S.F., CAL. Admin. Code §§ 5.260–5.260–5 (2012); see SAN FRANCISCO WAGE THEFT TASK FORCE, FINAL REPORT 3–6 (2013).

132 SAN FRANCISCO WAGE THEFT TASK FORCE, *supra* note 131, at 11, 23. The San Francisco Wage Theft Task Force reports that SF-OLSE resolves nearly all cases for at least the amount that it determines that the employer has underpaid employees, often with the assistance of a PIO. *Id.* at 16, 26 (SF-OLSE resolves 99% of its cases where it finds back wages owed, 91% of those

PIOs also use private remedies to resolve about a third of them with the employer without any required action by SF-OLSE.¹³³ The San Francisco grant process expressly directs SF-OLSE investigators to collaborate in cases with PIOs, which has resulted in large-scale enforcement. For example, CPA jointly worked on a case with SF-OLSE against Yank Sing Restaurant, where workers were initially too afraid to come forward and where the payroll records falsely showed no violations. CPA identified workers willing to show the public agency that the employer's payroll records were fabricated.¹³⁴ It brought a private lawsuit, and CPA, SF-OLSE and DLSE jointly negotiated with the employer for a \$4 million settlement and a "workplace change agreement" with Yank Sing that required the restaurant to recognize rights not guaranteed by law, including progressive discipline and paid holidays, monitored by CPA and Yank Sing employees.¹³⁵ In 2014 SF-OLSE, CPA and the employer issued a joint press release hailing the agreement.¹³⁶

Seattle adopted a similar model as San Francisco through its Office of Labor Standards (Seattle-OLS), created in 2014 to enforce local workplace laws, including the city's minimum wage and paid sick time requirements.¹³⁷ In September 2015, Seattle's City Council provided \$1 million for Seattle-OLS to fund a PIO to act as a hub to train ten other PIOs to provide outreach, education and technical assistance.¹³⁸ Seattle-OLS also approved a similar grant

for at least the owed amount of back wages.).

133 SAN FRANCISCO OFFICE OF LABOR STANDARDS ENFORCEMENT, MINIMUM WAGE ORDINANCE ANNUAL REPORT FY 2012-2013 (2013), https://sfgov.org/olse/sites/default/files/FileCenter/Documents/11225-FINAL%2069_MinWageOrdinanceReport_1213.pdf (PIOs independently resolved eight of the 25 minimum wage cases they referred to SF-OLSE.).

134 See FINE, CO-PRODUCTION, *supra* note 7, at 27.

135 Press release: Immigrant Workers Negotiate \$4 Million Settlement, CHINESE PROGRESSIVE ASS'N, (Nov. 9, 2015), <http://www.cpasf.org/press/settlement-yanksing>.

136 *Id.*

137 Elmore & Chishti, *supra* note 28, at 38–39. Seattle-OLS was initially housed in Seattle's Office for Civil Rights, which principally enforces anti-discrimination protections, but now is a stand-alone agency. *Id.* Like the SF-OLSE, Seattle-OLS has rulemaking authority and can issue administrative orders for restitution and penalties. SEATTLE, WASH., ADMIN. CODE §§ 14.19.075 (administrative order authority); 14.20.040(B) (rulemaking authority) (2015).

138 Seattle-OLS distributed these funds to ten PIOs, designating one lead PIO, the Fair Work Center, to provide technical assistance to nine other PIOs, to engage in "outreach, hosting community-based education events, developing training materials to educate workers and other organizations about Seattle's

to employer associations to educate small businesses about how to comply with the requirements.¹³⁹ Seattle additionally established the Labor Standards Advisory Commission, appointed by the Mayor and City Counsel and composed of PIOs representing the interests of businesses and workers. The Labor Standards Advisory Commission advises Seattle-OLS, the Mayor and City Council about labor standards, and provides feedback about Seattle-OLS's proposed rulemaking.¹⁴⁰

B. Evaluation of Collaborative Enforcement

This Article's account of collaborative enforcement reveals underdiscussed benefits of enforcement collaboration between PIOs and public agencies. First, these techniques can effectively and efficiently address capacity, expertise and remedy gaps, by integrating private and public enforcement tools within a single enforcement pyramid, channeling enforcement to particular sectors or regions of the economy, and by changing the structure of the law when existing remedies are inadequate. Collaborative enforcement also permits state and local government to channel private enforcement to specific low-wage sectors and to elaborate the requirements of a new labor standard. Third, collaborative enforcement improves upon the political accountability of public enforcement, which can insulate public agencies from capture by the regulated entities. Lastly, collaborative enforcement can facilitate sophisticated forms of tripartism, including social bargaining between worker and employer stakeholders in lawmaking and private, firm-specific bargaining to resolve noncompliance through codes of conduct and unionization.

1. Collaborative Enforcement Can Effectively and Efficiently Address Enforcement Pathologies

These examples show how PIOs and public agencies can create a more complete enforcement pyramid by integrating

labor standards, and providing labor rights intake, counseling, and referral for workers experiencing labor standards violations." 2015-2016 Community Outreach and Education Fund, OFFICE OF LABOR STANDARDS, <https://www.seattle.gov/laborstandards/outreach/community-fund/2015-2016-community-outreach-and-education-fund> (last visited Jan. 15, 2017); Elmore & Chishti, *supra* note 28, at 38-39.

139 Elmore & Chishti, *supra* note 28, at 38-39. The most recent budget has increased the annual grant to \$1.8 million for PIOs for worker education, outreach and referrals and \$800,000 for business trainings. *Id.*

140 *Id.*

enforcement tools that PIOs and public agencies could not access independently. First, collaboration can address enforcement pathologies by bringing complainants into the enforcement regime. In the remedial enforcement technique, PIOs' ties to complainants allowed public agencies to quickly identify targets. Large-scale detection in these industries then enabled the public agencies to use top-tier remedies to address problems that they would otherwise be unable to detect through independent enforcement. As with the SF-OLSE and Seattle-OLS examples, providing grants to PIOs to conduct education and outreach can also provide public agencies with access to complainants to identify enforcement gaps in new labor standards. The SF-OLSE requirement that PIOs assist workers in resolving some of the cases they triage before referring them permits the public agency to establish informal mediation as a bottom-pyramid enforcement step.

The intermediate public enforcement step of civil investigations by NYAG of car wash establishments and fast food franchisees allows PIOs to resolve potentially protracted wage-and-hour and retaliation cases. Investigating retaliation as a form of witness tampering or obstruction of justice enables public agencies to secure a remedy for retaliation long before the NLRB's hearing process could even begin, which the employers may have defeated through delay and attrition.

Third, collaborative enforcement also improves access to top-pyramid enforcement tools to incapacitate bad actors and to change the structure of the law where the existing remedies are inadequate. PIO mass referral of car wash complainants for public enforcement in California led the public agency to issue penalties in amounts greater than the cost of compliance. As in the Yank Sing example, availability of top-pyramid public enforcement remedies also strengthens PIO monitoring by backing it with a credible threat of public enforcement if the employer is unwilling to comply. Where enforcement tools are inadequate, remedial enforcement can change the structure of the law to create or improve a remedial or sanctions regime.¹⁴¹ After complaints from PIOs enabled the New

¹⁴¹ In this manner, agencies can use collaborative enforcement to coordinate with PIOs in lobbying for regulatory changes that make enforcement more responsive. See DANIEL CARPENTER, REPUTATION AND POWER: ORGANIZATIONAL IMAGE AND PHARMACEUTICAL REGULATION AT THE FDA 82 (2010) (describing coordination between U.S. Food and Drug Administration, which was legally prohibited from lobbying, and allied

York Attorney General to conduct a wide-ranging investigation into the pay practices of fast food franchisees, the NYAG now seeks law reform through litigation that, if successful, may establish the franchisor's joint liability for its franchisees' wage-and-hour law violations in New York State.¹⁴² In the car wash examples, access to labor bonds enables PIOs and public agencies to efficiently distribute owed restitution from a surety instead of a potentially insolvent employer. The car wash licensure requirement of labor standards compliance also improves public enforcement with a top-pyramid sanction for egregious violations.

Performing these functions within a single enforcement pyramid can be more efficient than independent enforcement, by avoiding duplication in the use of enforcement tools. PIOs often perform many of the functions otherwise performed by public agencies, such as witness interviews and triage, and informal PIO resolution permits public agencies to strategically reserve their resources. PIO collaboration in case resolution provides public agencies with access to workers through PIOs in monitoring compliance, which can be more effective and efficient than on-site inspections.

Collaborative enforcement also permits public agencies to channel private enforcement to specific enforcement gaps. As with SF-OLSE's PIO grants, public agencies in a grant-based enforcement approach may designate specific regions or sectors where education, outreach, referral and private enforcement may improve compliance. SF-OLSE's selection of CPA as a collaborating PIO, for instance, effectively channels enforcement capacity into San Francisco's Chinatown restaurants. PIOs may also channel public enforcement through remedial enforcement. PIO access to workers and use of private enforcement in the fast food and car wash examples provided NYAG and the California agencies with a compelling reason to prioritize those sectors to amplify the deterrent effect of their enforcement tools.

Lastly, collaborative enforcement attends to the problem of PIO or public agency exit from the collaborative relationship before its goals are met because of the unfunded costs of collaboration or

PIOs, in which the FDA publicized need for greater regulation over food and medicine while PIOs lobbied for more expansive FDA regulatory powers).

142 See Andrew Elmore, *The Future of Fast Food Governance*, 165 U. PA. L. REV. ONLINE 73, 79, 87 (2017), <http://www.pennlawreview.com/online/165-U-Pa-L-Rev-Online-73.pdf>.

the shifting priorities of PIOs and public agencies.¹⁴³ Premature exit wastes enforcement resources that would have been better invested in other forms of regulation that did not require coordination. The legislative grants of grant-based enforcement reduce the exit incentive by sustaining costly enforcement functions that PIOs otherwise lack the resources to perform. Remedial enforcement, in contrast, controls for exit by requiring a commitment of significant upfront resources. In the case of the car wash campaigns in Los Angeles and New York, the CLEAN and WASH campaigns invested in the collaborative relationship until they achieved the legislative solution of a bonding and licensing regime that they could monitor. In the fast food initiative in New York, Fast Food Forward invested in the collaborative relationship until the referred complaints had been successfully resolved, and the evidence of widespread noncompliance with labor standards developed a sufficient factual record for the NYAG to seek law reform through litigation. Once these goals were met, the exit costs accordingly dropped for the participating PIOs and public agencies.

2. Collaborative Enforcement Injects Political Accountability into Public Enforcement and Prevents Capture by the Regulated Entities

Collaborative enforcement can also inject political accountability into public enforcement decisions by making public agency enforcement priorities transparent to PIOs, and promote public agency independence by insulating public agencies from capture by the regulated entities.¹⁴⁴ This is in sharp contrast to independent public enforcement, in which there is no requirement that public agencies subject their enforcement priorities to public scrutiny, or even to tell the public what those priorities are.¹⁴⁵ Public agencies often need not disclose communications with private entities, and anti-lobbying rules exempt enforcement decisions.¹⁴⁶ The opacity of public enforcement has understandably raised

143 Early examples of collaborations between PIOs and public agencies failed to attend to high exit incentives. See, e.g. ESTLUND, *supra* note 7, at 114.

144 As Margaret Lemos argues, the challenge of public enforcement is “designing enforcement institutions in a way that promotes accountability while preserving a role for independent, professional judgment.” Margaret H. Lemos, *Democratic Enforcement? Accountability and Independence for the Litigation State*, 102 CORNELL L. REV. 929, 935 (2017).

145 *Id.* at 933.

146 *Id.* at 999.

concerns about the possibility of capture by the regulated entities.¹⁴⁷

In this backdrop of enforcement unaccountability, the presence of PIOs in the enforcement design can inject political accountability into public agency enforcement decisions.¹⁴⁸ Seattle's Labor Standards Advisory Commission, which provides guidance to Seattle OLSE about enforcement priorities and produces an annual report about Seattle OLSE's enforcement outcomes,¹⁴⁹ provides a greater level of public transparency. Seattle's approach of inviting public participation in public agency enforcement is a traditional role for local government, which "often operate[s] at the edge of a blurry line between governmental action and public participation,"¹⁵⁰ and, as Nestor Davidson argues, this type of public participation in local administration "may bolster the ability of local regulators to resist capture by outside regulated entities."¹⁵¹ The legislative funding of grant-based enforcement, and its formal grant-making procedures, also permit interested third party stakeholders—legislators, ethics groups, and employer associations—to monitor collaborative enforcement through required PIO and public agency disclosures about the use of enforcement resources. In remedial enforcement, elected officials and agency heads publicly state the goals of these high-profile enforcement initiatives, which subjects them to political accountability should the initiative fail or succumb to abuse.

PIOs are also more accountable to the public in collaboration with public agencies than in informal collaboration. In the SF-OLSE and Seattle-OLS examples, these public agencies require PIOs to adhere to a formal grant process, with limited terms, and

147 *Id.* at 949 & n.85.

148 In the words of Ayres and Braithwaite, "in the presence of empowered [PIOs] the firm must capture [PIOs] as well as the [public] agency to be effective." *AYRES & BRAITHWAITE*, *supra* note 5, at 71.

149 See *Labor Standards Advisory Commission*, SEATTLE.GOV, <https://www.seattle.gov/labor-standards-advisory-commission> (last visited Oct. 9, 2017); *Labor Standards Advisory Commission Meetings*, SEATTLE.GOV, <https://www.seattle.gov/labor-standards-advisory-commission/what-we-do/meetings>(last visited Oct. 9, 2017).

150 Davidson, *supra* note 15, at 572.

151 Davidson, *supra* note 15, at 630. David Freeman Engstrom echoes this argument in his empirical analysis of qui tam suits, finding that public agency delegation to private litigants pursuing enforcement actions "can improve, rather than degrade, democratic politics by offering a salutary counterweight to 'capture' and other patterns of political control within the legislative or administrative process." Engstrom, *supra* note 79, at 2003–4.

performance reviews with measurable outcomes.¹⁵² In jurisdictions in which multiple PIOs seek to deter employment law violations, the contested nature of the grant selection process creates a natural incentive for other PIOs to monitor the collaborating PIOs for abuse.¹⁵³ Abuse by one PIO is also limited by the SF-OLSE and Seattle-OLS reporting structure, in which one PIO grantee is the primary contractor accountable for the enforcement activities of the other PIOs grantees. These controls provide for a greater level of public accountability than exist in pure public enforcement.

3. Collaborative Enforcement Improves Effectiveness and Governance of Public-Private Enforcement Through Tripartism

Finally, these examples show how collaborative enforcement can facilitate sophisticated forms of tripartism unlikely to devolve into weak forms of self-regulation. The remedial enforcement of the CLEAN Carwash campaign and California state agencies is a paradigmatic example. Developing a public enforcement initiative around a PIO's enforcement priority, and the subsequent investigation findings of widespread noncompliance, created an incentive for the public agency to support the PIOs' legislative goal of strengthening labor standards in the sector. The license and bonding requirements now required in the California car wash sector permit PIOs to continue to coordinate with state agencies in seeking remedies for employment law violations and to incapacitate bad actors through license revocation. Grant-based enforcement can also facilitate durable forms of tripartite lawmaking. Seattle's Labor Standards Advisory Commission creates opportunities for tripartite lawmaking through its appointment of worker- and business-affiliated stakeholders to advise the legislature and Seattle-OLS

152 See Alfred C. Aman, Jr., *Privatization and Democracy*, in FREEMAN & MINOW, *supra* note 8, at 264–67 (arguing that the contract process can improve accountability by de-politicizing the privatization debate, and focusing the PIOs and public agency on the priorities and trade-offs of outsourcing a particular function).

153 See AYRES & BRAITHWAITE, *supra* note 5, at 73 (arguing that PIO abuse is less likely where the PIO's role in enforcement can be contested by other PIOs). In addition to serving as a check on PIO abuse of grant-based enforcement, third-party monitoring also permits PIOs to evaluate whether selected PIOs effectively represent the interests of the communities that they are charged with assisting, which the public agency may lack the sophistication to evaluate without outside PIO assistance.

about the agency's rulemaking.¹⁵⁴

Collaborative enforcement can also facilitate private firm-level bargaining to improve labor standards. In the car wash examples, remedial enforcement facilitated unionization as a way to regulate low-wage workplaces, and protect complainants from retaliation for complaining about employment law violations.¹⁵⁵ Unionized employers in underregulated sectors are also more likely to support industry compliance with labor and safety and health standards to put them on equal footing with non-union competitors.¹⁵⁶ These workers and employers are therefore more likely to participate in the kinds of tripartism championed decades ago, in which employers and workers share the motivation and expertise, and can privately bargain for their respective roles, in enforced self-regulation.¹⁵⁷

Even with more routine, episodic forms of enforcement, grant-based enforcement can improve the likelihood of tripartite resolutions by assigning clear and durable roles to state actors and PIOs. In the Yank Sing example, CPA's protests of labor conditions in the Yank Sing restaurant opened a pathway for the PIO and public agencies to negotiate a tripartite agreement with the employer that permitted a more complete set of prospective, enforceable labor standards than in a standard resolution of a private wage-and-hour-law suit. Facilitating tripartism, as illustrated by these examples, can also channel employers into private bargaining with PIOs for higher and more uniform industry standards.

Lastly, in addition to improving the effectiveness of enforcement, collaborative enforcement can serve the value of enhancing democratic participation by channeling private participation into public enforcement. SF-OLSE's intervention in

154 Elmore & Chishti, *supra* note 28, at 38–39.

155 In this light, unionization serves important compliance goals. Workers who have the assistance of a union are more likely to know about their labor and safety and health rights, and those with collective bargaining agreements that require good cause for discipline and terminations are more likely to complain about labor and safety and health violations. WEIL, *supra* note 1, at 20–36, 245–48. Firms subject to a collective bargaining agreement have lower rates of workplace fatalities and serious injuries, and are more likely to comply with wage-and-hour law. Morantz, *supra* note 10, at 520–22.

156 This is the experience, for example, of the Maintenance Cooperative Trust Fund, a Taft-Hartley trust fund created by unionized janitorial firms as a watchdog group to enforce labor standards in the industry to deter non-union firms from undercutting them by violating labor standards. Fine & Gordon, *supra* note 4, at 565–66 & n.97.

157 See generally ESTLUND, *supra* note 7, at 65.

the private litigation between CPA and Yank Sing Restaurant, and Seattle's creation of the Labor Standards Advisory Commission, for example, created incentives for employer-led PIOs to negotiate with worker-led PIOs for firm-level and sectoral standards.¹⁵⁸ CPA's joint negotiations with Yank Sing and public agencies enabled Yank Sing employees to participate in public agency enforcement and to monitor compliance with the settlement agreement. Public agency protection of these groups from retaliation can be a critical, often-missing element of worker engagement with public enforcement. Public agency protection of complainants from retaliation in the fast food and car wash remedial enforcement campaigns permitted these workers to speak openly about the problems of precarious employment to public agencies and to the media after the resolution of their cases.

To be sure, the examples of collaborative enforcement offered in this Part could be improved.¹⁵⁹ The New York State wage board only proposed lifting the state minimum wage, leaving unaddressed the question of how to enforce this new, higher standard. The San Francisco approach of tying collaboration to funding streams could, over time, erode public agency and PIO independence, creating disincentives to hold the other publicly accountable. The New York City carwash bonding opt-out provision for unionized establishments, at least for now, has been enjoined as preempted under the NLRA. These limitations and risks suggest the need for additional constraints on the private delegation and legislative aims of collaborative enforcement that will be elaborated in the next Part.

C. Proposal: Collaborative Enforcement to Amplify the Deterrent Effect of Enforcement

The previous sections have shown how collaborative

158 See AYRES & BRAITHWAITE, *supra* note 5, at 55. The New York Car Wash Association, for example, provides members with trainings on OSHA compliance and "a legal code and compliance program" for an attorney to review and correct members' payroll practices. Steve Rotlevi, *Mission Statement*, ASSOCIATION OF CAR WASH OWNERS, <https://nyccarwashassoc.herokuapp.com/about/> (last visited Oct. 9, 2017).

159 The state legislative backlash since 2015, in which employer-backed political campaigns have successfully lobbied many states to preempt local mandatory paid sick leave and minimum wage laws, offers another important limitation to local labor lawmaking. See generally Lori Riverstone-Newell, *The Rise of State Preemption Laws in Response to Local Policy Innovation*, 47 PUBLIUS: J. FEDERALISM 403 (2017).

enforcement can amplify the deterrent effect of PIO and public agency enforcement by integrating enforcement tools and improving legal remedies, improve the political accountability of public enforcement and facilitate tripartite lawmaking and case resolution. By analyzing these two approaches individually, these sections have also shown their distinct benefits. Grant-based enforcement is effective in governing ongoing collaborations, while remedial enforcement is more appropriate for intensive, short-term collaborations. Grant-based enforcement permits the legislature and public agency to direct enforcement into areas that the statute is intended to address, whether in an underregulated sector that public enforcement alone cannot reach, or to elaborate the requirements of a new labor standard. Remedial enforcement can stimulate the creation of new remedies if existing enforcement tools are inadequate.

Remedial and grant-based enforcement also have different PIO stakeholders. Unions are vital participants in remedial enforcement, while non-union PIOs are ideally suited for grant-based enforcement. Unions have newly repositioned themselves as lead PIOs in political campaigns for regional and sectoral labor standards in addition to seeking unionization in low-wage firms, top-pyramid tools available to unions but not to non-union PIOs. In contrast, non-union PIOs are well positioned to provide sophistication to public enforcement because of their superior access to information about systemic violations in underregulated sectors of the economy. They also often lack top-pyramid tools to address widespread noncompliance without access to public agency remedies.¹⁶⁰ Lacking the membership dues structure of unions, and often reliant on government funding,¹⁶¹ non-union PIOs are better equipped to sustain this role with public funding in a grant-based enforcement model. And, unlike unions, non-union PIOs are not subject to NLRA regulation, which can constrain unions' ability to protest employers because of labor standards violations.

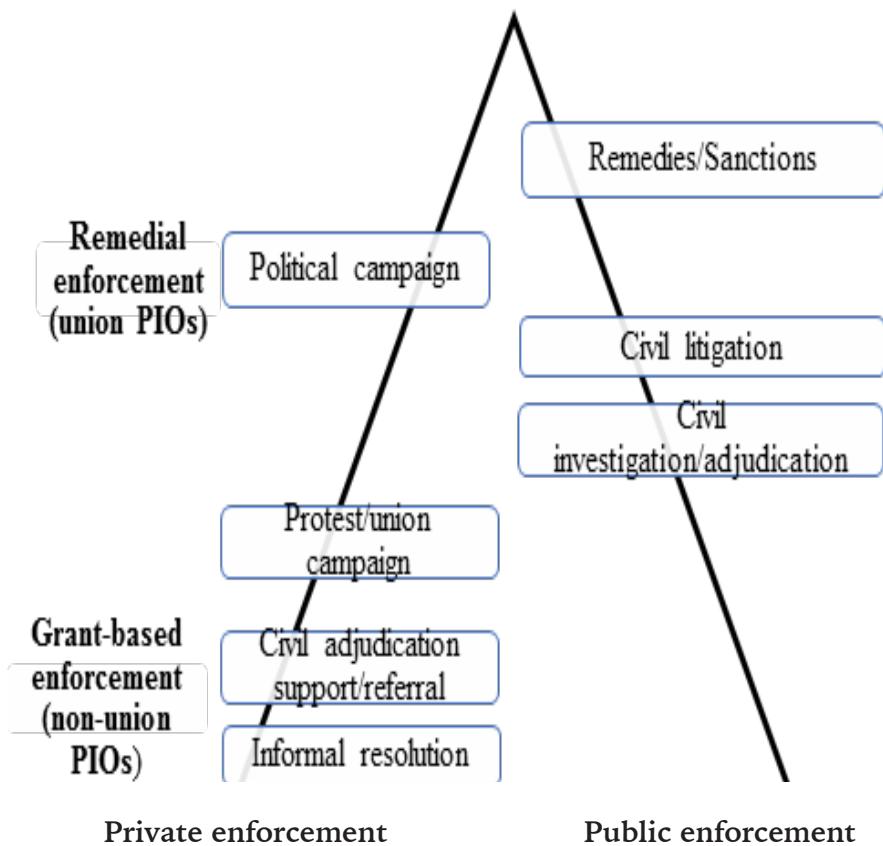
Even as remedial and grant-based enforcement stand apart based on their goals and composition, they are complementary and mutually reinforcing. As shown in Part II, underenforcement is often a result of pathologies that extend beyond the public-private divide. While grant-based enforcement primarily addresses capacity and

160 DE GRAAUW, *supra* note 13, at 51 (finding that only two of 100 studied community-based organizations serving immigrants in San Francisco had the resources to fund political advocacy).

161 *Id.* at 14; Gates et al., *supra* note 13.

expertise gaps, remedial enforcement is often necessary to change the structure of the law where existing remedies are inadequate. An integrated remedial and grant-based enforcement approach may be necessary to bridge all three.

Drawing from these examples, a combined approach appears as follows:



In a combined approach, public agencies collaborate with union-led PIOs to combine public and private enforcement tools to change the structure of the law where remedies and sanctions are inadequate. As those standards must then be enforced, public agencies can then rely on non-union PIOs in a grant-based enforcement technique. In both phases, enforcement is centered on sectors and regions characterized by underenforcement of labor standards. In it, the first goal would be to strengthen labor standards in those sectors and regions through civil enforcement and legislation. Once this goal is met, the enforcement priority would shift to channeling ongoing

enforcement resources to those sectors and regions to establish a norm of compliance. In this model, lawmaking achieved through remedial enforcement would task the public agency enforcing the new labor standard with delegating education, triage and referral to non-union PIOs.

Unlike other proposals for public-private collaboration in labor standards enforcement, which seek to encourage self-regulation backed by PIO oversight with the threat of state enforcement for high-risk employers,¹⁶² collaborative enforcement is an available strategy in sectors in which enforced self-regulation is not a realistic near-term strategy. The history of tripartism in these sectors is that public enforcement often lacks the enforcement tools to counteract entrenched norms of noncompliance. Instead, the goal of collaborative enforcement is to amplify the deterrent effect of public and private enforcement. Deploying PIO and public agency resources to develop, and coordinate the use of, remedies and sanctions in an integrated enforcement pyramid is more likely to impact employer behavior than public agency enforcement alone. These collaborative enforcement techniques, therefore, flip the tripartism focus to high-risk employers with the ultimate goal of deterring legal violations sufficiently to make enforced self-regulation possible.

In this framework, grant-based enforcement would restrict funding eligibility to non-union PIOs. Separating a non-union PIO's grant-based enforcement activities from direct participation in unionization campaigns in this manner would protect the legal regime from political attack as a subterfuge for government-funded unionization and non-union PIOs from attack as a "labor organization" subject to regulation by the NLRA.¹⁶³ In all examples,

162 For example, the "decentered proposal" offered by David Doorey proposes a two-track enforcement approach, a low-risk stream for employers willing to negotiate a code of conduct with worker representatives (not unlike the greengrocer code of conduct), and a high-risk stream for other employers, subject to heightened state enforcement. David J. Doorey, *A Model of Responsive Workplace Law*, 50 OSGOODE HALL L.J. 47, 77 (2012). This is similar to Gunningham's and Grabosky's proposed "two-track" regulatory system to regulate pollution by offering regulatory flexibility for companies that meet a high environmental standard and traditional regulation for firms that do not. See GUNNINGHAM & GRABOSKY, *supra* note 5, at 241–46.

163 29 U.S.C. § 152(5) (2012). The NLRB's Division of Advice has found that a PIO worker center that routinely negotiates with employers on behalf of employees to resolve wage-and-hour law violations is not a labor organization because the settlement agreements reached with employers "were discrete, non-recurring transactions with each [e]mployer." Michael C. Duff, *Alt-Labor*,

remedial enforcement would primarily address gaps in sanctions regimes by creating new sanctions and prophylactic remedies, such as bond requirements for restitution for victims of noncompliance and license revocation authority.

There are limitations to this proposal. Grant-based enforcement requires a public funding stream (or a PIO able to sustain collaboration without funding). Remedial enforcement requires PIOs capable of lobbying state or local governments. In jurisdictions lacking these elements, the traditional PIO roles of private enforcers and holding public agencies accountable for public enforcement will remain necessary.

The next Part addresses the critique that delegating public enforcement to PIOs undermines the independence of public agencies and PIOs and evaluates the vulnerability of collaborative enforcement to challenges under the NLRA and nondelegation doctrine. It responds that the threat to public agency and PIO independence can be mitigated by political and administrative controls. NLRA preemption and the nondelegation doctrine do not threaten the core requirements of collaborative enforcement, but do suggest outer boundaries for the scope of private delegation and its legislative aims.

IV. The Risk of Eroding Public and Private Independence and Doctrinal Limitations to Collaborative Enforcement

Proponents of formal collaboration between public agencies and PIOs have not yet addressed important critiques that public delegation to private entities may undermine public agency independence from PIOs or PIO independence as monitors of public agency enforcement. This Part will address these concerns as well as doctrinal limitations, namely the threats of preemption under the NLRA to the lawmaking aims of remedial enforcement and of

Secondary Boycotts, and Toward a Labor Organization Bargain, 63 CATH. U. L. REV. 837, 855–56 & n.121–24 (2014) (quoting NLRB, Office of the Gen. Counsel, Advice Memorandum, Cases 2-CP-1067, et al. (Nov. 30, 2006)). But as an agency determination subject to change, the line separating a worker center’s efforts to improve the working conditions of low-wage workers and those of a labor organization subject to NLRA regulation remains unclear and unstable. See Ben Penn and Tyrone Richardson, *Labor Department Looking Into Worker Center Scrutiny: Acosta*, BLOOMBERG (Nov. 15, 2017), <https://www.bna.com/labor-department-looking-n73014472124/> (reporting that Labor Secretary Acosta is currently reviewing the legal status of worker centers under the Labor-Management Reporting and Disclosure Act).

state nondelegation doctrine to tripartite rulemaking. This Part will conclude that while these normative concerns and doctrinal limitations do not threaten the core requirements of collaborative enforcement, they do suggest the need for controls to ensure that coordination does not erode public or private enforcer independence, and for limiting the scope of its private delegation and legislative aims to avoid litigation challenges.

A. *The Erosion of Independent Enforcement*

Legal scholars, noting the unprecedented scale and scope of privatization, have raised important objections to the contracting out of public services to private entities.¹⁶⁴ This scholarship often focuses on the ways that privatization can erode accountability through the delegation of public services to contractors with inadequate oversight, and by placing private contractors beyond the reach of public agency statutory, administrative and political checks.¹⁶⁵ Contracting out, with its sole focus on efficiency, may obscure democratic or other important values expressed in a public function.¹⁶⁶ Jon Michaels raises the separate danger that public agencies may contract out to private entities as a workaround to bypass political, statutory or administrative controls in order to aggrandize agency power.¹⁶⁷

Not all of these concerns are present in collaborative enforcement. Collaborative enforcement's limited delegation permits public agencies to retain their discretion to make value-laden, contestable decisions about priorities and case decisions, and limits the private enforcer role to the vindication of private interests.¹⁶⁸

164 FREEMAN & MINOW, *supra* note 8, at 3–6.

165 See, e.g., Martha Minow, *Outsourcing Power: Privatizing Military Efforts and the Risks to Accountability, Professionalism, and Democracy*, in *id.*, 110–147. Margaret Lemos, for instance, argues that the lack of accountability in the deputization of private attorneys to enforce on behalf of the state can lead to private attorney fraud and abuse. See Lemos, *supra* note 9, at 529, 578–82.

166 Sharon Dolovich, *How Privatization Thinks: The Case of Prisons*, in FREEMAN & MINOW, *supra* note 8, at 134.

167 Jon D. Michaels, *Privatization's Pretensions*, 77 U. CHI. L. REV. 717, 719 (2010) (In an administrative workaround, public agencies may turn to privatization to achieve “distinct public policy goals that – but for the pretext of technocratic outsourcing – would be impossible or much more difficult to attain in the ordinary course of nonprivatized public administration.”). See also Jon D. Michaels, *CONSTITUTIONAL COUP: PRIVATIZATION'S THREAT TO THE AMERICAN REPUBLIC* 81 (2017) (hereinafter “*CONSTITUTIONAL COUP*”).

168 The delegation in grant-based enforcement is more modest than in deputation programs, in which local agencies deputize union representatives

Its deeply intertwined nature subjects PIOs to meaningful public agency oversight.¹⁶⁹ Restricting grant-based enforcement to not-for-profit organizations that would suffer a reputational harm from fraud or abuse is another control against abuse.¹⁷⁰ Private delegations to PIOs do not necessarily undermine democratic values; on the contrary, collaborative enforcement can serve the democratic value of channeling worker voice into workplace regulation.

There is, however, the risk that even modest forms of delegation can undermine the independence of public agencies and PIOs. The funding of PIO enforcement may create incentives for PIOs to misuse collaboration by exploiting principal-agent asymmetries and by seeking to substitute instead of complement public agency enforcement. Grant-based enforcement may also create incentives for public agencies to use grants as a form of political patronage, eroding PIO independence.

This section will first address the concern that grant-based enforcement will create incentives for PIOs to abuse information asymmetries and for PIO enforcement to substitute for rather than complement public agency enforcement, and then turn to the

to conduct inspections in public work construction sites and to assist the procurement agency with audits, hearings and review conferences for the enforcement of prevailing wage laws. Fine & Gordon, *supra* note 4, at 563–65. While deputization provides the technocratic benefit of “formal power” to deputies that PIOs often lack in worksite regulation, *id.* at 565, it also creates incentives for PIOs to abuse the delegation by exercising public enforcement “with less notice, resistance, or legal consequence than if they were actually to join the governmental ranks or otherwise shed their private personas.” Jon D. Michaels, *Deputizing Homeland Security*, 88 TEX. L. REV. 1435, 1452 (2010).

- 169 Lemos, *supra* note 9, at 530 (arguing that abuse of enforcement delegations can be controlled where public enforcers “remain[] in the background, capable of filling in where private efforts fall short”). Engstrom’s empirical analysis of private litigants who enforce public law as whistleblowers, for instance, found no support for the claim that privatizing public enforcement claims leads to abuse. Engstrom, *supra* note 79, at 1963 (“In sum, the composite evidence points decisively away from widespread claims that qui tam enforcement efforts are in the midst of an inefficient ‘explosion.’”).
- 170 The identity of PIO enforcers is often shaped around assisting the communities most deeply impacted by enforcement gaps, which would be placed at risk by fraud or abuse of collaborative enforcement. Also, unlike for-profit entities, which have an incentive to exploit information asymmetries by charging unreasonably high rates or providing poor services, not-for-profit corporations have no “owner,” but are instead controlled by managers, paid in salary drawn from funders, who would not directly profit from abuse. See HENRY HANSMANN, THE OWNERSHIP OF ENTERPRISE, 227–30, 233–37 (1996).

risk that public agencies will misuse grants to silence PIOs that are potential watchdogs of public enforcement. The section will conclude that political and administrative controls will be necessary to preserve PIO and public agency independence.

1. Public Enforcement Independence

There are two primary ways that collaborative enforcement may erode public agency independence. The first is that PIOs in grant-based enforcement may misuse private delegations by exploiting information asymmetries. The second is that PIOs may seek funding for private enforcement that will replace public agency functions instead of contributing to them.

First, PIOs may seek to exploit public agency enforcement resources in unproductive ways.¹⁷¹ In grant-based enforcement, PIOs could triage and refer cases outside the agency's priorities but which may serve PIO recruitment or other goals.¹⁷² Delegation may also invite abuse for difficult-to-monitor functions.¹⁷³ It may be difficult, for example, for public agencies to monitor grants to PIOs to educate immigrant communities about labor standards in languages that public agency personnel cannot speak. While public agencies could control for these abuses by selecting only PIOs that would incur reputational harm for misuse of public funds, there may be few PIOs in a jurisdiction to choose from, and the public agency enforcer may be ill-equipped to identify which PIOs bear the most reputational risk, and may instead prefer PIOs without a strong reputation because of their docility.

In short, the intertwined relationship between PIOs and public agencies in grant-based enforcement suggests the need for political and administrative controls in the grant design to prevent

171 See Engstrom, *supra* note 79, at 2000–01 (In some instances, private attorneys have used the False Claims Act to exploit regulatory ambiguities rather than to reveal enforcement gaps.). See generally FREEMAN & MINOW, *supra* note 8, at 2–6 (2009); Michaels, *supra* note 167, at 718.

172 Matthew Stephenson and Howell Jackson have noted similar principal-agent problems in lobbyist contributions to public policy. See Matthew C. Stephenson and Howell E. Jackson, *Lobbyists as Imperfect Agents: Implications for Public Policy in a Pluralist System*, 47 HARV. J. LEGIS. 1 (2010).

173 Government contracting of services, such as education and health care, is often difficult for public agencies to monitor and evaluate because it involves complex tasks and the direct beneficiaries have no direct relationship with the public agency funding the services. HANSMANN, *supra* note 170, at 227–30, 233–37.

abuse in the PIO selection, training, monitoring and evaluation process.¹⁷⁴ Grant-based enforcement often comes at the direction of the local legislature,¹⁷⁵ creating the possibility of a formal legislative review process,¹⁷⁶ and legislative delegation of grant selection and monitoring to third-party government agencies and PIOs¹⁷⁷ to check against abuse.

Transparency is an important, additional control against abuse of grant-based enforcement. A legislative requirement that public agencies disclose their grant-based enforcement priorities and benchmarks prior to PIO selection will improve the screening of PIOs during the selection process by limiting consideration of PIOs to those that can meet pre-set selection criteria. It would also empower watchdogs and the regulated entities to identify and mobilize against perceived abuse in grant-based enforcement.¹⁷⁸

174 See Laura A. Dickinson, *Public Values/Private Contract*, in GOVERNMENT BY CONTRACT 336 (arguing that “the very government contracting that is the engine of privatization itself opens the space for an intriguing set of accountability mechanisms”).

175 See, e.g., Berkeley, Berkeley, Cal., Mun. Code Ch. 13.99, § 13.99.080 (2016) (“The Department of Finance shall seek out partnerships with community-based organizations and collaborate with the Labor Commission to facilitate effective implementation and enforcement.”); Los Angeles Cnty., Cal., Ord. No. 102703 amending County Code, Title 8 - Consumer Protection, Business and Wage Regulations relating to the enforcement of the County Minimum Wage Ordinance, § 8.101.090(G) (Apr. 26, 2016). (“The DCBA shall have the authority to contract, in accordance with County contracting rules and procedures, with Community Based Organizations for them to assist in the education and outreach related to the Los Angeles County Minimum Wage Ordinance and this Chapter.”); San Francisco, S.F., Cal., Ord. No. 140687, Amending S.F. Mun. Code § 12R.25 (July 17, 2014) (“The Office of Labor Standards Enforcement shall establish a community-based outreach program to conduct education and outreach to employees. In partnership with organizations involved in the community-based outreach program, the Office of Labor Standards shall create outreach materials that are designed for workers in particular industries.”).

176 Michaels, *supra* note 167, at 769–70. If the legislature identifies abuses in the grant-based enforcement it requires, “legislators can apply resistance in proportion to the perceived encroachment on their prerogatives.” *Id.* at 770.

177 As Ayres and Braithwaite argue, delegating monitoring to third-party PIOs can prevent abuse by making grant-seeking PIOs accountable to third parties that may replace the PIO in the grant-based relationship should the initial delegation fail. See AYRES & BRAITHWAITE, *supra* note 5, at 57 (arguing that contestable guardianship requires “a regulatory culture where information on regulatory deals is freely available to all individual members of a multitude of” PIOs).

178 Transparency may also improve training of public agency and PIO staff, by

Transparency may improve legislative review and ongoing screening and monitoring functions by permitting third-party monitoring, to evaluate PIO use of public funding against public agency priorities to detect funding misuse.

The second concern is that public funding of PIOs for services that could be provided by public agencies themselves can erode the public's view of the role of government in society. At least in some instances, this critique does not apply because collaboration entails the integration of private enforcement tools that are otherwise inaccessible to the public agency. In the case of remedial enforcement, the complementary nature of a PIO's top-tier enforcement tools—PIO public protests, unionization and political campaigns—is clear. But PIO collaboration is not *necessarily* complementary. In other instances, it is at least plausible that public enforcement regimes could be retooled to address (or avoid) a capacity or expertise deficit. A public agency, for example, could encourage complaints through an agency-run mediation program instead of by providing grants to PIOs to privately resolve them. In this instance, it may be that a PIO-delegated informal resolution is preferable to public agency mediation. A PIO may have a unique relationship with a particular worker community, its staff may have cultural and linguistic competencies that public agency staff lack, or a PIO may be better positioned to allay worker fears about retaliation than staff in public agencies, at least at the outset.¹⁷⁹ Or, it could be that the public agency would be better served by building internal expertise in cultural and linguistic competencies, and that an internal mediation program would permit public agencies to select targets for public enforcement while promoting civic engagement in public enforcement without PIO assistance.

The question of whether grant-based enforcement is complementary or substitutive, therefore, is necessarily contingent on the enforcement tool, the worker community and the PIO. This suggests an additional need for transparency about the complementary service the public agency seeks through collaboration. Requiring a public agency justification for delegation of responsibilities could restrict grants to services that the public agency could not provide

guiding personnel across the public-private divide about shared priorities. Seattle and San Francisco, for example, require PIO staff to undergo training with public agency staff about the information that the PIOs is charged with providing to the public. Dickenson, *supra* note 177, at 430-41.

179 DE GRAAUW, *supra* note 13, at 9.

itself and that would generate additional benefits if provided by the PIO, such as civic integration of groups that do not normally participate in the political process. The legislature could require public agencies seeking collaboration with PIOs to condition funds on the transfer of expertise to public agency staff, ultimately improving the provision of public services to underrepresented groups.

2. Harm to Private Enforcement

There is also the possibility, underdiscussed in the literature, that public delegations could also harm *private* enforcement by undermining PIO independence as monitors of public agency enforcement.¹⁸⁰ Even without imposing legal restrictions,¹⁸¹ a public agency could seek to use the benefits of collaboration to PIOs in order to co-opt them. In essence, public agencies may use grant-based enforcement as a workaround, but instead of evading statutory or administrative controls,¹⁸² the public agency may use grants to evade public accountability by chilling the speech of PIOs that have traditionally served as watchdogs for public agency enforcement. Public agencies could restrict funded collaboration to those PIOs that agree to advocate that the legislature expand the public agency's budgets or authority, or those PIOs that refrain from criticizing the public agency.¹⁸³ Or the public agency could

180 Whether government funding decreases the political activity of not-for-profit organizations has been the subject of sociological research, with the ambiguous conclusion that government funding creates incentives in both directions and has no clear net effect. See Chaves et al., *Does Government Funding Suppress Nonprofits' Political Activity?*, 69 AM. SOC. REV. 292, 313–15 (2004).

181 Some grant-based enforcement may impose broad lobbying restrictions on participating PIOs that suppress their role as monitor of public enforcement. New York City, for example, prohibits the use of discretionary funds for any form of "lobbying." See NEW YORK CITY COUNCIL, DISCRETIONARY FUNDING POLICIES AND PROCEDURES 7, <http://council.nyc.gov/budget/wpcontent/uploads/sites/54/2017/01/PoliciesProceduresJan2017.pdf> ("Funds may only be allocated for a public purpose and may not support political activities (including but not limited to lobbying, campaigns or endorsements) and/or private interests.").

182 Michaels points to ways that outsourcing may aggrandize agency executives by permitting them to, for example, bypass statutory privacy protections by contracting data mining operations to private contractors, and sideline politically insulated civil servants who may oppose the executive's priorities by outsourcing research and regulatory drafting responsibilities to private experts. Michaels, *supra* note 167, at 721–22.

183 See, e.g., DE GRAAUW, *supra* note 13, at 14 (finding from interviews with PIOs in San Francisco, that "[w]hen immigrant-serving nonprofits advocated too

seek collaboration to shift responsibility of the priority away from itself, potentially also shifting blame to the PIO if an enforcement effort fails. Some of these risks are mitigated by the nature of PIOs as non-profit organizations that serve communities that can hold PIOs accountable for cooptation by the public agency. PIOs may be less susceptible to adverse treatment by public agencies if they are equally or more deterred by the possibility of backlash from their membership or base. But, and particularly for PIOs insulated from backlash from the communities that they serve, cooptation is a genuine threat to their independence.

Political and administrative controls will be equally important to protect PIOs from public agency cooptation and unfair blame. As with protecting the independence of public agencies, transparency about the public agency's policy choices in the grant-making process would also enable the legislature to control against abuse. The risk of cooptation can also be reduced by assigning the grant-monitoring role to a third-party PIO and public agency. There is an additional need for explicit whistleblower protections in grant contracts sufficient to encourage PIOs to complain about public agency abuse.

In conclusion, while abuse is possible in any principal-agent relationship, grant-based enforcement is less susceptible to abuse because of its limited delegation, intertwined nature, and political accountability. Political control over the grant selection and evaluation in grant-based enforcement, transparency about the public agency's enforcement priorities, and administrative controls, such as third-party selection and monitoring of PIOs and whistleblower protections, can further limit the potential of abuse.

Of course, these political and administrative controls may supply an appearance of public agency and PIO independence without providing substantive protections.¹⁸⁴ The legislature may lack the interest or sophistication to identify fraud or abuse, and third-party PIOs and government agencies may find that objective evaluation of PIOs that seek grants is elusive. Abuses hidden in public agency disclosures may be difficult to unmask. To the extent that the

aggressively, they worried that city officials might react by endangering their tax status or government funding").

184 See Adrian Vermeule, *Our Schmittian Administrative Law*, 122 HARV. L. REV. 1095, 1096 (2009) (quoting DAVID DYSENHAUS, THE CONSTITUTION OF LAW: LEGALITY IN A TIME OF EMERGENCY 3 (2006)) (cautioning that many formal controls on administrative agencies can operate as mere façades that agencies can easily bypass).

funding of PIOs in grant-based enforcement may not overcome the risks of abuse, state and local governments could coordinate with PIOs without funding, and seek other ways to fund PIO enforcement outside of the collaborative relationship.¹⁸⁵

B. NLRA Preemption and State Nondellegation

There are only two legal doctrines that might prevent state or local agencies from collaborating with PIOs in enforcing labor standards. The most important of these is the NLRA, which preempts state and local laws that condition a public benefit on acceptance of union neutrality. The second, state nondellegation doctrine, constrains legislative grants to public agencies and private entities.¹⁸⁶

Other legal doctrines, such as other constitutional separation of powers principles and freedom of information laws, do not significantly restrict collaborative enforcement. The exercise of enforcement discretion is presumptively valid, and the selection and resolution of cases based in part on a PIO's private enforcement does not nullify a law or violate an express statutory command.¹⁸⁷ Any delegation entailed in grant-based enforcement is either expressly or impliedly authorized by statute, presenting no separation of powers problem either.¹⁸⁸ Nor do state and local freedom of information laws significantly constrain collaborative enforcement. A public agency generally need not disclose information about a pending

185 New York City has recently experimented, for example, with allowing fast food employees to self-fund PIO enforcement through voluntary deductions from paychecks, *see* N.Y.C. COUNCIL Intro. Bill No. 1384 (2016), and where unions have successfully unionized firms in remedial enforcement initiatives, they may seek to create employer-employee funded Taft Hartley funds to enforce labor standards among their competitors. *See* Elmore & Chishti, *supra* note 28, at 14 (describing formation of Taft Hartley fund in Illinois to police prevailing wage bids in the construction industry); ESTLUND, *supra* note 7, at 117–22 (describing formation of Maintenance Cooperative Trust Fund in California following the “Justice for Janitors” union campaign, “to identify and challenge labor standards violations among janitorial contractors”).

186 *Whitman v. Am. Trucking Ass’ns* 531 U.S. 457, 472 (2001) (“In a delegation challenge, the constitutional question is whether the statute has delegated legislative power to the agency.”).

187 *Heckler v. Chaney*, 470 U.S. 821, 831–35, 837–38 (1985) (holding that a public agency’s nonenforcement decision is presumptively nonreviewable).

188 *See Youngstown Sheet & Tube Co. v. Sawyer*, 343 U.S. 579, 635–36 (1952) (Jackson, J., concurring) (“When the President acts pursuant to an express or implied authorization of Congress, his authority is at its maximum.”).

investigation or litigation,¹⁸⁹ and open meetings laws would merely require the types of transparency in grant-based enforcement that San Francisco and Seattle already provide.

1. *NLRA Preemption*

While the NLRA contains no express preemption provision, the Supreme Court has found two “unquestionably and remarkably broad”¹⁹⁰ variations of implied preemption under the NLRA.¹⁹¹ Under the first, *Garmon* preemption, the NLRA preempts states and localities from regulating conduct that is “arguably” prohibited or protected by the NLRA.¹⁹² The second, *Machinists* preemption, prohibits any state and local regulation of union conduct that Congress intended to leave unregulated.¹⁹³

Most lawmaking sought by remedial enforcement identified in this Article does not implicate *Garmon* or *Machinists* preemption because they are laws of general applicability, which affects all employees equally, and “neither encourages nor discourages the collective-bargaining process.”¹⁹⁴ State and local lawmaking to establish wage-and-hour law standards above the federal minimum is expressly permitted by FLSA¹⁹⁵ and grant-based enforcement channels enforcement to all employers in the informal economy. Tripartite lawmaking of the type that resulted in the New York wage board proposal to increase fast food worker wages in the state is also a law of general applicability and does not constitute the kind of collective bargaining regulated by the NLRA.¹⁹⁶ Increasing local

189 See, e.g., N.Y. Pub. Off. Law § 87(2) (2015) (discussing interference with a law enforcement investigation exception).

190 Benjamin I. Sachs, *Despite Preemption: Making Labor Law in Cities and States*, 124 HARV. L. REV. 1153, 1164 (2011).

191 *Id.* at 1164–69.

192 *San Diego Bldg. Trades Council v. Garmon*, 359 U.S. 236, 259–60 (1959).

193 *Int'l Ass'n of Machinists v. Wis. Emp't Rel. Comm'n*, 427 U.S. 132 (1976); see also *Local 20, Teamsters v. Morton* 377 U.S. 252, 259–60 (1964) (State law that prohibits secondary boycotts permitted by the NLRA preempted; boycotts are a self-help weapon permitted under federal law that states cannot regulate.). Labor scholars criticize the breadth of these preemption doctrines for stifling state and local experimentation with labor law. See Estlund, *supra* note 56, at 1572; see also Sachs, *supra* note 190, at 1168–69 (describing the history and evolution of NLRA preemption doctrine).

194 *Metro. Life Ins. Co. v. Massachusetts*, 471 U.S. 724, 755 (1985).

195 29 U.S.C. § 218(a) (2012).

196 See Andrias, *supra* note 6, at 91–92 (arguing that finding this type of tripartite lawmaking to be preempted “would require a significant expansion of preemption law”).

minimum wages and establishing uniformly-applied remedies to operationalize these rights, such as licenses and labor bonds to operate a car wash, face no plausible NLRA preemption challenge.¹⁹⁷

Lawmaking that permits union workplaces to “opt-out” of remedies that apply to non-union workplaces, however, faces the plausible argument that *Machinists* preemption applies to the extent that they coerce employers to adopt a position of union neutrality. In New York City, following the passage of a local law requiring a labor bond and registration for car wash establishments, subject to an opt-out for unionized car washes, the local car wash association sued, claiming that the NLRA preempts the law.¹⁹⁸ The trial judge in that case, *Association of Car Wash Owners v. City of New York*,¹⁹⁹ agreed, finding that the opt-out “explicitly encourages unionization, and therefore impermissibly intrudes on the labor-management bargaining process” and is preempted by the NLRA.²⁰⁰ The case is now on appeal.²⁰¹

While the ultimate outcome of *Association of Car Wash Owners* remains uncertain, the trial court’s decision shows that NLRA preemption is a plausible threat to opt-outs, particularly if the opt-out imposes onerous requirements on non-union firms or if legislators seek an opt-out to advance union campaigns. The two Supreme Court cases that set the limits on state and local regulations that impact labor relations are *Golden State Transit Corporation v. Los Angeles*²⁰² and *Chamber of Commerce v. Brown*.²⁰³ In *Golden State Transit Corp.*, the Supreme Court held that the state’s conditioning of the renewal of a taxicab franchise on the settlement of a strike was preempted because it “destroyed the balance of power designed by Congress, and frustrated Congress’ decision to leave open the use of economic weapons.”²⁰⁴ In *Brown*, the Supreme Court held that California could

197 See *Am. Hotel and Lodging Ass’n v. City of Los Angeles*, 834 F.3d 958, 965–66 (9th Cir. 2016). (“We have consistently held that minimum labor standards do not implicate *Machinists* preemption.”).

198 *Ass’n of Car Wash Owners Inc. v. City of N.Y.*, 15-CV-08157 (AKH), Pl’s Mem. Supp. S.J. Mot. (S.D.N.Y. Oct. 16, 2015).

199 *Ass’n of Car Wash Owners Inc. v. City of N.Y.*, 15-CV-08157 (S.D.N.Y. Oct. 16, 2015).

200 *Ass’n of Car Wash Owners Inc. v. City of N.Y.*, 15-CV-08157 (AKH), Order Granting Pl’s S.J. Mot. (S.D.N.Y. May 26, 2017).

201 *Ass’n of Car Wash Owners Inc. v. City of N.Y.*, 15-CV-08157 (AKH), Notice of Appeal, (S.D.N.Y. June 12, 2017).

202 475 U.S. 608 (1986).

203 554 U.S. 60 (2008).

204 *Golden State Transit Corp.*, 475 U.S. at 619.

not condition public contracts on an employer's position of union neutrality.²⁰⁵ It found that Congress had "renounced" the state's policy judgment that partisan employer speech interferes with an employee's choice about whether to be represented by a labor union in the Taft-Hartley Act, and thus conflicted with federal policy.²⁰⁶ In both cases, lawmaking that seeks to coerce an employer to accept union neutrality, even indirectly through government licensing or procurement powers, is preempted under *Machinists*.²⁰⁷

A broad interpretation of *Association of Car Wash Owners* that would displace state laws that contain opt-outs for employers subject to a collective bargaining agreement because they "explicitly encourage[]" and "pressur[e] businesses to unionize"²⁰⁸ would require an extension to existing *Machinists* doctrine. To be sure, the NLRA displaces state laws that substantially interfere with a non-union firm's ability to operate unless it agrees to a collective bargaining agreement.²⁰⁹ But *Golden Gate* and *Brown* do not suggest that state and local labor standards that provide narrow opt-outs for employers subject to collective bargaining agreements are subject to NLRA preemption under *Machinists*, and courts have repeatedly

205 *Brown*, 554 U.S. at 69.

206 *Id.*

207 Project labor agreements are typically not preempted under the market participant doctrine because they advance only the proprietary, and not regulatory, interest of the state. *See Bldg. & Constr. Trades Council v. Assoc. Builders & Contractors of Massachusetts/Rhode Island*, 507 U.S. 218, 223, 231–32 (1993); *Michigan Bldg. and Constr. Trades Council v. Snyder*, 729 F.3d 572, 581–82 (6th Cir. 2013). But state and local governments cannot bypass NLRA preemption review by formally grounding its power to condition union neutrality in its procurement or tax authority. *See Brown*, 554 U.S. at 70 (rejecting the argument that using a procurement power to require union neutrality fell within the market participant exception to NLRA preemption); *Assoc. Builders and Contractors Inc. v. City of Jersey City*, 836 F.3d 412, 417–21 (3d Cir. 2016) (explaining that a city that conditions tax exemptions on union neutrality acts as a regulator rather than a market participant).

208 *Ass'n of Car Wash Owners Inc. v. City of N.Y.*, 15-CV-08157 (AKH), Order Granting Pl's S.J. Mot. at 6 (finding that "a fivefold increase in the amount of a surety bond required for car washing companies that are not parties to a collective bargaining agreement, or, alternatively, an independent monitoring scheme and large security deposits," amounts to a penalty on non-union car washes).

209 *See Golden State Transit Corp.*, 475 U.S. at 619; *Chamber of Commerce v. Bragdon*, 64 F.3d 497, 501 (9th Cir. 1995) (requiring prevailing wage terms in private contract *Machinists* preempted because terms so onerous it dictated collective bargaining process).

upheld them against NLRA preemption challenges.²¹⁰ States have “broad authority under their police powers to regulate the employment relationship,”²¹¹ even if they alter the economic balance between labor and management.²¹² The NLRA “cast[s] no shadow on the validity”²¹³ of an opt-out provision, even if it “provided an incentive to unionize or to remain non-union” and may have a “potential benefit or burden in application.”²¹⁴ The Ninth Circuit in *American Hotel and Lodging Association*,²¹⁵ for example, recently upheld a city ordinance’s waiver for collective bargaining in a minimum wage ordinance against a *Machinists* preemption challenge. The modest opt-out provision in *American Hotel* did not approach the overreaching enforcement action of *Golden State*, in which the state agency sought to use state regulation to intervene in a labor strike, or the state command of union neutrality in *Brown*, which went well beyond protecting state funds.²¹⁶ For the Ninth Circuit, unlike state action that “intrudes on the mechanics of collective bargaining,” like those found preempted in *Golden State* and *Brown*, opt-outs merely establish a labor standard that sets the stage for future bargaining, which is not preempted.²¹⁷ Thus, while a wage bond opt-out almost certainly creates a cost for non-union employers, the court appears to have ignored the second step of the analysis, to determine whether the cost sufficiently interferes with collective bargaining.²¹⁸ This is

210 See *Am. Hotel and Lodging Ass’n*, 834 at 965 (holding that union opt out provision in minimum wage standard not preempted by NLRA); *Viceroy Gold Corp. v. Aubry*, 75 F.3d 482, 489–90 (9th Cir. 1996) (holding that union carve out in state maximum hours legislation not preempted by NLRA); *Filo Foods, LLC v. City of SeaTac*, 183 Wash. 2d 770, 778, 796–97 (Wash. 2015) (upholding union waiver provision in SeaTac’s recent \$15 per hour minimum wage increase against *Machinists* preemption challenge).

211 *De Canas v. Bica*, 424 U.S. 351, 356 (1976).

212 *Fort Halifax Packing Co. v. Coyne*, 482 U.S. 1, 21 (1987); *Metro. Life Ins. Co.*, 471 U.S. at 754; see also *Concerned Home Care Providers, Inc. v. Cuomo*, 783 F.3d 77, 85 (2d Cir. 2015).

213 *Livadas v. Bradshaw*, 512 U.S. 107, 132 (1994).

214 *Viceroy Gold*, 75 F.3d at 490. See also *Livadas*, 512 U.S. at 132 n.26 (reasoning that it does not “seem plausible to suggest that Congress meant to pre-empt such opt-out laws, as ‘burdening’ the statutory right of employees not to join unions by denying nonrepresented employees the ‘benefit’ of being able to ‘contract out’ of such standards”).

215 834 F.3d 956, 965 (9th Cir. 2016).

216 *Id.* at 964–66.

217 *Id.* at 964.

218 *Castillo v. Toll Bros*, 197 Cal. App. 4th 1172, 1207 (Cal. Ct. App. 2011) (finding that to the extent that an opt-out imposed costs on firms that can

particularly the case, as the Supreme Court reasoned in *Fort Halifax*,²¹⁹ since opt-outs only apply to firms that have already bargained for a similar requirement.²²⁰

However, while a broad interpretation of *Association of Car Wash Owners* that would invalidate all opt-outs cannot be reconciled with established law, a narrower, more defensible interpretation is possible, that animating the court's decision is skepticism of the local law's purpose. The court, citing comments made by New York City councilmembers during the legislative hearing supportive of the PIOs' unionization campaign,²²¹ concluded that "a central purpose of [the wage bond bill] is to encourage unionization in the car wash industry."²²² This suggests the need for clarity in opt-out provisions

be reduced in collective bargaining, that impact on collective bargaining is "insufficiently substantial to trigger *Machinists* preemption"). Cf. *Rondout Elec., Inc. v. NYS Dept. of Labor*, 335 F.3d 162, 168 (2d Cir. 2003) ("While the [state] regulation may impose an additional cost on non-union employers . . . this increase does not affect the bargaining process that is the subject of the NLRA.").

- 219 *Fort Halifax Packing Co.*, 482 U.S. at 21–22 ("If a statute that permits *no* collective bargaining on a subject escapes NLRA pre-emption . . . surely one that permits such bargaining cannot be pre-empted."); *see also Livadas*, 512 U.S. 107 (1994) (distinguishing California policy not to enforce state law requiring immediate payment of wages due upon discharge for employees subject to collective bargaining agreements on this ground).
- 220 California's wage bond opt-outs only apply to collective bargaining agreements that obligate employers to pay particular wages and "An expeditious process to resolve disputes concerning nonpayment of wages." CAL. LAB. CODE § 2055(b)(4)(D) (West 2015). *See St. Thomas-St. John Hotel & Tourism Ass'n v. Government of U.S. Virgin Islands*, 218 F.3d 232, 245 (3d Cir. 2000) (upholding opt-out to minimum wage protection because it "does not force an employee to choose between collective bargaining and the protections of state law; rather, it protects all . . . employees, but gives employees the option of relinquishing the territorial statutory protections through the terms of a collective-bargaining agreement"); *Firestone v. Southern California Gas Co.*, 219 F.3d 1063, 1068 (9th Cir. 2000) (upholding California overtime opt-out provision because it "does not operate automatically to exempt virtually all union-represented employees from its coverage—it exempts only those who have bargained for an alternative overtime compensation scheme").
- 221 *Ass'n of Car Wash Owners Inc. v. New York*, 15-CV-08157 (AKH), Order Granting Pl's S.J. Mot., at 6 (S.D.N.Y. May 26, 2017).
- 222 *Id.* While it is unclear from the court's decision whether the councilmembers' statements were in reference to the opt-out provision, the decision correctly states that opt-out provisions must be motivated by a purpose other than to support or encourage unionization. *See Chamber of Commerce of the U.S. v. Brown*, 554 U.S. 60, 63 (2008) (quoting the statute's express legislative purpose to "prohibit an employer from using state funds and facilities for

that their intent is to permit private negotiation for remedies rather than to encourage unionization. Legislation permitting waiver of prophylactic remedies such as labor bonds could also clarify the low need for labor bonds in firms subject to collective bargaining agreements because they are less likely to violate employment laws than non-union firms. To mitigate the preemption threat, the cost of compliance with a statute without a waiver should also be directly tied to the enforcement problem and should not be so onerous that union neutrality is the only meaningful choice for employers in the sector or region.

Alternatively, one might seek to minimize the preemption threat by modifying a statutory opt-out by expanding the exemption to not only employers with a collective bargaining agreement, but also those non-union employers who can adequately demonstrate compliance with labor standards. However, such an approach would multiply the administrative complexity of the law. It may also provide incentives for employers to use the exemption to skirt the requirements of the law. Given these countervailing risks, and that these waivers are not essential to collaborative enforcement, should preemption law expand to encompass remedial differences between unionized and non-union firms, it may be preferable to avoid them entirely.

2. State Nondelegation Doctrine

Collaborative enforcement must also account for the nondelegation doctrine, which constrains administrative rulemaking, not enforcement authority, which is an executive function.²²³ As a result, nondelegation principally constrains remedial enforcement's use of tripartite rulemaking. Federal nondelegation doctrine and that of most states impose few constraints on private delegations. Courts have not found a violation of the federal nondelegation doctrine since 1936, and the distinction between public agency and private

the purpose of influencing employees to support or oppose unionization") (quoting 2000 Cal. Stat. ch. 872 § 1). In *Brown*, the Supreme Court relied on the state statute's preamble to conclude that it sharply conflicted with federal policy under the NLRA to permit noncoercive employer speech opposing unionization. *Id.* at 68.

223 Courts routinely uphold delegations of executive power, including the power to arrest, to private individuals. See David M. Lawrence, *Private Exercise of Governmental Power*, 61 IND. L.J. 648, 666 (1986) ("The power to arrest has been delegated to railway police, to humane society agents, and to bail bondsmen.").

stakeholder delegations has not figured prominently in federal nondelegation cases.²²⁴ While courts have invoked the nondelegation doctrine to strike down private delegations at the state and local levels,²²⁵ in most states delegations to private parties are permissible so long as the agency retains the ultimate authority to approve and the agency provides sufficient safeguards to prevent abuse.²²⁶

For these states, two cases that consider claims that prevailing wage laws unconstitutionally delegate prevailing wage rate setting power to local employers and unions illustrate the nondelegation doctrine constraints on state and local delegations of rulemaking power to private parties. In *Beary Landscaping, Inc. v. Costigan*,²²⁷ the Seventh Circuit considered a challenge by construction contractors of the Illinois prevailing wage law that set the wage rate based on collective bargaining agreements (CBA) for construction work in the county on the ground this constitutes an unconstitutional

224 Alexander Volokh, *The New Private-Regulation Skepticism: Due Process, Nondelegation, and Antitrust Challenges*, 37 HARV. J.L. & PUB. POL'Y 931, 956-62 (2014). Federal courts have declined to strike down delegations to private entities since *Carter v. Carter Coal Co.*, 298 U.S. 238 (1936). See Gillian E. Metzger, *Privatization and Delegation*, 103 COLUM. L. REV. 1367, 1438-40 (2003); *Curran v. Wallace*, 306 U.S. 1, 15-18 (1939) (Delegation by Secretary of Agriculture to private auction markets to inspect and sell tobacco products is constitutional because growers vote on creation of markets by referendum and Secretary may suspend designated market for lack of competent inspectors or insufficient quantity of tobacco.); *Todd & Co. v. S.E.C.*, 577 F.2d 1008, 1012-14 (3d Cir. 1977) (describing how delegated self-regulation of securities dealers by SEC to private, independent association not unconstitutional as permitted by Maloney Act, which permits the SEC to register private organizations to discipline brokers in order to prevent fraud and protect investors, with a right of appeal to the SEC).

225 See generally Davidson, *supra* note 15, at 622-24. Courts routinely strike down local zoning ordinances, for example, that delegate lawmaking power to private residents to limit neighbors' use of their property. See, e.g., *Marta v. Sullivan*, 248 A.2d 608, 610 & n.3 (Del. 1968).

226 See Jim Rossi, *Institutional Design and the Lingering Legacy of Antifederalist Separation of Powers Ideals in the States*, 52 VAND. L. REV. 1167, 1191-1201 (1999). The nondelegation doctrines of California and Washington both track federal nondelegation doctrine. See *Light v. State Water Res. Control Bd.*, 226 Cal. App. 4th 1463, 1494 (Cal. Ct. App. 2014) (finding that private delegation of development of programs to preserve wild salmon not unlawful delegation so long as agency exercised independent discretion in their evaluation); *Murray v. State, Dept. of Lab. & Ind.* 403 P.3d 949, (Wash. Ct. App. 2017) (finding that legislative delegation of determination regarding whether health technology is includable as a covered benefit in a health care program to a public authority composed of private physicians did not violate state nondelegation doctrine).

227 667 F.3d 947 (7th Cir. 2012).

delegation of rulemaking power to unions and employers subject to CBAs. In rejecting this argument, the Seventh Circuit reasoned that the Illinois statute, which required administrative and judicial review of a prevailing wage rate before it became final, sufficiently constrained private rulemaking power to satisfy the nondelegation doctrine.²²⁸ In contrast, the Second Circuit in *General Elec. Co. v. New York State Dep’t of Labor*²²⁹ held that an allegation that a union and employer colluded to raise prevailing wage rates established through their CBA to off-set lower private-sector wages (thereby shifting the disproportionately high rate to the taxpayers) without any administrative oversight sufficient to allege an unconstitutional delegation.²³⁰

Beary and *General Electric* suggest that in these states, the nondelegation doctrine requires state and local legislatures in legislating collaborative enforcement to refrain from delegating rulemaking power to PIOs unless it is guided by administrative agencies and subject to judicial review. Tripartite rulemaking in these states can follow the model of the New York wage board, which included government representatives, required public agency approval and afforded opponents the opportunity for judicial review.²³¹

While *Beary* and *General Electric* reflect the majority view of how states constrain tripartite regulation via the nondelegation doctrine, a minority of states have a stronger nondelegation doctrine,²³² and it is an open question how these states treat private delegations. At the outer end of private delegation skepticism is Texas, which crafted a unique set of criteria to “specifically impose some limits on delegations to private parties.”²³³ Reasoning that “private delegations clearly raise even more troubling constitutional issues than their public counterparts,” the Texas Supreme Court

228 *Id.* at 951.

229 936 F.2d 1448, 1457–59 (2d Cir. 1991).

230 *Id.* at 1457–59.

231 See *Nat'l Rest. Ass'n v. Comm'r of Labor*, 141 A.D.3d 185, 191–95 (3d Dept 2016). See also *Andrias*, *supra* note 6, at 65, 89–90.

232 This is the law in New York. See *N.Y. Statewide Coal. of Hisp. Chambers of Com. v. N.Y.C. Dept. of Health and Mental Hygiene*, 23 N.Y.3d 681, 711 (2014) (Read, J., dissenting) (characterizing New York’s nondelegation doctrine as “strong” in criticizing the majority decision striking down the Sugary Drinks Portion Cap Rule as an unconstitutional delegation of lawmaking).

233 Volokh, *supra* note 224, at 965.

in *Texas Boll Weevil Eradication Foundation v. Lewellen*²³⁴ held that in addition to judicial review and safeguards against abuse, that state's nondelegation doctrine requires that private delegations be free from potential conflicts,²³⁵ which is difficult to reconcile with the purpose of tripartite regulation to include all interested stakeholders in the regulatory process.²³⁶

One might discount the Texas nondelegation doctrine as an outlier, but it does reflect a broader concern of the deeper involvement of private actors in public regulation at the local level, raising judicial concerns about local corruption.²³⁷ Indeed, this skepticism is warranted, as is Jon Michaels's call for a stronger nondelegation doctrine in order to protect "the tripartite architecture of administrative power" from privatization.²³⁸ While this Article agrees that judicial supervision must account for the incentives for abuse in private delegations,²³⁹ as Nestor Davidson argues in proposing a functionalist local nondelegation doctrine, oversight of collaborative enforcement "should reflect the advantages as well as the risks that public involvement, knowledge, and accountability bring to local agencies."²⁴⁰ Collaborative enforcement seeks to integrate private participation in tripartite workplace regulation, a potentially

234 952 S.W.2d 454, 468 (Tex. 1997). In Lewellen, the Texas legislature created a non-profit organization to oversee boll weevil eradication efforts to protect the state's cotton industry, overseen by the state's agriculture commission. *Id.* at 457–62.

235 *Id.* at 469–72.

236 See Volokh, *supra* note 224, at 969–70.

237 Davidson, *supra* note 15, at 623–24.

238 MICHAELS, *supra* note 167, at 188–90; see also Christopher Slobogin, *Panvasive Surveillance, Political Process Theory, and the Nondelegation Doctrine*, 102 GEO. L.J. 1721, 1770–75 (2014) (arguing that "mission creep" and other potential abuses of large-scale surveillance and enforcement justify administrative law controls on law enforcement).

239 This Article agrees that the nondelegation doctrine should supervise private delegations that "disrupt the democratic inclusivity and heterogeneity of civil society." MICHAELS, *supra* note 167, at 190. But a test that hinges on the public/private distinction is difficult to apply to collaborative enforcement, which is neither fully public nor private, and collaborative enforcement is entitled to more deference than other private delegations that create a high risk of abuse. Specifically, a functional nondelegation doctrine would not prohibit delegations to private stakeholders that have an interest in the delegation, would instead determine whether the limited delegation and intertwined nature of the delegation were sufficient safeguards against abuse, and would only invalidate tripartite rulemaking if it lacked administrative oversight and an inclusive process to invite participation by private stakeholders.

240 Davidson, *supra* note 15, at 624.

democracy-enhancing benefit. Stimulating tripartite regulation of the low-wage workplace requires administrative interventions that encourage PIOs to channel worker voice into regulation. In most states, this benefit and the limited delegation and intertwined nature of collaborative enforcement should be sufficient for its private delegations to survive nondelegation review. But a strong nondelegation doctrine may ignore these benefits and controls, greeting collaborative enforcement with the same skepticism as privatization schemes that invite local corruption or administrative aggrandizement. As with opt-outs under the NLRA, in Texas and states with a similar nondelegation test for private delegations, remedial enforcement can avoid nondelegation challenges through a standard, non-tripartite rulemaking process.

V. Conclusion

This Article offers an important, emerging proposal for how state and local public agencies and PIOs can collaborate to make workplace regulation more effective and efficient, create a system of political accountability for public enforcement, and facilitate sophisticated forms of tripartism. It proposes positioning PIOs as private enforcers coordinating resources and creating and deploying sanctions with state and local public agencies to change the behavior of regulated entities with high rates of legal noncompliance. This inverts the standard account of the PIO role in regulation, from enforcing self-regulation by high-compliance firms to amplifying the deterrent effect of enforcement among low-compliance firms, and to elaborate new legal requirements. This refines previous theories of public-private regulatory experimentation by showing that collaboration can not only promote self-regulation where the regulated entities have the expertise and motivation to comply with the law, but also to deter violations where enforced self-regulation is unlikely in the near term. The transparency that collaborative enforcement requires makes public enforcers more political accountable for the value-laden choices they make, and can preserve their independent judgment by limiting the scope of private delegation and preventing capture by the regulated entities. While the private delegation of collaborative enforcement can create incentives for abuse, the limited delegation and intertwined nature of collaborative enforcement make abuse less likely than more familiar forms of private delegations, such as deputization, and manageable through political and administrative controls. This analysis is applicable to other areas, such as consumer

protection, fair housing and civil rights enforcement, in which PIOs cannot effectively vindicate private rights alone, and in which public agencies may benefit from PIO sophistication and access to private enforcement tools.

A final, underdiscussed theme of collaborative enforcement is the role of state and local government in mediating and channeling public participation into enforcement governance, and as a site for sophisticated tripartite regulation. Coalescing individuals into stakeholder groups is of particular value in regulating small, undercapitalized employers, where traditional regulation rarely reaches, and which otherwise would not participate in tripartite regulation of the workplace.²⁴¹ This suggests that the effectiveness of collaborative enforcement in facilitating tripartism will depend, in part, on where the collaborating agency is positioned in the federalist system.

* * *

241 See ESTLUND, *supra* note 7, at 141.

THE SUPREME COURT'S CONSTITUTIONAL "BRIGHT LINE":

PREEMPTING AUTHORITY OF 47 OF 50 STATES

*Steven Ferrey**

* Steven Ferrey is Professor of Law at Suffolk University Law School and was Visiting Professor of Law at Harvard law School in 2003. Since 1993, Professor Ferrey has been a primary legal consultant to the World Bank and the U.N. Development Program on their renewable and carbon reduction policies in developing countries, where he has worked extensively in Asia, Africa, and Latin America. He holds a B.A. in Economics from Pomona College, a Juris Doctor degree and a Masters degree in Regional Energy Planning both from U.C. Berkley, and was a post-doctoral Fulbright Fellow at the University of London between his graduate degrees. He is the author of seven books on administrative and regulatory energy and environmental law and policy, the most recent of which are *UNLOCKING THE GLOBAL WARMING TOOLBOX* (2010), *ENVIRONMENTAL LAW* (7th ed. 2016), and *THE LAW OF INDEPENDENT POWER* (44th ed., 2017). He also is the author of more than ninety articles on these topics. Professor Ferrey thanks research assistant Nebyu Retta for his assistance.

Table of Contents

I.	Supreme Court Intaglio	145
II.	State Action; Federal Court Reaction	148
A.	<i>Maryland Regulatory Program Asserts State Influence Over Wholesale Markets</i>	148
B.	<i>The Law on the “Line” - State and Federal Authority.....</i>	153
1.	<i>The Supremacy Clause and the Commerce Clause .</i>	153
2.	<i>Supreme Court Precedent Establishing the Jurisdictional Lines.....</i>	155
C.	<i>The Federal Trial Court Establishes the Legal Preemption Benchmark</i>	158
D.	<i>The Unanimous Federal Circuit Court Opinion.....</i>	161
III.	Dissecting a Supreme Court Unanimous 2016 Decision	165
A.	<i>Reaffirmation of the Traditional ‘Bright Line’ of Field Preemption</i>	165
B.	<i>Concurring ‘Brackets’ on the Opinion Refine its Foundation.....</i>	173
IV.	Within the Brackets: How the Opposing Parties Framed Key Legal Issues for the Court to Decide	176
A.	<i>Field Preemption.....</i>	177
1.	<i>Maryland Appellants</i>	178
a.	<i>State of Maryland, Appellant/Petitioner</i>	178
b.	<i>CPV, Appellant/Petitioner.....</i>	181
2.	<i>Respondent/Appellee Talen Energy Marketing</i>	184
3.	<i>Amici Allco Renewable Energy in Support of Respondents</i>	189
B.	<i>Conflict Preemption</i>	191
1.	<i>Maryland Petitioners/Appellants</i>	192
2.	<i>Respondent/Challengers</i>	194
3.	<i>Amici, Allco.....</i>	196
C.	<i>The Supreme Court Response to the Parties’ Framing ..</i>	196
V.	Constitutional Implications of Field Preemption for the Most Regulated Industry.....	198

I. Supreme Court Intaglio¹

In 2016, the Supreme Court reinforced a constitutional “bright line” by stripping 47 (but not affecting 3) of the 50 states of what the states thought were their regulatory prerogatives: electricity moving in the high voltage transmission network.² In this case, *Hughes v. Talen Energy Marketing, LLC*, the Supreme Court retched³ in high relief the most prominent and distinctive “bright lines” in American law.⁴ And it did so unanimously, overturning state jurisdiction and authority.⁵

The important question is whether this unanimous Supreme Court decision expands constitutional “field preemption,” to permanently withhold from 47 states this power, or only employs “conflict preemption” to preempt a single narrow type of state regulation. It is the answer to this question that will determine the enduring jurisprudential legacy of the holding in *Hughes v. Talen Energy* and the future of American power.⁶

Electricity is the foundation of the modern economy⁷ and has been called the second most important invention since the wheel.⁸ Electricity has a delivered value in the U.S. of approximately \$390

1 Intaglio is a printing process which uses a metal plate. Charlotte Jirousek, *Printmaking Processes, ART, DESIGN, AND VISUAL THINKING: AN INTERACTIVE TEXTBOOK* (1995), <http://www.char.txa.cornell.edu/media/print/print.htm>.

2 See *Hughes v. Talen Energy Marketing, L.L.C.*, 136 S. Ct. 1288 (2016).

3 In etching, “a strong metal plate is coated with a material called a ground. The artist then draws his [or her] design on the ground with a sharp needle[] that cuts through the ground to the metal below The artist etches on the plate those parts which will appear in the finished print as black or colored areas. Since the ground is soft, the artist is able to work more freely than is possible with engraving.” Charlotte Jirousek, *Printmaking Processes, ART, DESIGN, AND VISUAL THINKING: AN INTERACTIVE TEXTBOOK* (1995), <http://www.char.txa.cornell.edu/media/print/print.htm>.

4 *Hughes*, 136 S. Ct. 1288.

5 *Id.* at 1297, 1300 (though note that Justice Thomas concurred with the opinion only in part).

6 *Id.* at 1299.

7 STEVEN FERREY, ENVIRONMENTAL LAW: EXAMPLES & EXPLANATIONS, 580-82 (7th ed. 2016).

8 James Fallows, *The Fifty Greatest Breakthroughs Since the Wheel*, ATLANTIC MONTHLY (Nov. 2013). Electricity finished behind only the movable type printing press. Electricity is essential to operate seven other “top fifty” inventions of all time: The Internet, computers, air-conditioning, radio, television, the telephone, and semiconductors. *Id.*

billion annually,⁹ exceeding the total amount of corporate income taxes collected in the U.S.¹⁰ The high-voltage electric transmission network was recognized by engineers as the most important engineering feat of the 20th century.¹¹ This network is controlled by law. In its *Hughes v. Talen Energy* decision, the Supreme Court re-etched the architecture of American power.¹² This judicially-imposed Constitutional “bright line” determines *if, how, and by whom/what* the nation moves forward with this key technology and, consequently, the degree to which the planet’s climate warms.¹³

The Supreme Court’s unanimous decision in 2016 reconfigures the Constitution’s Supremacy Clause to eliminate state authority over electric capacity prices under both implied field preemption and conflict preemption. This Clause, through these preemption doctrines, is a core constitutional principle separating the branches of government and powers, a distinguishing feature of American government.¹⁴ In *Hughes*, the Justices had a choice to find and implement either field preemption or conflict preemption. The former creates an extensive “field” of authority removed permanently from the states, while the latter bars only the specifically challenged regulation. This distinction is as close as the Supremacy Clause comes to an “all or virtually none” enduring legacy of a decision. This article analyzes in extensive detail every aspect of this decision and concludes that this decision upheld the former, “field preemption,” to permanently crimp and curtail state authority in 47 of the 50

9 U.S. ENERGY INFORMATION ADMINISTRATION, TABLE 1.1. TOTAL ELECTRIC POWER INDUSTRY SUMMARY STATISTICS, 2015 AND 2014, https://www.eia.gov/electricity/annual/html/epa_01_01.html (last visited October 28, 2017). The average delivered price of all electricity nationwide in 2011 was \$0.0966/Kwh, and \$0.1109/Kwh for residential customers. See Public Policy Institute of New York State, “Average Retail Price of Electricity to Ultimate Customers by End-Use Sector, by State, Year-to-Date through February 2011 and 2010,” available at <http://ppinys.org/reports/jtf/2011/employ/average-retail-price-of-electricity2010-11.htm>.

10 Terence P. Jeffrey, \$3,248,723,000,000: *Federal Taxes Set Record in FY 2015; \$21,833 Per Worker; Feds Still Run \$438.9B Deficit*, CNS News (Oct. 15, 2015, 3:46 PM), <https://www.cnsnews.com/news/article/terence-p-jeffrey/3248723000000-federal-taxes-set-record-fy-2015-21833-worker-feds-0>.

11 Mason Willrich, *Electricity Transmission Policy for America: Enabling a Smart Grid, End to End*, 22 ELECTRICITY J. 77 (2009).

12 See *Hughes v. Talen Energy Mktg., L.L.C.*, 136 S. Ct. 1288 (2016).

13 See generally STEVEN FERREY, UNLOCKING THE GLOBAL WARMING TOOLBOX: KEY CHOICES FOR CARBON RESTRICTION AND SEQUESTRATION (2010).

14 U.S. CONST., art. VI.

states over a critical U.S. technology.

This distinction as to whether the Supreme Court extends field preemption or uses conflict preemption is important. Not only is this pivotal Supreme Court decision unanimous; it is unique. Only for electric power transactions, and no others, is the law separated by a legal “bright line” which allocates exclusive control to federal and state government depending on whether a power transaction is labeled as “wholesale” in interstate commerce or “retail” and not in interstate commerce. Nowhere else in American law, other than for electric power, is exclusive legal jurisdiction so distinctly and peculiarly divided.

An additional distinction of *Hughes* is that the opinion only preempts the jurisdiction and authority of 47 of the 50 states. Electricity is the movement of electrons in a metal line and commerce in electricity only occurs through a copper transmission or distribution line.¹⁵ Hawaii and Alaska are remote and not interconnected by transmission lines to any other states, and thus cannot engage in interstate commerce in electricity. Additionally, the majority of Texas is not interconnected by wire to the rest of the United States and does not engage in interstate commerce in electricity.¹⁶ *Hughes* interprets the Federal Power Act (FPA),¹⁷ which only applies the Constitution’s Supremacy Clause to those states engaging in interstate commerce.¹⁸ Thus, there is a different preempted legal rule for the regulatory authority of the 47 interconnected contiguous states than for the other 3 states that are not interconnected for purposes of interstate commerce in power. Therefore, a field preemption decision impact affects 47 states, and exempts the other 3.

After *Hughes*, this “bright line” of preemption is elevated to preempt what 47 states have attempted to regulate. This article

15 *Electrical Energy, Energy for Missouri: Today and Tomorrow - Educator's Guide*, Mo. DEP'T OF ECON. DEV., DIV. OF ENERGY, https://energy.mo.gov/docs/default-source/energy_division/ElectricalEnergy.pdf (last visited July 17, 2017). Electric circuits convey energy in a force field to different locations, always within the line or attachments to it.

16 See *New York v. F.E.R.C.*, 535 U.S. 1 (2002). “Two of these grids-the "Eastern Interconnect" and the "Western Interconnect"-are connected to each other. It is only in Hawaii and Alaska and on the "Texas Interconnect"-which covers most of that State-that electricity is distributed entirely within a single State. In the rest of the country, any electricity that enters the grid immediately becomes a part of a vast pool of energy that is constantly moving in interstate commerce.” *Id.* at 7.

17 Federal Power Act, 16 U.S.C. §§ 791a – 828c (2012).

18 *Hughes v. Talen Energy Mktg., L.L.C.*, 136 S. Ct. 1288, 1291–93 (2016).

traces the “how” and “why” of the Supreme Court decision to elect between field preemption and conflict preemption, as well as the key legal implications of *Hughes*. Section II starts by examining the specific state attempt to regulate energy which was challenged in *Hughes*. It examines each level of jurisprudence in the lower federal courts at the trial and circuit levels before reaching the Supreme Court. Section II also analyzes the interstices of the Supremacy Clause and its ongoing application to critical technology in America.¹⁹ The article analyzes how and why the decisions were unanimous at both the district court²⁰ and the Fourth Circuit Court of Appeals,²¹ both holding that the state regulation violated the Supremacy Clause and was preempted in its entirety.

Section III examines the appeal of the case to the Supreme Court. Section III analyzes, through different preemptive lenses, the substantive Supremacy issues presented before the Supreme Court. The Article looks at the nuance of the two doctrines of implied preemption: field preemption and conflict preemption. Section III peels back the legal layers to analyze carefully both the primary opinion, as well as two critical, but to date largely overlooked, concurring opinions whereby two Justices “bracket” the majority opinion of the Court. The details and nuance of these concurrences matter in reestablishing the bottom-line of this constitutional “bright line” going forward in American law.²²

Context matters in a constitutional challenge of state authority. Section IV examines in depth how the legal issues were presented to the Court by the two opposing parties and key amici. These frame the Constitutional issues before the Court and inform the scope of the decision. Finally, section V charts what a single Supreme Court opinion in 2016 signals for the second most important invention in history against the Supremacy Clause of the Constitution. This determines the powers of different levels of government over electric power in the United States.

II. State Action; Federal Court Reaction

A. Maryland Regulatory Program Asserts State Influence Over Wholesale Markets

19 See *infra* Section II B.

20 See *infra* Section II C.

21 See *infra* Section II D.

22 See *infra* Section III B.

An energy program in Maryland motivated a challenge under the Supremacy Clause and the Commerce Clause of the Constitution.²³ The state of Maryland adopted competitive retail markets in 1999, and authorized its utilities to participate in the PJM Regional Transmission Organization (“PJM”).²⁴

More recently, Maryland required its utilities to enter long-term “contract for differences” (“CfD”), a form of power purchase agreement (“PPA”), only with certain designated independent power producers willing to locate their new electricity generation capacity in Maryland or the District of Columbia.²⁵ The Maryland law required its utilities to enter into a 20-year CfD with the winning in-state project bidder.²⁶ Commercial Power Ventures, Maryland (“CPV”) won the bid.²⁷ This CfD established the final wholesale rates that CPV would receive for a 20-year period — in part in capacity payment from the PJM capacity auction and in part from Maryland augmenting that winning price, in return for in-state locus.²⁸ The PJM interstate wholesale “capacity auction” is designed to ensure that winning electricity generation will be constructed to meet future demand, taking bids approximately 3 years in advance of need and selecting future generator bids from proposals which agree to install generating capacity by three years hence.²⁹ The Maryland CfD provided that regardless of the price set, pursuant to Federal Energy

23 PPL EnergyPlus, L.L.C. v. Nazarian, 974 F. Supp. 2d 790, 796 (D. Md. 2013), *aff'd*, 753 F.3d 467 (4th Cir. 2014), *aff'd sub nom.* Hughes v. Talen Energy Marketing, L.L.C., 136 S. Ct. 1288 (2016).

24 *Deregulation, BALT. GAS AND ELEC. Co.*, <http://www.taxpolicycenter.org/taxfacts/displayafact.cfm?Docid=203> (last visited Aug. 25, 2016).

25 Hughes v. Talen Energy Mktg., L.L.C., 136 S. Ct. 1288, 1294–97 (2016).

26 *Id.*

27 *Id.* at 1294.

28 *Id.* at 1295. If CPV’s winning bid for capacity payments was less than the Maryland contract price, Maryland utilities would pay it the difference; if the reverse, CPV would pay the Maryland utilities the difference. Consequently, CPV had no incentive to submit its true competitive auction bid.

29 *Id.* at 1293. If a capacity owner’s bid is accepted, it is said to have “cleared” the auction. Capacity owners that have cleared the market by remaining in as the bid price bar is lowered are all paid the clearing price for capacity, which is the price of the highest accepted bid. FERC’s Minimum Offer Price Rule requires new generators to submit their initial bids at a price set by the grid operator unless they can demonstrate that this price exceeds their actual costs. Once a capacity owner has submitted a clearing bid, in subsequent rounds of the auction, it can submit a lower bid of its own choosing, and may be able to take advantage of the New Entry Price Adjustment (“NEPA”) rule, which guarantees stable prices for three years. *Id.* at 1294, 1296.

Regulatory Commission (FERC) approval,³⁰ by the relevant PJM wholesale market capacity auction,³¹ the Maryland utilities would assure that the Maryland-selected in-state power projects received a guaranteed price, in part augmented by state funds and fixed by long-term contractual formula.³²

PJM operates the “largest centrally dispatched power market . . . in the world,” covering 60 million customers and 185,000 megawatts (MW) of power generation, including the District of Columbia and all or part of thirteen states including Maryland.³³ The PJM market regulates and controls all wholesale sale of power through an interstate, federally-regulated power market. PJM provides capacity payments to winning generation facilities for the siting of new power generation as needed throughout this thirteen-state area.³⁴ PJM operates pursuant to the “Open Access Transmission

-
- 30 Regional Transmission Organizations (RTOs) or Independent System Operators (ISOs) are FERC-approved and regulated entities which facilitate commercial electricity transfers, through a private corporation that functions as a tariff administrator. RTOs are responsible for managing both electrical and financial transactions, including scheduling transmission transactions, dispatching generation, and managing the entire accounting for the grid capacity and energy charges and transmission fees. See STEVEN FERREY, LAW OF INDEPENDENT POWER, §§ 8:10, 10:87, 10:91 (44th ed. 2017); STEVEN FERREY, THE NEW RULES: A GUIDE TO ELECTRIC MARKET REGULATION, 49–50 (2000).
- 31 PJM, an ISO, is a FERC-created and authorized entity. In the PJM ISO, which serves multiple Eastern states, there are two retail energy markets, a real-time (spot) and a day-ahead (forward) market. The basis of calculating the electricity price in either market is Locational Marginal Pricing. PJM’s capacity-market model, the Reliability Pricing Model, was implemented in 2007 as the successor to its Capacity Credit Market design, as a series of auctions for a delivery year approximately three years in the future. PJM’s demand curve, the Variable Resources Requirement, defines the price for a given capacity commitment relative to the applicable reliability requirement, defined for each constrained Locational Delivery Area. See *Amended and Restated Operating Agreement of PJM Interconnection, L.L.C.*, INTRA-PJM TARIFFS (2011), <http://www.pjm.com/directory/merged-tariffs/oa.pdf>. See also *Electric Power Markets: PJM*, FERC, <http://www.ferc.gov/market-oversight/mkt-electric/pjm.asp> (last updated Aug. 1, 2017).
- 32 PPL EnergyPlus, L.L.C., et al. v. Nazarian, 974 F. Supp. 2d 790, 831–33 (D. Md. 2013), *aff’d*, 753 F.3d 467 (4th Cir. 2014), *aff’d sub nom.* Hughes v. Talen Energy Marketing, L.L.C., 136 S. Ct. 1288 (2016).
- 33 PPL EnergyPlus, L.L.C., et al., v. Hanna, 977 F. Supp. 2d 372, 378–79 (D. N.J. 2013), *aff’d sub nom.* PPL EnergyPlus, L.L.C. v. Solomon, 766 F.3d 241 (3d Cir. 2014).
- 34 *Id.* at 395.

Tariff” approved by FERC.³⁵

The Maryland-enacted CfD contained provisions which enabled the selected supplier to receive its proposed “contract price” for each unit of energy and capacity sold to PJM in the PJM markets, up to a CfD ceiling amount.³⁶ Maryland electricity ratepayers would supply any cash supplement price between the uniform winning PJM bid price for capacity payments and the price established by the CfD rates approved by Maryland. Because the state subsidizes the price difference, this encourages and indemnifies the selected Maryland generators to bid lower than they might otherwise in order to win the PJM capacity auction, and suppresses artificially the winning bid(s) price cleared for all generators.³⁷ The challenger’s complaint argued that the Maryland program encouraged the state-selected new power generator, CPV, to:

submit capacity bids to the PJM forward capacity auction at a price lower than its actual competitive cost of producing power,

thereby artificially driving down the market-clearing price for all producers, and

usurping the otherwise winning bid and position of other power market competitors.³⁸

The law did affect the geographic situs of the generating facility; however, there was no requirement for the owner of the in-state facility to be an in-state company. The capacity payment is not to cover one’s variable costs of operation, but the capital and fixed costs of construction: One is paid by PJM the full 100% value for capacity installed, even if the facility is never called on to operate.³⁹

Putting aside the possible tax revenue, jobs, and other benefit that the host state could receive from additional large projects sited in-state,⁴⁰ these programs shift, in part, the cost of in-state electric

35 *Id.* at 378.

36 *Nazarian*, 974 F. Supp. 2d at 821–22.

37 For more on “second price” electricity auctions see STEVEN FERREY, LAW OF INDEPENDENT POWER, § 9:26 (44th ed. 2017). For more on the operation of ISOs, such as PJM, see *id.* at § 10:106.

38 Brief for Respondent at 52, *Hughes v. Talen Energy Mktg., L.L.C.*, 136 S. Ct. 1288 (2016) (No. 14-614).

39 Exelon Corp., 150 FERC 61,067, ¶ 30 (2015).

40 See FERREY, *supra* note 37 at §§ 10:78–10:90 (State and local tax consequences

capacity projects from in-state ratepayers to ratepayers who pay for PJM capacity costs in the 13 states comprising the largest independent system operator (ISO) in the country. In most, but not all, of the ISOs in the U.S., power generators are paid a capacity payment for just being there and operational, even if the plant does not generate power in fact.⁴¹ As Woody Allen was fond of saying: "Showing up is 80 percent of life."⁴² All the power in PJM is in interstate commerce as "the PJM region encompasses all or part of thirteen states and the District of Columbia."⁴³

After capacity payments, there is a second component of payment for power generated and delivered. Bids to generate power are accepted in order of ascending price; in the PJM ISO, only the lowest-cost power bids for each day are asked to deliver power.⁴⁴ Therefore, payments for energy reflects the lowest-cost producers of the amount of power that is actually demanded, purchased, and used by consumers.

The legal issue around the Maryland program is whether one should be able to collect that capacity payment through the PJM ISO which bills all consumers in the 13 PJM states and the District of Columbia. When any state subsidizes location of power generation capacity in its state not simply with state grants or tax benefits,⁴⁵ but through an imposed contractual condition, FERC does *not* authorize that the power generation facility first recover its capacity payments from consumers in all the 13 states through an interstate PJM ISO. This state condition and cross-subsidization of in-state power generation resources interferes with a wholesale capacity market exclusively within federal FERC authority; this issue was raised by allegedly disadvantaged wholesale power generators to the federal courts.⁴⁶

-
- after acquisition of in-state power generation assets).
- 41 See id. at § 10:106 (operation of ISOs which conduct capacity market operation).
- 42 See William Safire, *On Language; The Elision Fields*, N.Y. TIMES MAG. (Aug. 13, 1989), <http://www.nytimes.com/1989/08/13/magazine/on-language-the-elision-fields.html>.
- 43 PPL Energyplus, L.L.C. v. Hanna, 977 F. Supp. 2d 372, 377-79 (D. N.J. 2013), *aff'd sub nom.* PPL Energyplus, L.L.C. v. Solomon, 766 F.3d 241, 247 (3d Cir. 2014).
- 44 See generally FERREY, *supra* note 37 at § 10:106; see F.E.R.C. v. Elec. Power Supply Ass'n, 136 S. Ct. 760 (2016).
- 45 For discussion of state grants and tax benefits, See FERREY, *supra* note 38 at §§ 10:82, 10:114, 10:151, 10:153.
- 46 Hughes v. Talen Energy Mktg., L.L.C., 136 S. Ct. 1288, 1295 (2016).

B. The Law on the “Line” – State and Federal Authority

1. The Supremacy Clause and the Commerce Clause

The separation of U.S. government power is established and mediated by the Commerce Clause and the Supremacy Clause of the Constitution. From 1937 until 1995, the Supreme Court rebuffed every Commerce Clause challenge to federal statutes.⁴⁷ In *United States v. Lopez*,⁴⁸ the Supreme Court established the following canonical test:

“Congress may regulate use of the channels of interstate commerce”⁴⁹

Congress may “regulate and protect the instrumentalities of interstate commerce, or persons or things in interstate commerce, even though the threat may come only from intrastate activities”⁵⁰

Congress may regulate activities, even intrastate ones, that alone or in the aggregate “substantially affect” interstate commerce.”⁵¹

That the Supremacy Clause of the Constitution creates a separation of state and federal power is one of the most litigated areas of Constitutional law. The Supremacy Clause establishes preemption of federal law over and in lieu of state and local regulation: “[The laws of the United States . . . shall be the supreme Law of the Land; and the Judges in every State shall be bound thereby, anything in the Constitution or laws of any State to the Contrary notwithstanding.”⁵² Preemption doctrine enforces the supremacy of the higher or highest (*i.e.* federal) levels of government jurisdiction. Three types of preemption can occur:

federal law could expressly establish the lines for state

47 See Lainie Rutkow & Jon S. Vernick, *The U.S. Constitution's Commerce Clause, the Supreme Court, and Public Health*, 126 PUB. HEALTH REP. 750, 751 (2011).

48 *United States v. Lopez*, 514 U.S. 549, 558–60 (1995).

49 *Id.* at 558.

50 *Id.*

51 *Id.* at 559.

52 U.S. CONST. art. VI, cl. 2.

preemption;⁵³

without any express preemption, “state law may be preempted if it regulates conduct in a field Congress intended the federal government to occupy exclusively, either because the federal regulatory scheme is ‘so pervasive’ that a court may infer Congress left ‘no room for the States to supplement it’;”⁵⁴ or

even if not *per se* barred from regulating in a specific area, state laws enacted could actually conflict with federal law.⁵⁵

For purposes of electric energy in America, the Federal Power Act (“FPA”) of 1935 is the vehicle of preemption jurisdiction.⁵⁶ Section 201(a) of the FPA states that federal regulation under the statute “only extend[s] to those matters which are not subject to regulation by the States.”⁵⁷ FERC does not regulate the construction of transmission facilities themselves, only economic tariffs for transactions moving power over them.⁵⁸ Section 201(b) of the FPA gives FERC jurisdiction over all “facilities” for the transmission or wholesale sale of electric energy; it does not have jurisdiction “over facilities used for the generation of electric energy” or “used in local distribution or only for the transmission of electric energy in intrastate commerce,” or “facilities for the transmission of electric energy consumed wholly by the transmitter.”⁵⁹ Sections 205 and 206

53 English v. Gen. Elec. Co., 496 U.S. 72, 78–79 (1990).

54 Entergy Nuclear Vt. Yankee, L.C.C. v. Shumlin, 838 F. Supp. 2d 183, 218. *See also* English, 496 U.S. at 79.

55 Entergy Nuclear, 838 F. Supp. 2d at 218.

56 See Federal Power Act, 16 U.S.C. §791a (2017). *See e.g.* Hughes v. Talen Energy Marketing, L.L.C., 136 S. Ct. 1288 (2016).

57 16 U.S.C. § 819 (repealed 1935); 16 U.S.C. §824(a) (2015).

58 *See Transmission Planning and Cost Allocation by Transmission Owning and Operating Public Utilities*, Order No. 1000, 76 Fed. Reg. 49,842 (Aug. 11, 2011), 136 FERC ¶ 61,051 at para. 287 (2011). This pertains only to Commission-jurisdictional tariffs or agreements and does not require removal of references to such state or local laws or regulations from Commission-approved tariffs or agreements. *See Transmission Planning and Cost Allocation by Transmission Owning and Operating Public Utilities*, Order No. 1000, 76 Fed. Reg. 49,842 (Aug. 11, 2011), 136 FERC ¶ 61,051 at para. 57, 231 (2011). FERC noted that Order 1000 does not address the prudence of investment decision nor determine which particular entity should construct any particular transmission facility, but merely to allow more entities to be considered for potential construction responsibility. *See* 136 FERC ¶ 61,051 at para. 290.

59 16 U.S.C. § 819 (repealed 1935); 16 U.S.C. §824(b) (2015).

of the FPA⁶⁰ empower FERC exclusively to regulate the commerce and rates for the interstate and wholesale sale and transmission of electricity in the United States.⁶¹

2. Supreme Court Precedent Establishing the Jurisdictional Lines

With regard to regulation of electricity, the Supreme Court originally held that the Commerce Clause of the Constitution, while empowering the federal government to regulate interstate commerce, prohibited state and local regulators from regulating many interstate transactions, including interstate wholesale sales.⁶² More than eighty years ago, the FPA of 1935 filled this regulatory gap by providing when enacted that FERC has jurisdiction over interstate and wholesale power sales.⁶³ However, its authority does not extend to “any other sale of electric energy.”⁶⁴ States retain “most matters they traditionally regulated, including local utility rates and the siting of power plants.”⁶⁵

In interpreting the FPA and the Supremacy Clause, the U.S. Supreme Court established a “bright line” between federal and state jurisdiction. The Court held that “Congress meant to draw a bright line, easily ascertained and” not requiring case-by-case analysis, “between state and federal jurisdiction.”⁶⁶ Where a transaction is subject to exclusive federal FERC jurisdiction, inferior state regulation is preempted under the Supremacy Clause pursuant to a long-standing and consistent line of Supreme Court holdings.⁶⁷

60 16 U.S.C. §§ 824d and 824e.

61 See Pub. Util. Dist. No. 1 of Snohomish Cty. v. FERC, 471 F.3d 1053, 1058–59, 1066–67 (9th Cir. 2006), *aff’d in part and rev’d in part sub nom.* Morgan Stanley Capital Group, Inc. v. Pub. Util. Dist. No. 1, 554 U.S. 527 (2008).

62 Pub. Util. Comm’n of R.I. v. Attleboro Steam & Elec. Co., 273 U.S. 83, 89–90 (1927); See 16 U.S.C. § 824a(d) (2012).

63 16 U.S.C. § 824a; PPL EnergyPlus, L.L.C. v. Hanna, 977 F. Supp. 2d 372, 408 (D. N.J. 2013), *aff’d sub nom.* PPL EnergyPlus, L.L.C. v. Solomon, 766 F.3d 241 (3d Cir. 2014).

64 16 U.S.C. § 824(b)(1) (2012).

65 North Dakota v. Heydinger, 15. F. Supp 3d 912, 914 (8th Cir. 2016) (citations omitted).

66 Fed. Power Comm’n v. S. Cal. Edison Co., 376 U.S. 205, 215–16 (1964).

67 See, e.g., New England Power Co. v. New Hampshire, 455 U.S. 331, 338–44 (1982). The Supreme Court overturned an order of the New Hampshire Public Utilities Commission which restrained within the state, for the financial advantage of in-state ratepayers, low-cost hydroelectric energy produced within the state: “Our cases consistently have held that the Commerce Clause of the Constitution precludes a state from mandating that its residents be

This line applies to transactions in interstate commerce. FERC lacks jurisdiction over the siting, construction, or ownership of transmission facilities, which are exclusively within state jurisdiction.⁶⁸ The rates, terms, and provisions of any wholesale sale or transmission of electricity in interstate commerce are solely within federal jurisdiction and control, not state authority:⁶⁹ "FERC has exclusive authority to set and to determine the reasonableness of wholesale rates."⁷⁰ The FPA defines "sale of electric energy at wholesale" as any sale to any person for resale,⁷¹ which are beyond any state authority.⁷² "It is common ground that if FERC has jurisdiction over a subject, the States cannot have jurisdiction over the same subject."⁷³

The Supreme Court held that "it is difficult to conceive of a more basic element of interstate commerce than electric energy, a product used in virtually every home and every commercial or manufacturing facility. No State relies solely on its own resources in this respect."⁷⁴ The FPA "delegated to . . . the Federal Energy Regulatory Commission, exclusive authority to regulate the transmission and sale at wholesale of electric energy in interstate

given a preferred right of access, over out-of-state consumers, to natural resources located within its borders or to the products derived therefrom."

Id. at 338 (citations omitted). See also *Montana-Dakota Utils. Co. v. Nw. Pub. Serv. Co.*, 341 U.S. 246, 251 (1951); *Nantahala Power & Light Co. v. Thornburg*, 476 U.S. 953 (1986); *Miss. Power & Light Co. v. Mississippi ex rel. Moore*, 487 U.S. 354 (1988); *Entergy La., Inc. v. La. Pub. Serv. Comm'n*, 539 U.S. 39 (2003).

68 *Piedmont Envtl. Council v. FERC*, 558 F.3d 304, 309–10, 313 (4th Cir. 2009).

69 *New England Power Co.*, 455 U.S. at 340. The Supreme Court concluded that "§ 201(b) simply saves from pre-emption under Part II of the Federal Power Act such state authority as was otherwise 'lawful'" and that "[n]othing in the legislative history or language of the statute evinces a congressional intent 'to alter the limits of state power otherwise imposed by the Commerce Clause.'" *Id.* at 341 (citation omitted).

70 *Miss. Power & Light Co.*, 487 U.S. at 371 ("FERC has exclusive authority to determine the reasonableness of wholesale rates."); *accord*, *Public Utility District No. 1 of Snohomish County, Washington v. FERC*, 471 F.3d 1053, 1066; *aff'd in part and rev'd in part sub nom. Morgan Stanley Capital Grp., Inc. v. Pub. Util. Dist. No. 1*, 554 U.S. 527 (2008).

71 16 U.S.C. § 824(d) (2012) ("[S]ale of electric energy at wholesale' . . . means a sale of electric energy to any person for resale.").

72 *Fed. Power Comm'n v. S. Cal. Edison Co.*, 376 U.S. 205, 214; *Indep. Energy Producers Ass'n v. Cal. Pub. Util. Comm'n*, 36 F.3d 848 (9th Cir. 1994); *S. Cal. Edison Co.*, 70 FERC ¶ 61,215, 61,675 (1995).

73 *Miss. Power & Light Co.*, 487 U.S. at 377 (Scalia, J., concurring).

74 *FERC v. Mississippi*, 456 U.S. 742, 757 (1982).

commerce, without regard to the source of production.”⁷⁵ Where a state imposes a wholesale contract rate different than that set by FERC by either “law or policy,” the “contracts will be considered to be void *ab initio*.⁷⁶

When a transaction is subject to exclusive federal FERC jurisdiction and regulation, state regulation is preempted as a matter of federal law and the U.S. Constitution’s Commerce Clause, according to a long-standing and consistent line of rulings by the U.S. Supreme Court.⁷⁷ The Supreme Court in 1986,⁷⁸ 1988,⁷⁹ 2003,⁸⁰ and 2008⁸¹ reaffirmed the Supremacy Clause applied to negate state energy regulation over wholesale transactions otherwise subject to FERC’s exclusive authority. Under the FPA, a “bright line” exists between state and federal jurisdiction, with wholesale power sales ... falling on the federal side of the line.⁸² “Congress meant to draw a bright line, easily ascertained” and not requiring case-by-case analysis “between state and federal jurisdiction.”⁸³

75 *New England Power Co. v. New Hampshire*, 455 U.S. 331, 340 (1982) (citing *United States v. Pub. Utils. Comm’n of Cal.*, 345 U.S. 295, 311 (1953)); *see also Nantahala Power & Light Co. v. Thornburg*, 476 U.S. 953, 956 (1986) (stating that the Commission “has exclusive jurisdiction over interstate wholesale power rates.”).

76 *Conn. Light & Power Co.*, 70 FERC ¶ 61,012, 61,029–61,030 (1995).

77 *New England Power Co.*, 455 U.S. at 341–42. The Supreme Court overturned an order of the New Hampshire Public Utilities Commission that restrained within the state, for the financial advantage of in-state ratepayers, low-cost hydroelectric energy produced within the state. *Id.* at 344. It held this to be an impermissible violation of the dormant Commerce Clause of the U.S. Constitution, art. I, § 8, cl. 3, and the Federal Power Act, 16 U.S.C. §§ 791–828: “Our cases consistently have held that the Commerce Clause of the Constitution . . . precludes a state from mandating that its residents be given a preferred right of access, over out-of-state consumers, to natural resources located within its borders or to the products derived therefrom.” *New England Power Co.*, 455 U.S. at 338. *See also Entergy La., Inc. v. La. Pub. Serv. Comm’n*, 539 U.S. 39, 49–50 (2003); *Miss. Power & Light Co.*, 487 U.S. at 377; *Nantahala Power & Light Co.*, 476 U.S. at 963–64; *Montana-Dakota Utils. Co. v. Nw. Pub. Serv. Co.*, 341 U.S. 246, 251 (1951).

78 *Nantahala Power & Light Co.*, 476 U.S. at 966.

79 *Mississippi Power & Light Co.*, 487 U.S. at 370.

80 *Entergy La., Inc.*, 539 U.S. at 49–50.

81 *See Morgan Stanley Capital Grp.*, 554 U.S. at 531.

82 Pub. Util. Dist. No. 1 of Snohomish Cty. v. FERC, 471 F.3d 1053, 1066 (9th Cir. 2006), *aff’d in part, rev’d in part sub nom. Morgan Stanley Capital Grp. v. Pub. Util. Dist. No. 1 of Snohomish Cty. et al.*, 554 U.S. 527 (2008) (citing the separate Supreme Court opinions in *Nantahala Power & Light Co.*, *S. Cal. Edison*, and *Miss. Power & Light Co.*).

83 *Fed. Power Comm’n v. S. Cal. Edison Co.*, 376 U.S. 205, 215–16 (1964).

FERC is a quasi-judicial regulatory agency which also renders enforceable decisions and orders; over the past half century FERC case law establishes exclusive jurisdiction over the “transmission of electric energy in interstate commerce,” over the “sale of electric energy at wholesale in interstate commerce,”⁸⁴ and over “all facilities for such transmission or sale of electric energy.”⁸⁵

The critical precedential distinction in the case is whether these actions establish a case of conflict preemption, which only excludes states as to certain regulations, or field preemption, which excludes anything a state might regulate or impose. The trial court chose the latter, and the Court of Appeals added the other.⁸⁶

C. The Federal Trial Court Establishes the Legal Preemption Benchmark

In the *Hughes* case, the challengers argued that the Maryland CfD was preempted by field preemption, because the federal government fully occupies the electricity wholesale regulatory field, and conflict preemption, because the Maryland regulatory action impedes the FERC federal regulatory wholesale market as applied in the 13 PJM states.⁸⁷ Maryland retail utilities, which were required to divest their power generating facilities, must purchase energy for their retail customers on the federally regulated wholesale PJM interstate market.⁸⁸

CPV and the successful winning bidders under this in-Maryland energy regulation countered these claims by arguing that while the plant location was geographically limited, an out-of-state

84 16 U.S.C. § 824(b) (2012); *see e.g.*, Pa. Power & Light Co., 23 FERC ¶ 61,006 (F.E.R.C.) at 61,018, *reh'g denied*, 23 FERC ¶ 61325 (F.E.R.C.) (1983); N. Ind. Pub. Serv. Co., 66 FERC ¶ 61,213 at 61,488 (1994); Progress Energy, Inc., 97 FERC ¶ 61,141 at 61,628 (2001); Aquila Merchant Servs., Inc., 125 FERC ¶ 61175 at ¶ 17 (2008).

85 16 U.S.C. § 824(b); *see e.g.*, Pa. Power & Light Co., 23 FERC ¶ 61,006, at 61,018, *reh'g denied*, 23 FERC ¶ 61,325 (1983); N. Ind. Pub. Serv. Co., 66 FERC ¶ 61,213, at 61,488 (1994); Progress Energy, Inc., 97 FERC ¶ 61,141, at 61,628 (2001); Aquila Merchant Servs., Inc., 125 FERC ¶ 61,175 at ¶ 17 (2008).

86 *See infra* Section II. C–D.

87 PPL EnergyPlus, L.L.C. v. Nazarian, 974 F.Supp.2d 790, 801, 825–26 (D. Md. 2013), *aff'd*, 753 F.3d 467 (4th Cir. 2014), *aff'd sub nom.* *Hughes v. Talen Energy Marketing, LLC*, 136 S. Ct. 1288 (2016).

88 *Id.* at 824. Like the federal court in New Jersey at the same time (see *infra* at notes 44 and 102), this court cited the preemption holding of *Gade v. Nat'l Solid Wastes Mgmt. Ass'n*, 505 U.S. 88, 108 (1992). *Id.* at 854.

company could compete to build the plant as long as the facility itself was situated within Maryland.⁸⁹ They argued that the geographic situs of the commerce was overshadowed by the lack of geographic requirement for the owner of the in-state facility to be an in-state company.⁹⁰

The district court found the state Public Utilities Commission Order was preempted under the theory of field preemption.⁹¹ The trial court determined that Maryland's CfD, which required local utilities to enter into long term PPAs, was an impermissible intrusion of state regulation into wholesale rates, disrupting FERC-approved wholesale power markets.⁹² The court assessed whether the CfD mechanism impermissibly set wholesale prices for the regulated state utilities' energy and capacity power sales into the regional PJM markets.⁹³ The court concluded that when Maryland even indirectly manipulates the prices of wholesale power markets, the utilities, and correspondingly Maryland ratepayers, are directly affected by the resulting wholesale prices determined on the federally regulated wholesale PJM markets.⁹⁴ The court held that the Maryland regulation violates the Supremacy Clause of the United States Constitution by virtue of field preemption:⁹⁵ "The doctrine of field preemption forecloses state regulation in a field occupied entirely by the federal government, even if the state's purpose is admirable or the state regulation does not conflict with achievement of the federal scheme."⁹⁶

89 *Id.* at 853. The in-state requirement doesn't limit others' abilities to sell into the PJM market, so *Pike* discrimination was not shown. *Id.* at 854.

90 *Id.* at 853.

91 PPL EnergyPlus, L.L.C. v. Nazarian, 974 F.Supp.2d 790, 840 (D. Md. 2013), *aff'd*, 753 F.3d 467 (4th Cir. 2014), *aff'd sub nom.* Hughes v. Talen Energy Marketing, L.L.C., 136 S. Ct. 1288 (2016).

92 *Id.* at 834–35, 840 (D. Md. 2013). The court was persuaded in part by expert testimony explaining that the CfD went beyond a mere financing arrangement because it reflected the same factors typically used to establish rates and dictated the manner in which CPV (the winning bidder) could participate in PJM markets. *Id.*

93 *Id.* at 835, 837 ("The court does not find that the lack of physical delivery of energy between the parties to the CfD. . . insulates the [state order] from a field preemption attack.").

94 *Id.* at 795. The trial court cites the *West Lynn* case: "Such state economic protectionism 'violates the principle of the unitary national market by handicapping out-of-state competitors.' *W. Lynn Creamery, Inc. v. Healy*, 512 U.S. 186, 193 (1994)." *Id.* at 841.

95 *Id.* at 840.

96 *Id.* at 840. See *Arizona v. United States*, 132 S. Ct. 2492, 2502 (2012).

While Maryland stated its purpose in enacting the regulation was to cause the construction of sufficient reliable electric energy for and within the state of Maryland, the court held that no rationale permits a state to cross the “bright line” limiting jurisdiction or “invasion into a federally occupied field.”⁹⁷ States cannot dictate the ultimate price received for wholesale capacity sales in the PJM markets under the FPA and the Supremacy Clause:⁹⁸

[T]he PSC’s objective certainly fell within that traditional state purview continually referenced by Defendants, but the manner in which the PSC accomplished that objective involved establishing the amount received by CPV for its wholesale activity in the PJM Markets. Regulating in the field of wholesale price-setting is occupied by FERC, so therefore the strong presumption against preemption is not present.⁹⁹

The federal district court held that the FERC wholesale market field-preempts the Maryland scheme.¹⁰⁰ The court reasoned that the Maryland CfD prices have the effect of setting the ultimate price that the selected Maryland generator ultimately receives in the PJM market.¹⁰¹ The district court held “that the CfD payments had the effect of setting the ultimate price that CPV receives for its sales in the PJM auction, thus intruding on FERC’s exclusive authority to set interstate wholesale rates.”¹⁰² The court found field preemption,

97 *PPL EnergyPlus, L.L.C.*, 974 F. Supp. 2d at 828–30 (“Where a state action falls within a field Congress intended the federal government alone to occupy, the good intentions and importance of the state’s objective are immaterial to the field preemption analysis. Field preemption requires the state to ‘yield to the force of federal law . . . , notwithstanding that [the state’s action] is constructed upon values familiar to many and cherished by most, and notwithstanding that it may fit neatly within or alongside the federal scheme.’ See *French v. Pan Am Exp., Inc.*, 869 F.2d 1, 6 (1st Cir. 1989)”).

98 *Id.* at 829–30, 840. The state regulation compensates for delivery of energy and capacity to PJM in the PJM markets, “which provides further evidence that the CfD is not a purely financial contract generally considered to be outside FERC’s jurisdiction.” *Id.* at 837.

99 *Id.* at 840, n.60. See *United States v. Locke*, 529 U.S. 89, 108 (2000).

100 *PPL EnergyPlus, L.L.C.*, 974 F. Supp. 2d at 829.

101 *Id.* at 833.

102 *PPL EnergyPlus, L.L.C. v. Nazarian*, 753 F.3d 467, 474 (4th Cir. 2014) *aff’d sub nom.* *Hughes v. Talen Energy Marketing, L.L.C.*, 136 S. Ct. 1288 (2016).

and therefore chose not to reach whether there also was conflict preemption prohibiting the state's action. The court avoided finding a violation of the dormant Commerce Clause by making several key subjective determinations.¹⁰³ Similarly, a federal district court in New Jersey reached a determination of unconstitutionality of state energy regulation based on the Supremacy Clause on similar facts of state regulation of location.¹⁰⁴

D. The Unanimous Federal Circuit Court Opinion

On Maryland's appeal to the Fourth Circuit, the U.S. Solicitor General, in an amicus brief and at oral argument, said that FERC's wholesale electric rates preempt a different rate set by the Maryland Public Service Commission order being challenged.¹⁰⁵ The Fourth Circuit Court of Appeals unanimously affirmed and held that the Maryland program was "field preempted because it functionally set [] the rate that CPV receive[d] for its sales in the [wholesale] auction."¹⁰⁶ This was underscored because the contract payments by Maryland's order to its regulated utilities were conditional on CPV, the chosen power generator, entering and clearing (winning) the wholesale PJM capacity auction. This determined that the payments:

[p]lainly qualify as compensation for interstate sales at wholesale, not simply for CPV's construction of a plant . . . [and the state order] ensures . . . that CPV receives a fixed sum for every unit of capacity and energy [it sold in the wholesale auction, and] [t]he scheme thus effectively supplant[ed] the rate

103 *PPL EnergyPlus, L.L.C.*, 974 F. Supp. 2d at 841 (quoting *West Lynn Creamery, Inc. v. Healy*, 512 U.S. 186, 193 (1994)). The district court chose to apply the *Pike* balancing test and determined that because Maryland only provided incentives for location of facilities in the state, but did not restrict whether the electricity output of these facilities was sold in-state or out-of-state, it did not unduly burden interstate commerce. The statute was found to restrict the location of power facilities in-state, but not to restrict either facially or in its practical effect their commercial sales in commerce. *Id.* at 854-55.

104 *See id.*; *PPL EnergyPlus, L.L.C. v. Hanna*, 977 F. Supp. 2d 372 (D. N.J. 2013), *aff'd sub nom. PPL EnergyPlus, L.L.C. v. Solomon*, 766 F.3d 241 (3d Cir. 2014).

105 The petitioners arguing the case included the Maryland Public Service Commission and CPV Maryland; the respondents included Talen Energy Marketing, and the U.S. Solicitor General on behalf of FERC. Rebecca Kern, *Justices Appear Convinced State Subsidy Enters FERC Turf*, BLOOMBERG BNA (Feb. 25, 2016), <https://www.bna.com/justices-appear-convinced-n57982067793/>.

106 *PPL EnergyPlus, L.L.C.*, 753 F.3d at 476.

generated by the auction with an alternative rate preferred by the state.¹⁰⁷

Relying on “[a] wealth of case law [that] confirms FERC’s exclusive power to regulate wholesale sales of energy in interstate commerce,”¹⁰⁸ the Circuit Court found that “if FERC has jurisdiction over a subject, the States cannot have jurisdiction over the same subject.”¹⁰⁹ The court also noted that the Supreme Court made certain federal authority plenary and thus impliedly field preempted for states:

Congress meant to draw a bright line easily ascertained, between state and federal jurisdiction. . . . This was done in the [FPA] by making [FERC] jurisdiction plenary and extending it to all wholesale sales in interstate commerce except those which Congress has made explicitly subject to regulation by the States.¹¹⁰

The Fourth Circuit reasoned that allowing Maryland’s scheme to stand “compromises the integrity of the federal scheme and intrudes on FERC’s jurisdiction.”¹¹¹ As determined by the Fourth Circuit, the Maryland program’s effect “on matters within FERC’s exclusive jurisdiction is neither indirect nor incidental. Rather, [it] strikes at the heart of the agency’s statutory power to establish rates for the sale of electric energy in interstate commerce”¹¹² The Circuit Court affirmed the trial court holding of field preemption.¹¹³

A state even indirectly crossing the state-federal line is enough to make the action invalid under field preemption. The Fourth Circuit Court of Appeals noted that “if FERC has jurisdiction over a subject, the States cannot have jurisdiction over the same subject.”¹¹⁴

107 *Id.*

108 *Id.* at 475.

109 *Id.* at 476 (quoting *Miss. Power & Light Co. v. Mississippi ex rel. Moore*, 487 U.S. 354 (1988)).

110 *Id.* at 475 (alteration in original) (citing *Nantahala Power & Light Co. v. Thornburg*, 476 U.S. 953, 966 (1986)).

111 *Id.* at 476.

112 *Id.* at 478.

113 *Id.* at 480.

114 *Id.* at 475 (citing *Miss. Power & Light Co. v. Miss. ex rel. Moore*, 487 U.S. 354, 377 (1988) (Scalia, J., concurring)).

The Fourth Circuit decision notes that in recent times, the role of federal regulation has become increasingly prominent.¹¹⁵ No state is compelled to confront this line. It depends in significant degree on whether a particular state and its utilities elect to participate in an ISO subject to federal jurisdiction, or declines to do so. FERC has the authority to approve all Regional Transmission Organizations (“RTO”) and ISO terms of service and the financial tariffs.¹¹⁶ The Circuit Court stressed that Maryland chose to abandon its prior state model “and throw in its lot with the federal interstate markets . . . effectively compelling” participation in the interstate market.¹¹⁷ Because there is a “bright line” between state and federal authority, as more shifts to the area of exclusive federal regulation, the application of state regulation is displaced.

This state election is an important distinction. No state is compelled to participate in an integrated interstate market, and many states do not. Figure 1 below illustrates that approximately half of the states have not committed to participation in an interstate independent system operator wholesale power market.¹¹⁸ Because there is a “bright line” between state and federal authority, as more power sales shift to the area of exclusive federal regulation, the application of state regulation is displaced. Moreover, the federal courts have found even indirect state “external tampering” with these transactions to be impermissible.¹¹⁹

While traditional presumption against preemption of state authority exists in some areas, it does not for electric power markets. The Circuit Court of Appeals in the Maryland *Nazarian* opinion noted:

However, the presumption “is not triggered when the State regulates in an area where there has been a history of significant federal presence.” The presumption “is almost certainly not applicable here because the federal government has long regulated wholesale electricity rates.” Nevertheless, even were we to apply the presumption, we would find it

115 *Id.* at 472 (citing *New York*, 535 U.S. 1, 7 (2002)).

116 STEVEN FERREY, THE NEW RULES: A GUIDE TO ELECTRIC MARKET REGULATION, *supra* note 30, at 49–50.

117 *PPL EnergyPlus, L.L.C.*, 753 F.3d at 473.

118 See Regional Transmission Organizations (RTO)/ Independent System Operators (ISO), FERC (Dec. 21, 2017), <https://www.ferc.gov/industries/electric/indus-act/rto.asp>.

119 See *supra* text accompanying note 109.

overcome by the text and structure of the FPA, which unambiguously apportions control over wholesale rates to FERC.¹²⁰

With field preemption, which the FPA and the Supremacy Clause extend to the wholesale energy market:

The federal scheme thus “leaves no room either for direct state regulation of the prices of interstate wholesales of [energy], or for state regulations which would indirectly achieve the same result.” “Even where state regulation operates within its own field, it may not intrude indirectly on areas of exclusive federal authority.”¹²¹

The PJM market is regarded by the Circuit Court as a federally regulated interstate and wholesale market.¹²² The Court of Appeals in the PJM Maryland decision concluded that the FERC-administered PJM market is “a finely wrought scheme . . . carefully calibrated to protect a host of competing interests. It represents a comprehensive program of regulation that is quite sensitive to external tampering.”¹²³

The Fourth Circuit took a major step further than the trial court in applying the preemption doctrine by also finding that the Maryland program was conflict preempted because it (a) set CPV’s wholesale rate and (b) guaranteed it a fixed price for more than the federal PJM three-year period.¹²⁴ The Fourth Circuit concluded that conflict preemption reinforces a finding of field preemption, because the Maryland scheme “presents a direct and transparent impediment to the functioning of the PJM markets . . .”¹²⁵

In terms of finding a state statute preempted, it makes no ultimate difference whether it is disqualified on a single ground, or on both field preemption and conflict preemption grounds. Either

120 *PPL EnergyPlus, L.L.C.*, 753 F.3d at 477 (citations omitted).

121 *Id.* at 475–76 (citations omitted).

122 *Id.* at 472.

123 *Id.* at 473.

124 *Id.* at 479.

125 *Id.* at 480. The court suggested that a conflict preemption holding is narrow and to the particular case circumstances, and “not every state regulation that incidentally affects federal markets is preempted.” *Id.* at 479.

way, it is unconstitutional.¹²⁶ While the court again noted that “not every state regulation that incidentally affects federal markets is preempted,” it found the Maryland program “a bridge too far.”¹²⁷

III. Dissecting a Supreme Court Unanimous 2016 Decision

This matter was important enough to be reviewed on petition for *certiorari* by the Supreme Court, where most petitions for *certiorari* do not advance to the Supreme Court.¹²⁸ In 2016, the Supreme Court upheld the Fourth Circuit opinion, and did so unanimously on what had often been a closely divided Court.¹²⁹ The Supreme Court found that the Maryland statute intrudes on exclusive FERC wholesale market authority:¹³⁰ “Maryland’s program sets an interstate wholesale rate, contravening the FPA’s division of authority between state and federal regulators.”¹³¹ The Court noted that PJM operates both a capacity market and energy markets for the actual power. This case contested only the former – in which the project had been awarded a winning entitlement – and, while it could not enter the latter to participate in the power market until after it was constructed and operating, the Maryland CfD *could* affect the market for both.¹³²

The Court distinguished the Maryland contract for differences from more common bilateral contracts for power sale in subtle, but important, ways: “The contract for differences does not transfer ownership of capacity from one party to another outside the auction. Instead, the contract for differences operates within the auction; it mandates that LSEs and CPV exchange money based on the cost of CPV’s capacity sales to PJM.”¹³³ This distinction is important as it demarcates and reinforces the line over which the state illegally invades the exclusive field of federal energy jurisdiction.

126 PPL EnergyPlus, L.L.C. v. Nazarian, 974 F. Supp. 2d 790, 841 (D. Md. 2013), *aff’d*, 753 F. 3d 467 (4th Cir. 2014) *aff’d sub nom.* Hughes v. Talen Energy Marketing, L.L.C., 136 S. Ct. 1288 (2016). See also PPL EnergyPlus, L.L.C. v. Hanna, 977 F. Supp. 2d 372, 405-10 (D. N.J. 2013), *aff’d sub nom.* PPL EnergyPlus, L.L.C. v. Solomon, 766 F.3d 241 (3d Cir. 2014).

127 *PPL EnergyPlus, L.L.C.*, 753 F.3d at 479–80.

128 *The Justices’ Caseload*, SUP. CT. U.S., <https://www.supremecourt.gov/about/justicecaseload.aspx> (cert granted in about 80 cases out of 7,000-8,000 cases filed).

129 Hughes v. Talen Energy Mktg., L.L.C., 136 S. Ct. 1288, 1297 (2016).

130 *Id.* at 1298.

131 *Id.* at 1297.

132 *Id.* at 1290.

133 *Id.* at 1299.

A. *Reaffirmation of the Traditional ‘Bright Line’ of Field Preemption*

For the future of power and law, it matters whether the Supreme Court affirmed the field preemption found by both courts below, or acted on other grounds. There is a significantly different scope of the restriction of state authority depending on whether conflict preemption or field preemption is found. Both the majority and concurring Court opinions together “bracket” the legal foundation for this important Supremacy Clause decision.

First, the Court relied on its prior field preemption precedent. The Supreme Court identifies that the Maryland program presents the same legal constitutional problems that the Court identified three decades before in two seminal opinions, *Mississippi Power & Light*, and *Nantahala Power & Light Co.*¹³⁴ In *Nantahala Power & Light Co. v. Thornburg*, the Supreme Court found field preemption of state retail rate regulation that conflicts with wholesale power rates approved by FERC.¹³⁵ In reversing the court of appeals, the Supreme Court applied the “filed rate doctrine” enforcing the Supremacy Clause.¹³⁶ The Court held that “FERC clearly has exclusive jurisdiction over the rates to be charged [] interstate wholesale customers.”¹³⁷ The Court held that preemption under the “filed rate doctrine” is not “limited to ‘rates’ per se” but also includes practices that “directly affect[]” wholesale rate jurisdiction.¹³⁸

The Supreme Court notes that the Fourth Circuit opinion, which the Court affirms, relied on the Supreme Court’s *Mississippi Power* decision from 1988.¹³⁹ It then chooses to rely on this same opinion and the earlier consistent opinion in *Nantahala*.¹⁴⁰ For the current 2016 deregulatory environment, the Supreme Court cites both of these cases without any qualification as to their continued relevance in designating unconstitutional state regulation actions in

¹³⁴ *Id.* at 1288–90.

¹³⁵ *Nantahala Power & Light Co. v. Thornburg*, 476 U.S. 953, 966 (1986) (“plenary” jurisdiction of FERC concerning interstate wholesale rates; not limiting state role regarding prudence of buyer’s decision to purchase power).

¹³⁶ *Id.* at 963.

¹³⁷ *Id.* at 966.

¹³⁸ *Id.* at 966–67 (citing *Northern Natural Gas Co. v. Kansas Corporation Comm’n*, 372 U.S. 84, 90–91 (1963)).

¹³⁹ *Hughes*, 136 S. Ct. at 1296–97.

¹⁴⁰ *Id.* at 10, 13–14.

the current power market.¹⁴¹

These are more than casual references by the Court in its opinion; these are two pivotal Supreme Court opinions which created and reaffirmed the immutable “bright line” field preemption between state and federal authority with respect to electric power in America.¹⁴² In each case, state regulation affecting the wholesale power sector was overturned as being preempted and unconstitutional.

Of note, the Supreme Court decided each of these two cases before any state retail deregulation or utility asset divestiture began in the late 1990s,¹⁴³ and before there was any significant amount of wholesale power sales; wholesale sales now characterize the current electric power environment.¹⁴⁴ The current era of state deregulation of retail power in 17 states is depicted in Figure 1.¹⁴⁵ In 2014, what the U.S. Information Administration terms “independent power producers”¹⁴⁶ generated nearly forty percent of U.S. electricity, an increase of almost 400% from ten percent two decades earlier.¹⁴⁷ If this legal “bright line” applied when wholesale sales were a minor aspect of the power industry, the precedent becomes even more relevant when the wholesale power sales, both from merchant power plants and wholesale sales between utilities, are now a dominant aspect of power markets.¹⁴⁸

141 *Id.* at 13–14.

142 While the Supreme Court opinion also cites some of their decisions pursuant to the Natural Gas Act for general principles of federal exclusive authority, the key element is that they reinforce unchanged the two cases which established the “bright line” of exclusive federal authority in the electric market, which state regulation cannot cross in any regard. This is a reaffirmation, unchanged, of that existing “bright line.” The Court does this in the body and in the legal conclusion of its decision. *Id.* at 1297–99.

143 STEVEN FERREY, *supra* note 37, § 10:1, at 10-10 to -13.

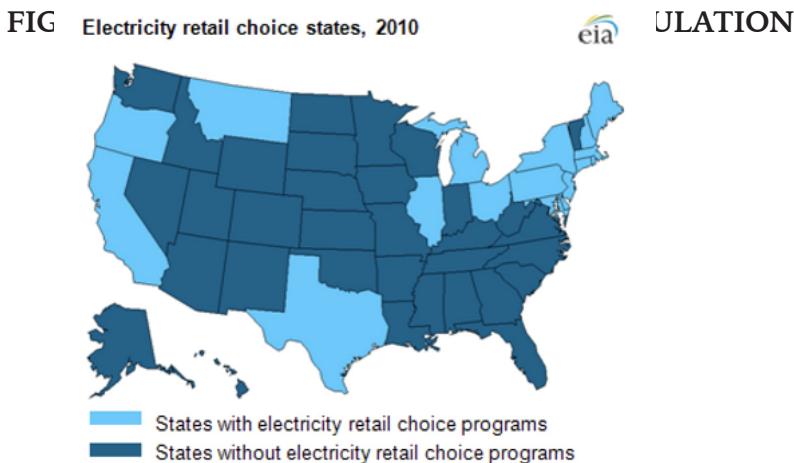
144 See STEVEN FERREY, *supra* note 7, at 608.

145 *Map of Deregulated Energy Market (Updated 2017)*, ELECTRIC CHOICE, <https://www.electricchoice.com/map-deregulated-energy-markets/> (last visited Oct. 28, 2017).

146 See U.S. ENERGY INFO. ADMIN., *Electric Power Monthly*, at Tbls. 1.2–1.5, http://www.eia.gov/electricity/monthly/current_year/august2015.pdf.

147 Federal Energy Regulatory Commission, Order No. 888, 61 Fed. Reg. 21,540, 21,550 (1996).

148 See *id.* at 21,550; Severin Borenstein & James Bushnell, *The U.S. Electricity Industry After 20 Years of Restructuring*, 7 ANN. REV. ECON. 437, 443–44 (2015) (“In 1997 only 1.6% of U.S. electricity was produced by generation owned by firms classified as Independent Power Producers. That figure rose to 25% by 2002 and was just under 35% in 2012. The share of nuclear generation owned by IPPs rose from zero in 1997 to almost 50% in 2012.”).



Second, while some Court observers forecast that this decision might blur or erase the line between state and federal jurisdiction over power,¹⁴⁹ the opinion reinforces the traditional line between exclusive federal wholesale and state retail authority. The opinion affirms what prior articles articulated at the beginning of the

¹⁴⁹ See Joel B. Eisen, *FERC v. EPSC and the Path to a Cleaner Electricity Sector*, 40 HARV. ENVTL. L. REV. F. 1, 9 n.60 (2016) (summarizing position of observers who state that a bright line split is no longer workable, mentioning Joel B. Eisen, *FERC's Expansive Authority to Transform the Electric Grid*, 49 U.C. DAVIS L. REV. 1783, 1789-90 (2016)); Jim Rossi, *The Brave New Path of Energy Federalism*, 95 TEX. L. REV. 399 (2016); Robert R. Nordhaus, *The Hazy "Bright Line": Defining Federal and State Regulation of Today's Electric Grid*, 36 ENERGY L.J. 203 (2015); Sharon Jacobs, *Bypassing Federalism and the Administrative Law of Negawatts*, 100 IOWA L. REV. 885, 940-41 (2015); Hannah J. Wiseman, *Moving Past Dual Federalism to Advance Electric Grid Neutrality*, 100 IOWA L. REV. BULL. 97, 99 (2015).

current millennium,¹⁵⁰ at the beginning of the current decade,¹⁵¹ and when Maryland was taking these now stricken steps,¹⁵² that state techniques, apart from indirect manipulation of wholesale prices, were permissible mechanisms for renewable energy incentives:

We therefore need not and do not address the permissibility of various other measures States might employ to encourage development of new or clean generation, including tax incentives, land grants, direct subsidies, construction of state owned generation facilities, or re-regulation of the energy sector . . . [that are] “untethered to a generator’s wholesale market participation.”¹⁵³

Merely stating that there are other ways for states to provide incentives for certain forms of power generation does not change the traditional FPA “bright line” or alter its contours in any manner. States have always been able to provide subsidies with state funds, as long as they do not commandeer or affect the operation of wholesale power markets.¹⁵⁴ Articulating that states cannot indirectly impede federally regulated wholesale markets excludes the entire wholesale market as “bright line” field preempted, rather than announcing that it is permissible unless there is a particular conflict with some detail, resulting in conflict preemption. It has always been permissible for states to provide in-state power generation facilities a tax break,¹⁵⁵

150 See e.g., Steven Ferrey, *Renewable Subsidies in the Age of Deregulation*, PUB. UTIL. FORT., Dec. 1997, at 22; Steven Ferrey, *Exit Strategy: State Legal Discretion to Environmentally Sculpt the Deregulating Electric Environment*, 26 HARV. ENVTL. L. REV. 109 (2002); Steven Ferrey, *Sustainable Energy, Environmental Policy, and States' Rights: Discerning the Energy Future Through the Eye of the Dormant Commerce Clause*, 12 N.Y.U. ENVTL. L. J. 507 (2004).

151 Steven Ferrey, *Carbon and the Constitution: State GHG Policies Confront Federal Roadblocks*, PUB. UTIL. FORT., April 2009; Steven Ferrey, *Goblets of Fire: Potential Constitutional Impediments to the Regulation of Global Warming*, 35 U.C. BERKELEY ECOLOGY L. Q. 835 (2008); Steven Ferrey, Chad Laurent & Cameron Ferrey, *Fire and Ice: World Renewable Energy and Carbon Control Mechanisms Confront Constitutional Barriers*, 20 DUKE ENVTL. L. & POL'Y J. 125 (2010).

152 Steven Ferrey, *Pentagon Preemption: The 5-Sided Loss of Energy and Power*, ILL. J.L. TECH., & POL'Y 393, 413–14 (2014).

153 *Hughes v. Talen Energy Mktg.*, L.L.C., 136 S. Ct. 1288, 1299 (2016).

154 FERREY, *supra* note 37, at § 10:114 (44th ed. 2017) (a detailed treatment of legal system benefit charges to support particular state power choices).

155 *Id.* at §§ 10:151, 10:78 to 90.

direct subsidies, which about a dozen states have done for decades,¹⁵⁶ or take advantage of the proprietary exception to construct facilities oneself.¹⁵⁷

In fact, six years before the Supreme Court decided *Hughes*, FERC struck a prior California regulatory enactment affecting wholesale power rates and, consistent with the *Hughes* decision six year later, held that California could have subsidized and provided financial incentives for the development of certain kinds of power generation facilities with more tax subsidies¹⁵⁸ or through certain renewable portfolio standards,¹⁵⁹ but not with direct or indirect alteration of the wholesale price of energy.¹⁶⁰ *Hughes* did not alter anything in terms of the bright line separating what is and is not allowed through state regulation. Wholesale power markets are not within state authority.

Third, the Court agrees that there is field preemption, consistent with all prior judges at both the federal trial and circuit levels rendering decisions on the *Hughes* matter. The Supreme Court opinion discusses both field preemption and conflict preemption.¹⁶¹ While the Court did not expressly state whether it relied on field or conflict preemption to find Maryland regulation impermissible, the language of the opinion suggests it determined both doctrines to be violated by the Maryland action:

FERC has approved the PJM capacity auction as the sole ratesetting mechanism for sales of capacity to PJM, and has deemed the clearing price per se just and reasonable. Doubting FERC's judgment, Maryland — through the contract for differences — requires CPV to participate in the PJM capacity auction, but guarantees CPV a rate distinct from the clearing price for its interstate sales of capacity to PJM. By adjusting an interstate wholesale rate, Maryland's program

156 *Id.* at § 10:114.

157 *Id.* at §§ 10:151, 10:78 to 90.

158 For information on tax subsidies, see STEVEN FERREY, *supra* note 38, at Tables 3.13, 3.15, 3.19.

159 Steven Ferrey, *Threading the Constitutional Needle with Care*, 7 TEX. J. OIL, GAS, & ENERGY L. 59 (2012).

160 See Order Granting Clarification and Dismissing Rehearing, 133 FERC ¶ 61,059 (Oct. 21, 2010).

161 *Hughes v. Talen Energy Marketing, L.L.C.*, 136 S. Ct. 1288, 1290 (2016).

invades FERC's regulatory turf.¹⁶²

The Court quotes FERC's prior determination that when one state has such a state subsidy program, it " . . . has the effect of disrupting the competitive price signals that PJM's [capacity auction] is designed to produce, and that PJM as a whole, including other states, rely on to attract sufficient capacity."¹⁶³ The Supreme Court articulates that "[w]e agree with the Fourth Circuit's judgment that Maryland's program sets an interstate wholesale rate, contravening the FPA's division of authority between state and federal regulators."¹⁶⁴ This is an original statement by the Supreme Court of conflict preemption by "contravening" the authority reserved exclusively for federal regulators. It also is a statement of field preemption established in a statutory "division of authority."¹⁶⁵

There does not need to be direct state regulatory interference: It is enough that a state regulation results in "adjusting" an area reserved as "FERC's regulatory turf."¹⁶⁶ This is language of both field preemption ("regulatory turf") and conflict preemption (where an adjustment changes the federally designed system). It also impliedly invokes field preemption where the Court states that "[t]he FPA leaves no room either for direct . . . or for regulation that would indirectly achieve the same result."¹⁶⁷ In addition to venturing into a constitutionally prohibited field, where that result is contrary to the federally mandated system, it is also conflict preemption.

One notes that the Court did not invalidate the Maryland program on the grounds that "it interferes with the capacity auction's price signals," or "counteracts FERC's refusal to extend the NEPA's duration."¹⁶⁸ Instead, the Court concluded that "[s]o long as a State does not condition payment of funds on capacity clearing the [wholesale] auction, the State's program would not suffer from the fatal defect that renders Maryland's program unacceptable."¹⁶⁹ It is not the fact that this involves interstate power sales that imperils the state action. As a more important fact, the Maryland program expressly tied contract payments to the generator entering and

162 *Id.* at 1297.

163 *Id.* at 1296.

164 *Id.* at 1297.

165 *Id.*

166 *Id.*

167 *Id.*

168 *Id.* at 1299 n.13.

169 *Id.* at 1299.

clearing the wholesale PJM interstate auction. State orders or policies that affect wholesale power markets or wholesale prices, directly or indirectly, cross the “bright line” of disallowed, field-preempted, state action.

This is important to note because the ability to transmit power is still through a monopoly utility, in both the transmissions and distribution systems. There is only one set of wires in the U.S., and they are owned by utilities, with a few exceptions.¹⁷⁰ This has implications for wider issues of the federalist division of authority, in which FERC regulates transmission transactions and the states regulate distribution transactions, which can occur sequentially over the same copper lines.¹⁷¹

It is the action of a regulatory order or approval of state energy regulatory agencies that implicates the vulnerability of state energy sector orders. Conflict preemption is narrower than field preemption: Field preemption rules off-limits an entire area; conflict preemption only bars state regulation which affects or interferes with federal regulation on a narrower case-by-case showing. This is implied strongly in both the language and the structure of the majority opinion.

Lest there be any ambiguity in what the majority opinion determines regarding a broad determination of field preemption, its author, Justice Kagan, at oral argument on this case stated:

I'm not sure why it is that when you say it was subject to FERC's jurisdiction, that doesn't end the case right

170 Operating merchant transmission projects in the United States include the Cross Sound Cable from Shoreham, New York to New Haven, Connecticut, Neptune RTS Transmission Line from Sayreville, N.J., to Newbridge, N.Y, and Path 15 in California. See *Transmission Lines*, U.S. SOLAR ENERGY INST., <https://ussolarinstitute.com/glossary/transmission-lines/> (last visited Sept. 17, 2017); see generally United States Electric Industry Primer, U.S. DEP'T OF ENERGY (July 2015) <https://www.energy.gov/sites/prod/files/2015/12/f28/united-states-electricity-industry-primer.pdf>; *The Project, NEPTUNE REG'L TRANSMISSION Sys.*, <http://neptunerts.com/the-project/> (last viewed Sept. 17, 2017).

171 See 18 C.F.R. § 35.28(a), (c) (2011); FERC Order 888 (establishing the foundation for the development of competitive bulk power markets. Promoting Wholesale Competition Through Open Access Non-discriminatory Transmission Services by Public Utilities); Recovery of Stranded Costs by Public Utilities and Transmitting Utilities, Order No. 888, FERC Stats. & Regs. ¶ 31,036, 61 Fed.Reg. 21,540 (1996). See *New York v. FERC*, 535 U.S. 1, 19-20, 28 (2002) (ratifying core provisions of Order 888).

there against you. . . [it is FERC's authority] to set the rates and other terms of wholesale sales, and that's not for the states to do. So that means you're preempted.¹⁷²

Such a statement of basic lines of authority delineates the field of exclusive federal jurisdiction. The Court concludes early in the opinion: "We affirm the Fourth Circuit's judgment."¹⁷³ It is important to note that the Fourth Circuit's judgment upheld the trial court's opinion that this state regulation unconstitutionally violated field preemption. The Court asserted the clearest reinforcement of the traditional "bright line" between state and federal electric sector jurisdiction. The Supreme Court conclusion did two things:

First, it characterized the state conditioning its subsidy on participation and successful result in the wholesale market as a "fatal" crossing of the line into exclusive federal authority.¹⁷⁴ This renders this type of relatively minor state interference an always lethal constitutional defect, not a case-specific anomaly.

Second, in its final sentence, the Court unanimously affirms the prior Fourth Circuit holding without qualification. That Fourth Circuit opinion itself unanimously upheld the trial court field preemption and added conflict preemption of the state action.

These are two separate defects; either one is constitutionally fatal.

B. *Concurring 'Brackets' on the Opinion Refine its Foundation*

While the *Hughes* opinion declares that the problem occurs where "a state does . . . condition payment of funds on capacity clearing the auction," there is even more precise context provided by the two concurring opinions written separately by Justices Thomas and Sotomayor. They act as "brackets" around the principal opinion of Justice Ginsburg for the Court. It is informative to consider each separately, to construe the precise context of the holding.

172 Rebecca Kern, *Justices Appear Convinced State Subsidy Enters FERC Turf*, BLOOMBERG BNA (Feb. 25, 2016) <https://www.bna.com/justices-appear-convinced-n57982067793/>.

173 *Hughes*, 136 S. Ct. at 1292.

174 *Id.* at 1299.

On one side of the Court majority opinion is Justice Sotomayor, stating that she thinks this line between state and federal authority is more of an interactive line on a case-by-case conflict basis.¹⁷⁵ On the other side of the majority opinion is Justice Thomas stating that the line between state and federal jurisdiction over electric power is express in the statute, and need not be implied by the Court to create its existence.¹⁷⁶ These two concurring opinions bracket what the majority Ginsburg opinion articulates when it reaffirms the *Mississippi Power* and *Nantahala Power* decisions of the Court with no alteration of the “bright line” separating state and federal authority.¹⁷⁷

First, Justice Sotomayor writes separately because, while she also finds the Maryland statute unconstitutional and preempted, she views the FPA as more interactive for states and FERC: “In short, the Federal Power Act, like all collaborative federalism statutes, envisions a federal-state relationship marked by interdependence.”¹⁷⁸ She thereafter cites some cases outside the electric energy sector.¹⁷⁹ Justice Sotomayor mistakenly refers to “the PJM . . . capacity auction” as the sale of “energy at wholesale,” which it, in fact, is not; rather, it is the sale of dedicated capacity without any energy included in any such transactions.¹⁸⁰ “Capacity’ is not electricity itself but the ability to produce it when necessary.”¹⁸¹

Justice Sotomayor’s concurring opinion, although she joined the opinion of the full Court,¹⁸² serves as one contextual end bracket on the opinion. Justice Sotomayor distinguishes her view from that of the full Court, which is not joined by any other member of the Court, that there is interdependence of federal and state regulation.¹⁸³ Justice Sotomayor is not defining a solid line between state and federal jurisdiction; she sees some degree of flexible, overlapping jurisdiction.¹⁸⁴ She would only find conflict preemption and not any field preemption where a state, on a case-by-case instance, takes actions which conflict with or veto the federal decision of FERC. At oral

175 *Id.* at 1300 (Sotomayor, J., concurring).

176 *Id.* at 1303 (Thomas, J., concurring).

177 *Id.* at 1288-90.

178 *Id.* at 1300 (Sotomayor, J., concurring).

179 *Id.*

180 *Id.*

181 *North Dakota v. Heydinger*, 825 F.3d 912 at n. 2 (8th Cir. 2016) (citing *Conn. DPUC v. FERC*, 569 F.3d 477, 479 (D.C. Cir. 2009)).

182 *Hughes*, 136 S. Ct. at 1300 (Sotomayor, J., concurring).

183 *Id.*

184 *Id.*

argument on the *Hughes* case, Justice Sonia Sotomayor identified this case as an example of conflict preemption, stating that “Maryland’s subsidy program intruded into FERC’s [PJM ISO] mechanism for setting wholesale electricity rates”; however, reserving judgment, Justice “Sotomayor said that she wasn’t convinced this case was necessarily an example of field preemption,” which would reinforce the “bright line” preempting “an entire area, or field, so it leaves no room for state regulation.”¹⁸⁵

Justice Sotomayor did not believe that the FPA establishes broad implied field preemption; she believes that there is only much narrower conflict preemption where a state regulation countermands the federal decision regarding wholesale power, which she finds, and on that basis joins the unanimous majority finding preemption. However, Justice Sotomayor voices an opinion not joined by any other members of the Court regarding wholesale power jurisdiction. All of the other Justices joined the majority opinion (or concurred with it) believing that the FPA establishes much broader and more defined field or express preemption, with Justice Sotomayor alone reaching the more restricted conclusion of conflict preemption.¹⁸⁶

Even in the absence of field preemption, federal law can still supersede state law based on conflict preemption if the state law interferes with a federal goal.¹⁸⁷ The majority opinion, in contrast to the concurring opinion of Justice Sotomayor, importantly underlines *Hughes*’ unanimous opinion that the “bright line” of jurisdictional separation, established previously by the Court and cited again in this opinion, is reaffirmed in both its brightness and its definitional distinction.

The second concurring opinion, written by Justice Thomas, forms a different end ‘bracket’ on the majority opinion of Justice Ginsburg. Justice Thomas states that the “bright line” distinction between state and federal jurisdiction over electric power need not even invoke, contrary to the Fourth Circuit’s holding, either implied field preemption or implied conflict preemption because there is express preemption in the FPA itself.¹⁸⁸ His concurring opinion also

¹⁸⁵ Rebecca Kern, *Justices Appear Convinced State Subsidy Enters FERC Turf*, BLOOMBERG BNA (Feb. 25, 2016), <https://www.bna.com/justices-appear-convinced-n57982067793/>.

¹⁸⁶ See *Hughes*, 136 S. Ct. at 1288.

¹⁸⁷ See *Hines v. Davidowitz*, 312 U.S. 52, 67 (1941) (State law will be preempted if it stands as an obstacle to the accomplishment and execution of the full purposes or objectives of Congress).

¹⁸⁸ *Hughes*, 136 S. Ct. at 1303 (Thomas, J., concurring at 1).

cites the same three decades old prior decisions of the Supreme Court in *Mississippi Power* and in *Nantahala Power* cited by Justice Ginsburg's principal opinion.¹⁸⁹ Justice Thomas specifically cites to the Court's opinion earlier in 2016 in *FERC v. EPSA* as also reinforcing a line that cannot be crossed even indirectly through "fiddling with the effective . . . price."¹⁹⁰

The theory and scope of this 2016 unanimous Supreme Court opinion are particularly important in establishing its legacy and precedent for Constitutional law. All nine Justices, as well as all four federal judges in the earlier decisions, found this state regulation to violate the Supremacy Clause of the Constitution.¹⁹¹ Eight Justices found any such regulation by the state to cross the bright line of field preemption, one Justice took a more limited view finding situational conflict preemption, while another Justice went further, articulating an explicit barrier of express preemption inherent in the FPA which distinguishes federal and state power over power. This reinforces the bright line excluding state regulation intruding into wholesale market transactions.

IV. Within the Brackets: How the Opposing Parties Framed Key Legal Issues for the Court to Decide

To fully understand whether the "bright line" of precedent was somehow impliedly erased or altered by this decision, one looks behind the unanimous decision of the Court to examine how the parties, on appeal, framed the legal issues for decision. From this analysis of what was submitted to the Court and how the Court framed the decision, there is no indication that the "bright line" separating exclusive fields of state and federal jurisdiction has been changed or weakened in any way.

The federal trial court determined that there was field preemption and struck the regulation.¹⁹² The Fourth Circuit Court of Appeals concurred that the state regulation was field preempted and also found that there was conflict preemption.¹⁹³ Field preemption

189 *Id.*

190 *Id.* For more on the *EPSA* case, see *infra* text accompanying note 304.

191 See Rebecca Kern, *Justices Appear Convinced State Subsidy Enters FERC Turf*, BLOOMBERG BNA (Feb. 25, 2016), <https://www.bna.com/justices-appear-convinced-n57982067793/>.

192 PPL EnergyPlus, L.L.C. v. Nazarian, 974 F.Supp.2d 790, 840 (D. Md. 2013), *aff'd*, 753 F.3d 467 (4th Cir. 2014), *aff'd sub nom.* Hughes v. Talen Energy Marketing, L.L.C., 136 S. Ct. 1288 (2016).

193 PPL EnergyPlus, L.L.C. v. Nazarian, 753 F.3d 467, 479–80 (4th Cir. 2014), *aff'd*

was the constant in the foundation of this decision, although both types of implied preemption were contested by appellants on appeal.

A. *Field Preemption*

As the Supreme Court approached preemption, it proceeded on the foundational decision of the Fourth Circuit. The Circuit Court sharpened the finding of field preemption relying on cases establishing and reinforcing the state-federal “bright line.”¹⁹⁴ Field preemption draws a bright line to exclude an entire area of regulation from any state regulation or authority:

Actual conflict between a challenged state enactment and relevant federal law is unnecessary to a finding of field preemption; instead, it is the mere fact of intrusion that offends the Supremacy Clause. “If Congress evidences an intent to occupy a given field, any state law falling within that field is preempted.”¹⁹⁵

The Fourth Circuit articulated that when an area is across a jurisdictional line, the entire field, regardless of state motive or the individual purpose of the state statute, is entirely preempted with no specific conflict with federal statute required to be demonstrated:

[T]he Supreme Court has expressly rejected the proposition that the “scope of [FERC’s] jurisdiction . . . is to be determined by a case-by-case analysis of the impact of state regulation upon the national interest.” Instead, “Congress meant to draw a bright line easily ascertained, between state and federal jurisdiction This was done in the [FPA] by making [FERC] jurisdiction plenary and extending it to all wholesale sales in interstate commerce except those which Congress has made explicitly subject to regulation by the States.”¹⁹⁶

The arguments of the parties frame whether this indirect

sub nom. *Hughes v. Talen Energy Marketing, L.L.C.*, 136 S. Ct. 1288 (2016).

194 *Hughes*, 136 S. Ct. at 1296–97.

195 *Nazarian*, at 753 F.3d 474 (citations omitted).

196 *Id.* at 475 (citations omitted).

state regulatory impact on wholesale markets is preempted by the FPA operating within the requirements of the Supremacy Clause of the Constitution. The Fourth Circuit unanimously upheld the challenge of the parties who were the appellees before the Supreme Court.¹⁹⁷ First, we examine the argument of the state of Maryland appellants.

1. Maryland Appellants

The briefs of the appellants positioned lack of any implied preemption as the primary issue for determination.

a. State of Maryland, Appellant/Petitioner

Led by chairmen of the Maryland Public Utilities Commission, the petitioners for the state of Maryland bifurcated the preemption issues as: (1) whether the FPA field preempts a state order directing retail utilities to enter into the contract, and (2) whether FERC's approval of an "annual [PJM] regional capacity auction preempt states from requiring retail utilities to contract at fixed rates with sellers who are willing to commit to sell into the auction on a long-term basis?"¹⁹⁸ Specifically, Maryland contended in their brief:

Congress implicitly [did not] foreclose[] states from doing what Maryland did: decide that it needed new generation facilities, conduct a competitive procurement, and direct state-regulated utilities to accept the best proposal by a generation developer to build a new facility and offer its capacity and energy to a regional market for twenty years.¹⁹⁹

In supporting this stance against field preemption, Maryland pointed to the "FPA's text, statutory structure, and forty years of [the] Court's precedent" decisions.²⁰⁰ Maryland argued that FERC's "wholesale rate authority establishes the contours of the federal field and the test for whether a state infringes it."²⁰¹ The petitioners then went on to declare that "so long as a state neither controls a FERC-

197 *See id.*

198 *See Brief for Petitioner Hughes at i, Hughes v. Talen Energy Mktg., L.L.C., 136 S. Ct. 1288 (2016) (No. 14-614).*

199 *Id.* at 24.

200 *Id.*

201 *Id.*

jurisdictional seller's decision whether and on what terms to sell, nor determines or second-guesses FERC's determination whether that rate is reasonable, the state remains on its side of the line."²⁰²

While conceding in their argument that the Fourth Circuit struck down Maryland's directive as field *and* conflict preempted by the FPA, the petitioners argued that the Circuit Court's decision should simply be reversed.²⁰³ Maryland argued that the Supreme Court should reverse because their state actions did not violate the statute, nor impede on FERC's turf or statutory authority.²⁰⁴ Moreover, Maryland contended that what they did was constitutionally sound because they merely "directed state-jurisdictional utilities to accept a willing seller's offer."²⁰⁵

The state petitioners devoted eight pages in their brief to arguing that Maryland's procurement was not field preempted.²⁰⁶ In their argument against field preemption, Maryland took issue with the Fourth Circuit's order finding field preemption because it "functionally set[] the rate that CPV receives for its sales in the PJM auction."²⁰⁷ The state argued that the Fourth Circuit did not explain how Maryland's regulation of its distribution utilities "functionally" set CPV's proposed rate.²⁰⁸ Maryland's brief continued to challenge the Fourth Circuit decision, arguing that the Circuit may have applied the wrong test for field preemption, and instead more correctly needed to "allow the FERC-jurisdictional sellers to set their rates (including by entering contracts) and empower FERC to modify those it finds unlawful."²⁰⁹ Maryland argued that they were not field preempted because Maryland neither exercised nor constrained the exercise of the rate-setting and review authority of FERC.²¹⁰

Maryland also argued that their procurement was congruent with the FPA framework.²¹¹ Maryland outlined that "[w]hen the legislature enacted the FPA, it left jurisdictional sellers' rate-making and rate-changing powers 'unaffected,'" but did not require that

202 *Id.*

203 *Id.* at 2.

204 *Id.* at 25.

205 *Id.* at 25.

206 *Id.* at 26-33.

207 See *id.* at 26 (quoting Pet.App.19a).

208 *Id.* at 26-27.

209 *Id.*

210 *Id.*

211 *Id.* at 27-31.

sellers initiate the contracting process.²¹² The Maryland petitioners pointed to “decades of experience. . .and FERC’s own regulations” which suggest that “sellers may set rates by responding to competitive solicitations and (if they win) [may] enter[] into the resulting contracts.”²¹³ Accordingly, the state argued that what Maryland did was a function left entirely to the states.²¹⁴

More specifically, Maryland argued that once it determined that it needed more generation facilities using long-term contracting to obtain them, it had exercised inherent state powers to control in-state generation hardware which the FPA left to the states.²¹⁵ The Maryland petitioners further emphasized that the division of responsibilities between FERC and the states is congruent with the spirit of the FPA’s plain meaning and common sense.²¹⁶ The state petitioners stressed that Maryland is in the best position to determine the kind and amount of new generation capacity it needs (650-700 megawatts of new natural gas fired capacity) which is a traditional state function.²¹⁷ The petitioners cited the rule of law in *California Public Utilities Commission*,²¹⁸ as binding precedent allowing Maryland’s selection of the winning bidder and its discretion that Maryland regulated distribution utilities accept the winner’s bid.²¹⁹ The petitioners assembled these arguments to submit that Maryland did not set CPV’s wholesale power rate.²²⁰

The Maryland state petitioners additionally argued that the Fourth Circuit’s approach undermined the framework of the FPA.²²¹ They characterized the Fourth Circuit’s holding as an “extra-statutory” limit on FERC-jurisdictional power sellers.²²² Maryland conceded that FERC is charged with the reasonable development of electricity at reasonable prices, but “cannot order development of

212 *Id.* at 27–28 (citing *United Gas Pipe Line Co. v. Mobile Gas Serv. Corp.*, 350 U.S. 332, 343 (1956)).

213 *Id.* at 28 (citing 18 C.F.R. § 35.27, JA 909).

214 *Id.* at 29.

215 *Id.* at 28.

216 *See id.* at 29.

217 *Id.* at 30 (citing *New York v. FERC*, 535 U.S. 1, 24) (noting that Maryland used a wide variety of studies that included change in customer requirements, resource adequacy and various other econometrics).

218 S. California Edison Co., 134 FERC ¶ 61,044, para. 30 (2011) (order denying rehearing).

219 Brief for Petitioner Hughes at 30, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

220 *Id.* at 31.

221 *Id.* at 31–33.

222 *Id.* at 32.

new generation facilities directly, even when necessary to remedy inadequate interstate service.”²²³ Maryland argued that “FERC thus depends on voluntary generation development, which itself depends in turn on the sellers’ ability to enter long-term contracts providing revenue stability.”²²⁴ The state argued that as a consequence of the Fourth Circuit’s holding, “FERC-jurisdictional sellers may not enter into contracts with []parties that act on state orders,” and, if they do, they risk their contract being nullified “by a court at some future date.”²²⁵ Hence, Maryland argued in their brief that this circuit court holding discourages generation development and “diminish[es] the usefulness of such contracts.”²²⁶

b. CPV, Appellant/Petitioner

In its litigation position before the Supreme Court, the winning Maryland power producer, CPV, took the position that no form of implied preemption was present in the case-at-bar.²²⁷ Either type of preemption of the Maryland statute would cause CPV to not be successful on appeal of the case. CPV Maryland attempted to convince the Supreme Court that neither field preemption nor conflict preemption was present, both of which the Fourth Circuit had found.²²⁸ CPV addressed and outlined their positions against both field and conflict preemption in 27 pages of the CPV brief, a majority of its total pages.²²⁹ Specifically, CPV outlined four arguments against field preemption in seventeen pages of their brief,²³⁰ and dedicated ten additional pages to argue that there was no conflict preemption.²³¹ The bulk of the effort was to overturn the finding of field preemption, although petitioners argued against any implied preemption of Maryland regulation.

The petitioners argued that the Court has shown “reluctance to find field preemption based merely on an agency’s comprehensive regulatory framework, particularly where . . . the agency has [failed

223 See *id.* at 32. See also NAACP v. FPC, 425 U.S. 662, 670 (1976). E.g., 16 U.S.C. § 824(f) (2012).

224 Brief for Petitioner Hughes at 32, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

225 *Id.* at 32–33.

226 *Id.* at 33.

227 Brief for Petitioner CPV Maryland, LLC at 33, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

228 *Id.* at 47.

229 See *id.*

230 See *id.*

231 See *id.*

to] affirmatively state[] its intention to preempt.”²³² Additionally, they argued that because federal agencies usually make their regulations in a very detailed manner, FERC would therefore make their intentions clear if they wanted their regulations in a particular subset of the energy market to be exclusively federal and leave no state space in the field at issue.²³³ Field preemption, when present, prohibits any state regulation in the particular preempted field.

Residual powers. CPV also asserted in its brief to the Court that the FPA reserves certain powers to the states.²³⁴ They tried to position this argument to support that states have an important role in supporting construction of new generation where and whenever the state determines it to be appropriate. Furthermore, CPV noted that states retain authority over operations of their local retail utilities, pursuant to 16 U.S.C § 824e(a).²³⁵ CPV suggested that this led to the conclusion that FERC is acting outside its allowed realm of authority and a court must guard against an expansive interpretation of FERC’s authority that would undermine Maryland’s inherent state power.²³⁶

CPV argued that Maryland’s means of regulating utilities was within their state field of allowed jurisdiction under the FPA. They noted that nothing in the FPA limited the states’ authority “to direct the planning and resource decisions of utilities under their jurisdiction.”²³⁷ Maryland sought to rely on the Second Circuit opinion in *Entergy Nuclear Vt. Yankee, LLC v. Shumlin*.²³⁸ However, in respect to CPV’s citation to the Second Circuit *Shumlin* case for this

232 *Id.* at 29. See *Hillsborough Cnt., Fla. v. Automated Med. Labs., Inc.*, 471 US. 707, 717–718 (1985) (pointing to the precedent as standing for the proposition that “if an agency does not speak to the question of pre-emption”, the court “will pause before saying that the mere volume and complexity of its regulations indicate that the agency did in fact intend to preempt”).

233 See Brief for Petitioner CPV Maryland, L.L.C. at 30, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

234 See *id.* (standing for the argument that FERC does not have authority over new energy generation facility construction).

235 *Id.* at 31.

236 See *id.*; *Nw. Cent. Pipeline Corp. v. State Corp. Comm’n of Kan.*, 489 U.S. 493, 506–07 (1989) (noting that states and FERC each have broad authority within their assigned spheres).

237 Brief for Petitioner CPV Maryland, L.L.C. at 33, *Hughes*, 136 S. Ct. 1288 (No. 14-614). See *Entergy Nuclear Vt. Yankee, L.L.C. v. Shumlin*, 733 F.3d 393, 417 (2d Cir. 2013).

238 Brief for Petitioner CPV Maryland, L.L.C. at 33, *Hughes*, 136 S. Ct. 1288 (No. 14-614) (citing *Entergy Nuclear Vt. Yankee, L.L.C. v. Shumlin*, 733 F.3d 393, 417 (2d Cir. 2013)).

proposition of local planning authority, it should be noted that, in *Shumlin*, both the trial court and Second Circuit Court of Appeals held that federal preemption was present so as to strike the state energy regulation.²³⁹ Thus, when relying on *Shumlin* for the proposition that states can direct utility planning in the state, the *Shumlin* case simultaneously establishes that the state cannot do so in a way that affects directly or indirectly wholesale energy prices which would be preempted by the Constitution.

CPV argued that because it was “exercising its authority over the contracting decisions of its local utilities, [CPV] used its authority over retail sales to ensure that any rebate or subsidy if earned by the in-state generators would be passed on to retail ratepayers.”²⁴⁰ CPV also took the stance that they did not engage in rate setting reserved to FERC because Maryland only directed procurement through competitive bids.²⁴¹

Binding precedent. As to prior binding Supreme Court wholesale energy preemption precedents, the CPV petitioners attempted to find and identify distinctions. CPV asserted that the *Mississippi Power & Light* case cannot support the Fourth Circuit’s field preemption theory or any other preemption theory because in *Mississippi Power* and in *Nantahala* before the Supreme Court, “FERC had approved the appropriate wholesale price the utility paid to a generator [initially], and the State’s [subsequent] action fundamentally contradicted FERC’s [prior] order” on the same

239 Entergy Nuclear Vt. Yankee, L.L.C. v. Shumlin, 733 F.3d at 407, 422 (2d Cir. 2013).

240 Hughes v. Talen Energy Marketing, L.L.C., 136 S. Ct. 1288, 1296 (2016); Brief for Petitioner CPV Maryland, L.L.C. at 34-35, *Hughes*, 136 S. Ct. 1288 (No. 14-614) (“there is no basis to distinguish between supporting new generation by directing a utility to buy capacity under a long-term agreement, and achieving the same end by directing the utility to enter into a long-term [contract for differences] to further the same permissible goal of building new generation facilities in the State.”).

241 Brief for Petitioner CPV Maryland, L.L.C. at 38, *Hughes*, 136 S. Ct. 1288 (No. 14-614). Maryland made the argument that they were not rate setting because they did not dictate the price. Therefore, they did not “intrude on both the seller’s authority to determine the rate at which it chooses to sell and FERC’s authority to determine whether the rate is just and reasonable.” *Id.* Petitioners cited two cases in support of their “bright line” interpretation. See e.g., Allegheny Energy Supply Co., 108 FERC ¶ 61,082, PP 15, 20, 21 (2004) (approving state-supervised procurement where “[w]inning bidders received the actual price in their offers”); Conectiv Energy Supply, Inc., 115 FERC ¶ 61,199, PP 4-6, 20 (2006). See Doswell Ltd., 50 FERC ¶ 61,251, 61,756-59 (1990) (approving contract prices set by competition).

matter.²⁴² They argued that those ordinal temporal conditions were not present in Maryland.²⁴³ Moreover, they argued that unlike these Supreme Court precedents, “Maryland guaranteed that the local utilities would recover any [contract-for-differences]-related costs from their own [in-state] ratepayers.”²⁴⁴

In essence, this argument sought to convert the foundational Supreme Court decisions in *Mississippi Power* and in *Nantahala* from field preemption “bright line” decisions into conflict preemption cases limited to situations where FERC first sets a price and the state subsequently refuses to follow it after the fact, and not with other timing. CPV framed the legal issue for the Court specifically to make a particular determination that *Mississippi Power* and *Nantahala* were limited to their specific temporal facts, rather than foundational Supremacy Clause statements that a state could not take an action interfering with federal wholesale power authority.²⁴⁵ As discussed later, the Court did not either alter their constant “bright line” or recast or reconstruct the cases on which it relied, *Mississippi Power* and *Nantahala*, from field preemption to only narrow holdings of conflict preemption.

2. Respondent/Appellee Talen Energy Marketing

The appellants clearly framed and presented the context of the preemption issue to the Supreme Court. When asked directly by Justice Sotomayor during oral argument whether this is a case of conflict preemption or field preemption, Paul Clement, attorney for the challenger independent power generators, stated:

I think it's both, Your Honor. I think part of the reason this is difficult is because it's like an overdetermined equation. There is both field preemption here, but there's also conflict preemption, and there are multiple evidences of conflict preemption because nobody's tried to do anything like this before.²⁴⁶

242 Brief for Petitioner CPV Maryland, L.L.C. at 44, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

243 *Id.*

244 *Id.* (arguing that the ends which they are trying to achieve are good, just and legitimate). Of note, this argument was later contradicted in the respondents’ brief.

245 *Id.*

246 Transcript of Oral Argument, *Hughes*, 136 S. Ct. 1288 (No. 14-614), http://www.supremecourt.gov/oral_arguments/argument_transcripts/14-614_6536.pdf.

The respondents/challengers called into legal question Maryland's scheme based on the Court's state/federal "bright line" precedent in *Nantahala* and *Mississippi Power*.²⁴⁷ The successful respondents/challengers noted that while one could attempt to characterize key Supreme Court energy preemption precedent as conflict preemption or field preemption, it does not matter where the state action collaterally affects the function of wholesale power markets:

Yet this Court did not hesitate to conclude that the states' efforts to use retail rates to prevent utilities from recouping their full wholesale rates constituted impermissible incursions on FERC's "exclusive authority to determine the reasonableness of wholesale rates." Whether those cases are better viewed as field or conflict preemption cases is beside the point; either way, they confirm that even the powers expressly reserved to states by the FPA may not be used to "collaterally attack[]" FERC's regulation of wholesale sales.²⁴⁸

Wholesale power transactions are protected by the bright line segregating the exclusively federally regulated wholesale field from state retail regulation.²⁴⁹ CPV, the petitioner on appeal, tried to distinguish these two Supreme Court precedents²⁵⁰ because Maryland announced their plans to "top off" wholesale market prices to FERC ahead of time, unlike the prior precedent cases in which the state refused to pass through wholesale FERC rates into consumer retail rates after the wholesale rates had been set by FERC.²⁵¹ In other words, Maryland argued that who goes first matters and here, by going first, the state was due deference.

The respondents/challengers rebutted petitioners' emphasis

Cf. F.E.R.C. v. Elec. Power Supply Ass'n, 136 S. Ct. 760, 776 (2016) (noting this case, although not cited in the brief) ("that the wholesale and retail markets in electricity . . . are not hermetically sealed from each other").

247 Brief for Respondent Talen at 32-34, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

248 *Id.* at 33 (citations omitted).

249 See, 16 U.S.C. § 824(a) (2012).

250 See Brief for Petitioner CPV Maryland L.L.C. at 42-45, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

251 *Id.* at 33-34; *Hughes*, 136 S. Ct. at 1298.

on order of events by stating in their brief that an announcement in advance that one is going to violate the Supremacy Clause does not negate the Constitution. “[T]he relevant question is whether, not when, the state is attempting to override the FERC-approved rate for a wholesale sale.”²⁵² The respondent/challengers’ brief used the FPA to frame the type of preemption at issue, noting that 16 U.S.C. § 824(a), a provision of the FPA, sets forth a clear jurisdictional line demarcating federal government authority from that of the states.²⁵³ “FERC has exclusive authority to regulate all ‘rates and charges made, demanded, or received. . . for or in connection with’ interstate wholesale sales.”²⁵⁴ The challengers highlight that the Court has routinely established that for a half century “the FPA left ‘no power in the states to regulate . . . sales for resale in interstate commerce.’”²⁵⁵

The successful challengers argued that the traditional line between federal FERC and state authority had, and has, not changed.²⁵⁶ They noted that the FERC-approved PJM capacity auction is a mechanism through which FERC exercises its statutory authority to ensure that wholesale prices are “just and reasonable.”²⁵⁷ The challengers argued that the contested Maryland regulation is altering the final price of a sale to PJM and therefore, by definition, attempting to alter the wholesale rates approved by FERC.²⁵⁸ The challengers argued that by doing so, Maryland attempted to occupy the exclusive federal field because “the exclusive and pervasive federal regulation of wholesale rates is [very] clear” when it comes

252 Brief for Respondents at 33, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

253 Brief for Respondents at 27, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

254 *Id.*

255 *Id.* at 27. See *FPC v. S. Cal. Edison Co.*, 376 U.S. 205 (1964) “Congress meant to draw a bright line easily ascertained, between state and federal jurisdiction, making unnecessary . . . case-by-case analysis [of conflicts]. This was done in the Power Act by making FPC jurisdiction plenary and extending it to all wholesale sales in interstate commerce except those which Congress has made explicitly subject to regulation by the states.” (emphasis added) *Id.* at 215-16. See also, e.g., *New York v. F.E.R.C.*, 535 U.S. 1, 21 (2002); *Miss. Power & Light Co. v. Miss. ex rel. Moore*, 487 U.S. 354 (1988). See also, *Nantahala Power & Light Co. v. Thornburg*, 476 U.S. 953, 966 (1986) (“FERC clearly has exclusive jurisdiction over [wholesale rates]”).

256 See, Brief for Respondent at 24-25, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

257 *Id.* at 24-25, 27-28.

258 *Id.* at 23. The brief stresses that the Maryland initiative should be shut down because it was telling the local generators what they were to do in order to earn substantial subsidies from the state for signing the twenty-year contracts. *Id.* at 23, 55.

to wholesale sales of energy and capacity to wholesale-market operators such as PJM.²⁵⁹

The respondent/challengers attacked the constitutionality of Maryland's scheme for intruding across the unpermitted line demarcating federal jurisdictional field authority. They submitted that the federal government not only occupies the field for what PJM should pay for power capacity, but also what CPV "receive[s] . . . for or in connection with" its sales to PJM, which are undeniably wholesale sales.²⁶⁰ The successful challengers refuted Maryland's argument that this scheme was permitted by 16 U.S.C. § 824d(b) because it was done by Maryland with the intention to spur and incentivize new generation only if it located in Maryland. Respondents emphasized that:

[N]o matter how broad the states' authority to incentivize new generation may be, it does not encompass the power to regulate the 'rates and charges made, demanded, or received' by a generator 'for or in connection with' wholesale sales—let alone to alter the rates that FERC has already approved for those wholesale sales.²⁶¹

The respondents/challengers framed the issue as one of field preemption by noting that "even if the statute were less clear, there could be no serious dispute that the express grant of authority to FERC over the rate for a sale would preclude state authority to set a different rate for the exact same sale."²⁶² This is a fundamental principle in field analysis in energy. The federal government jurisdiction is so extensive regarding all wholesale power transactions, that there is no room impliedly left for the states even to indirectly regulate small areas not covered by the statute.²⁶³ The respondents note that

259 *Id.* at 27.

260 *Id.* at 35 (citing 16 U.S.C. § 824d(a)).

261 *Id.* at 30 (emphasis and citations omitted).

262 *Id.* at 30–31.

263 The federal energy regulatory scheme "leaves no room either for direct state regulation of the prices of interstate wholesales of [energy], or for state regulations which would indirectly achieve the same result." *N. Nat. Gas Co. v. State Corp. Comm'n of Kan.*, 372 U.S. 84, 91 (1963) (citation omitted). "Even where state regulation operates within its own field, it may not intrude indirectly on areas of exclusive federal authority." *Pub. Utils. Comm'n v. FERC*, 900 F.2d 269, 274 n. 2 (D.C. Cir. 1990) (internal quotation marks

this is an easier case than was *McCullough v. Maryland*; unlike that older precedent, here “Congress made clear that the states’ reserved authority over generation cannot displace FERC’s ability to set wholesale rates.”²⁶⁴

The respondent/challengers noted that Maryland’s argument regarding the generic nostrum that there is a general presumption against preemption.²⁶⁵ This presumption had not been successful at the prior level before the Fourth Circuit.²⁶⁶ The challengers rebut such a presumption here applied to the FPA by emphasizing that there is no ambiguity in the statute regarding respective roles on either side of the line of exclusive jurisdiction between the state and federal governments.²⁶⁷ The challengers argued that the traditional presumption against preemption to limit the extent of federal jurisdiction under the FPA should not apply here.²⁶⁸

The successful respondent/challengers countered the petitioner’s argument that the Court’s 2015 decision in *Oneok, Inc. v. Learjet, Inc.* suggests otherwise.²⁶⁹ The challengers stress that *Oneok* does not stand for permission for states to “enact measures that are deliberately designed to alter the rate received for fully consummated wholesale sales, so long as they are doing so in service of achieving ‘objectives within the State’s own jurisdictional field.’”²⁷⁰ They emphasized that the cases relied on in *Oneok* “expressly reject[ed] the notion that states may use their ‘undoubted jurisdiction over retail sales’ to [undermine] FERC’s regulation of wholesale sales.”²⁷¹ Of note, *Oneok* was a natural gas and not an electricity case. Natural gas is governed by a separate federal statute.²⁷² The “bright line”

omitted).

264 Brief for Respondent at 31, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

265 *Id.*

266 *See supra*, at Section II D.

267 Brief for Respondent at 31, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

268 *Id.*

269 *Id.* at 31–32 (“*Oneok* instead simply held that FERC’s exclusive authority over wholesale rates did not preempt application of state antitrust laws to an effort to manipulate wholesale rates when the state antitrust laws were ‘not aimed at natural-gas companies in particular,’ but instead had ‘broad applicability’ beyond the field of natural gas.”). *See Oneok, Inc. v. Learjet, Inc.*, 135 S. Ct. 1591, 1601. *See also* Brief for Petitioner at 32, *Hughes*, L.L.C., 136 S. Ct. 1288 (No. 14-614).

270 Brief for Respondent at 32, *Hughes*, 136 S. Ct. 1288 (No. 14-614) (citing Brief for Petitioner at 32).

271 *Id.* at 33.

272 *See* Natural Gas Act, 15 U.S.C. § 717 (2012).

separating electric power jurisdiction remains.

The respondents/challengers further rebutted Maryland's argument that the Maryland statute was motivated to incentivize the construction of new in-state generation, and that this should alter the preemption analysis:

Mississippi Power and Nantahala both involved state efforts to achieve ends that were perfectly legitimate under the FPA, via means—namely, the alteration of wholesale rates—that were just as plainly verboten. And a long line of this Court's cases extending well beyond those two precedents confirms that states cannot regulate wholesale rates for the best of reasons, the worst of reasons, or any reason in between Maryland has intruded upon FERC's exclusive jurisdiction over the "rates and charges made, demanded, or received . . . for or in connection with" interstate wholesale sales.²⁷³

3. *Amici Allco Renewable Energy in Support of Respondents*

Allco is a "New York-based renewable energy firm that develops and invests in solar [energy] projects."²⁷⁴ Allco, as amici in support of respondent, explicitly framed for the Court that Maryland is preempted because it is intruding on a field that Congress has designated for, and FERC has served with, exclusive federal regulation.²⁷⁵ Allco asserted that the Fourth Circuit correctly decided this case, and cited case law which previously established FERC's exclusive power to regulate wholesale sales of energy

273 Brief for Respondent at 34-35, *Hughes*, 136 S. Ct. 1288 (No. 14-614) (citations omitted).

274 *Investing in the Future of Energy*, ALLCO, www.allcous.com/allco/.

275 Brief of Amicus Curiae ALLCO Renewable Energy Ltd. in Support of Respondents at 17-19, *Hughes*, 136 S. Ct. 1288 (No. 14-614). See *Arizona v. United States*, 567 U.S. 387, 401 (2012) (pointing to the precedent as primary authority for field preemption) "[w]here Congress occupies an entire field. . . even complementary state regulation is impermissible." See *Silkwood v. Kerr-McGee Corp.*, 464 U.S. 238, 248 (1984) (also citing this case as further authority for defining the contours of field preemption) "If Congress evidences an intent to occupy a given field, any state law falling within that field is preempted." *Id.*

in interstate commerce.²⁷⁶ Allco also includes in a footnote their counterargument to the petitioner's position that the "presumption against preemption" should apply here, given the explicit facts present in this matter.²⁷⁷

Allco reinforced its field preemption argument by establishing that a bright-line rule was expressed in Section 201(b) (1) of the FPA.²⁷⁸ Allco argued that whether viewed as a power purchase agreement or a separate CFD, Maryland's action intruded into federal territory.²⁷⁹ They contended that the plain language of the statute sets a bright-line between a state's authority over its facilities and interstate wholesale sales.²⁸⁰ Thus, it concluded that, as in *Silkwood v. Kerr-McGee Corp.*, Maryland intruded into a "field of exclusive federal jurisdiction, and its action [was] pre-empted and

276 Brief of Amicus Curiae ALLCO Renewable Energy Ltd. in Support of Respondents at 17–19, *Hughes*, 136 S. Ct. 1288 (No. 14-614) (providing a landmark court decision upholding state's limitations in the field of wholesale sales of energy; "Congress meant to draw a bright line, easily ascertained, between state and federal jurisdiction.... This was done ... by making [FERC] jurisdiction plenary and extending it to all wholesale sales in interstate commerce except those which Congress has made explicitly subject to regulation by the States.") (citation omitted). See also New England Power Co. v. New Hampshire, 455 U.S. 331, 340 (1982) (citing an earlier decision that also reinforces the challengers position that the FPA prevents states from occupying the field of wholesale sales of electricity. "[FERC] [has] exclusive authority to regulate the transmission and sale at wholesale of electric energy in interstate commerce, without regard to the source of production.").

277 Brief of Amicus Curiae ALLCO Renewable Energy Ltd. in Support of Respondents at 18 n. 13, *Hughes*, 136 S. Ct. 1288 (No. 14-614) ("[A]nd thus the Federal Power Act does not displace the state's traditional police powers."). See *New York v. FERC*, 535 U.S. 1, 6 (citing the precedent for the proposition that wholesale electricity sales in the stream of commerce are not subject to state regulation), *United States v. Locke*, 529 U.S. 89, 108 (2000) (pointing to the precedent as establishing pervasive federal presence in an area).

278 Brief of Amicus Curiae Allco Renewable Energy Ltd. in Support of Respondents at 19–20, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

279 *Id.* at 19.

280 See *id.* at 19–20 (delineating the language in the statute which grants FERC exclusive jurisdiction over the field of wholesale electricity). Federal Power Act, 16 U.S.C § 824 (2012) ("[FERC] shall not have jurisdiction, except as specifically provided in this Part and the Part next following, over facilities used for the generation of electric energy or over facilities used in local distribution or only for the transmission of electric energy in intrastate commerce, or over facilities for the transmission of electric energy consumed wholly by the transmitter.")

the [associated Maryland power sale] contracts void.”²⁸¹

Allco emphasized that the plain language of the FPA extends to interstate wholesale sales and rules affecting or pertaining to rates and charges, which is more of a field preemption contention.²⁸² Also somewhat smacking of field preemption, they argue that “FERC’s goal of establishing a competitive market designed to meet demand at least cost would be frustrated,” if Maryland were to go forward with its scheme:

[Maryland] compelled the utility to enter a contract with the [Maryland’s] chosen winner, and thereby mandated a wholesale sale of electricity that would not have taken place absent the State’s compulsion. And this Court has held that the Federal Power Act “left no power in the states to regulate . . . sales for resale in interstate commerce.”²⁸³

Furthermore, Allco stressed that the FPA disallows “even the possibility of such state interference” in the field of wholesale sales.²⁸⁴

B. Conflict Preemption

Conflict preemption does not describe the jurisdictional areas demarcated by a jurisdictional “bright line” separating state and federal authority. Rather, conflict preemption occurs when a particular mechanism or legal structure conflicts with a superior regulation, even if such regulation was allowed as within basic state jurisdiction. Conflict preemption is specific to the effect to contradict federal regulation, rather than crossing a prohibited jurisdictional “bright line.”

281 See Brief of Amicus Curiae ALLCO Renewable Energy Ltd. in Support of Respondents at 20, *Hughes*, 136 S. Ct. 1288 (No. 14-614) (concluding that the facts in *Silkwood* should be looked at as convincing persuasive authority: “If Congress evidences intent to occupy a given field, any state law falling within that field is preempted.”). See *Silkwood v. Kerr-McGee Corp.*, 464 U.S. 238, 248 (1984).

282 See Brief of Amicus Curiae ALLCO Renewable Energy Ltd. in Support of Respondents at 28–29, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

283 *Id.* (citations omitted).

284 *Id.* at 32. (citing the first line of the Federal Power Act, specifically, as a point where the legislature drew the bright-line between federal and state territories). See 16 U.S.C. § 824 (2012).

Conflict preemption does not draw a “bright line,” and instead looks for a case-specific conflict case-by-case regarding individual state orders and contrary federal orders.²⁸⁵ In *Hughes*, the Circuit Court of Appeals also held that “conflict preemption applies ‘where under the circumstances of a particular case, the challenged state law stands as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress.’”²⁸⁶ The Fourth Circuit held that “[h]ere, ‘the impact of state regulation of production on matters within federal control is so extensive and disruptive of’ the PJM markets that preemption is appropriate.”²⁸⁷

1. Maryland Petitioners/Appellants

The state of Maryland argued to the Court that the Fourth Circuit’s conflict preemption rulings identify no actual conflict with a federal statute or regulation.²⁸⁸ Instead they argued that the ruling on appeal relied solely on unfounded assumptions about purposes underlying the PJM interstate capacity auction and auction rules that would supposedly be obstructed by Maryland’s support for constructing and operating CPV’s power plant.²⁸⁹ Given the FPA’s framework of explicitly divided authority over energy regulation, the Court has cautioned that the inevitable “jurisdictional tensions” between state and federal initiatives in energy regulation should not be mistaken for a conflict giving rise to preemption.²⁹⁰

Maryland prefaced the second prong of its argument by

285 See *Gade v. Nat'l Solid Waste Mgmt.*, 505 U.S. 88, 98 (1992) (citing *Florida Lime & Avocado Growers, Inc. v. Paul*, 373 U.S. 132, 142-43 (1963) and *Hines v. Davidowitz*, 312 U.S. 52, 67 (1941)) (standing for the proposition that conflict preemption occurs when a state law conflicts with a valid federal law so that it is physically impossible to comply with both or when the state law stands as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress).

286 *PPL EnergyPlus, L.L.C. v. Nazarian*, 753 F.3d 467, 478 (4th Cir. 2014), *aff'd sub nom. Hughes v. Talen Energy Marketing, L.L.C.*, 136 S. Ct. 1288 (2016) (quoting *Crosby v. Nat'l Foreign Trade Council*, 530 U.S. 363, 373 (2000)) (citations omitted).

287 *Id.* (citing *Nw. Cent. Pipeline Corp. v. State Corp. Comm'n*, 489 U.S. 493, 517-18 (1989)).

288 Brief for Petitioner *Hughes*, at 45-46, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

289 *Id.* at 46-48.

290 Brief for Petitioner CPV Maryland L.L.C. at 45, *Hughes*, 136 S. Ct. 1288 (No. 14-614) (quoting *Nw. Cent. Pipeline Corp.*, 489 U.S. at 515) (“[C]onflict pre-emption analysis must be applied sensitively in this area, so as to prevent the diminution of the role Congress reserved to the States . . .”).

addressing the conflict preemption.²⁹¹ The state noted that the Fourth Circuit decision declared field and conflict preemption as being “mutually reinforcing.”²⁹² The Maryland petitioners noted that the Circuit Court arrived at this holding “because the allegedly state-set contract rate differed from the single-year rate set by PJM’s auction.”²⁹³ The state argued that this conclusion was based on a “faulty analogy” to *Mississippi Power & Light Co. v. Mississippi ex rel. Moore and Nantahala Power & Light Co. v. Thornburg*.²⁹⁴ The Maryland petitioners first point out this alleged error in a footnote distinguishing that “[t]he [Fourth Circuit] court of appeals advanced the analogy to support its field-preemption ruling, but this Court later clarified that *Mississippi* is ‘best read as a conflict pre-emption case.’”²⁹⁵ In those earlier cases, Maryland argued that “retail regulation occurred after FERC accepted a wholesale rate, complicating FERC’s efforts to police its jurisdiction.”²⁹⁶ This footnote suggests that Maryland was advancing an argument against conflict preemption as well.

Maryland continued by arguing that state regulation contributed “to forming a wholesale rate.”²⁹⁷ “FERC’s ability to review the contracts here and its decision [outlining] how state contract-backed-resources like CPV’s participate in PJM’s auction, precluded any possibility of a conflict.”²⁹⁸ The Maryland petitioners further argued that the analogies drawn from *Mississippi* and *Nantahala* are asymmetrical to the facts-at-bar.²⁹⁹ They explained that “[those earlier cases involved a single FERC-approved wholesale rate and the question of whether states had to respect it.]”³⁰⁰ Maryland further argued that “[h]ere, there [was] no single legal rate for capacity sales in the PJM region or to the system operator,” and that was a material distinction.³⁰¹ The state argued that a “PJM tariff providing for it to buy capacity cannot prohibit FERC-jurisdictional sellers from exercising their statutory right to sell on different terms subject

291 Brief for Petitioner Hughes at 33, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

292 *Id.* (citing Pet. App. 24a).

293 *Id.* (citing Pet. App 25a-26a).

294 *Id.*

295 *Id.* at 33 note 28 (citing Oneok Inc. v. Learjet, Inc., 135 S. Ct. 1591, 1593 (2015)).

296 *Id.* at 34.

297 Brief for Petitioner at 34, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

298 *Id.*

299 *Id.*

300 *Id.*

301 *Id.*

to FERC review.”³⁰² Maryland petitioners argued that there was neither conflict preemption nor field preemption.³⁰³ Maryland took the position on appeal that the state’s procurement did not conflict with FERC authority.³⁰⁴ They argued that the Fourth Circuit failed to draw a line between internal auction policies and procedures and when these policies would preempt state actions outside of FERC.³⁰⁵ Because of this alleged lack of specificity in the Fourth Circuit decision, petitioners asserted that there was no basis for upholding the Fourth Circuit opinion on conflict preemption.³⁰⁶

2. Respondent/Challengers

Conflict preemption is a narrower case-specific state intrusion on federal authority.³⁰⁷ As it briefed this matter, the successful respondents/challengers only needed to play a form of legal defense because they had prevailed at both the trial court and Fourth Circuit levels. They did not need to concoct any new legal arguments. They also did not care whether the Supreme Court upheld conflict preemption, field preemption, or both. In the respondents/challengers’ brief, they take the broad position, ultimately encompassed by the unanimous Supreme Court opinion, that however one labels it, incursion into any element of wholesale rates by the states is preempted:

Thus, whether Maryland’s order is viewed through the lens of field preemption or of conflict preemption, the result is the same: It is a blatant incursion on FERC’s exclusive jurisdiction over the rates “received … for or in connection” with wholesale sales that produces a blatant conflict with the policies FERC has pursued in approving those wholesale rates.³⁰⁸

For example, the Maryland program resulted in Maryland guaranteeing its selected developer, CPV, 20 years of wholesale sale rate stability for this private business’s private sales of its wholesale

302 *Id.* at 33.

303 *Id.*

304 *Id.* at 45–46.

305 *Id.*

306 *Id.* at 46–48.

307 See *supra*, note 284 and accompanying text.

308 Brief for Respondents at 41, *Hughes*, 136 S. Ct. 1288 (No. 14-614) (citation omitted). See 16 U.S.C. § 824(a).

capacity and power to the interstate wholesale market operated by PJM, despite the fact that the FERC itself rejected Maryland's plea for such a rate guarantee longer than three years.³⁰⁹ The respondent appellees noted this as a deliberate state conflict with federal wholesale power markets: "Maryland's direct interference with FERC's exclusive authority to determine what a generator will receive in connection with a wholesale sale to PJM could hardly be accidental."³¹⁰

Respondents categorized this specific Maryland program as "squarely . . . intentional" and an "unapologetic" and "deliberate effort to override federal policy judgments that Maryland tried unsuccessfully to change."³¹¹ The respondents/challengers submitted that a state-directed, competitively-procured contract to support construction of a power plant with a long-term pricing structure providing incentives for that construction is different from, and conflict with, the "price signals" generated by a FERC-supervised yearly forward capacity auction.³¹²

As the matter neared decision, respondents/challengers submitted a supplemental brief to address the Supreme Court's interceding 2016 decision in *FERC v. Electronic Power Supply Association* ("EPSA")³¹³ as persuasive authority for upholding the Fourth Circuit Court's decision in *Hughes*:

The EPSA decision bolsters the conclusion reached by every federal judge to consider the matter that Maryland's attempt to dictate what a wholesale seller will receive for, or in connection with its wholesale sales of electricity in a wholesale-market operator's auction is preempted.³¹⁴

Moreover, the respondents/challengers noted that

309 Brief for Respondents at 35, 41, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

310 *Id.*

311 *Id.* at 35, 38.

312 *Id.* at 37-39. Under the Maryland scheme, challengers assert that the successful bidder sets the ultimate rate that it will receive, while in a competitive PJM interstate (federal) procurement, the resulting rate is set in a competitive market. Brief for Petitioner CPV Maryland L.L.C. at 38, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

313 F.E.R.C. v. Elec. Power Supply Ass'n, 136 S. Ct. 760 (2016)

314 Supplemental Brief for Respondents at 1, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

"Maryland's effort to ensure that CPV received different [payment] amounts in connection with its sales to PJM is just as focused on the wholesale market, which is a fatal problem under both EPSA and broader principles of preemption."³¹⁵ The respondents/challengers also argued that the EPSA decision confirms that Maryland cannot get around preemption by forcing, through its regulation, its regulated third party utilities to be the ones to make side payments to CPV in connection with its wholesale sales to PJM (instead of trying to mandate what PJM itself must pay CPV for those sales).³¹⁶ "The FPA 'leaves no room either for direct state regulation of the prices of interstate wholesales' or for regulation that 'would indirectly achieve the same result.'"³¹⁷

3. Amici, Allco

The Allco amici devotes ten pages of its brief³¹⁸ arguing that Maryland's scheme is conflict preempted because it conflicts with implementation of FERC's market-based regulatory scheme.³¹⁹ Specifically, Allco asserts that, otherwise, one enters a world "in which state commissions can compel entry into a wholesale electricity contract, and do so at a price that is neither the FERC-regulated market price resulting from PJM, nor a price that is either freely negotiated between seller and purchaser, or the price permitted by [the Public Utility Regulatory Policies Act]."³²⁰ Finally, Allco argued that Maryland's scheme would likely lead to states "supplanting [the FERC's] chosen regulatory approach."³²¹ They contend that this has been clearly prohibited by the FPA.³²²

C. The Supreme Court Response to the Parties' Framing

Challenger/appellee Talen drew on precedent supporting both

315 *Id.* at 3.

316 *Id.*

317 *Id.* at 3–4 (quoting *N. Natural Gas Co. v. State Corp. Comm'n of Kan.*, 372 U. S. 84, 91 (1963)). *See also* 16 U.S.C. § 824d(a) (noting that the statute on its face also preempts this) What matters under the FPA is whether a state is trying to dictate the price that a wholesale seller will "receive[] ... for or in connection with" interstate wholesale sales. *Id.*

318 Brief of Amicus Curiae ALLCO Renewable Energy Ltd. in Support of Respondents at 26–36, *Hughes v. Talen Energy Mktg., L.L.C.*, 136 S. Ct. 1288 (2016) (No. 14-614).

319 *Id.* at 28.

320 *Id.*

321 *Id.* at 32.

322 *Id.* at 32.

field and conflict preemption. The challengers were supported by the trial court and Fourth Circuit decisions supporting field preemption, as well as conflict preemption added in the decision of the Fourth Circuit. It was not critical which type of preemption the courts upheld on appeal regarding state motivations to exercise control over the location, pricing revenue, and operations of wholesale generators: “There are, however, grave risks of state intrusion into an exclusively federal field and state frustration of FERC’s policies with respect to what a generator should receive for or in connection with its sales to PJM.”³²³ Talen broadly positioned Maryland’s actions steering a particularly chosen “generator [to] bid its energy and capacity into PJM” although the Maryland action would have it “receive a price for that energy and capacity different from that provided by the federally approved” PJM protocol, without labeling it as a particular type of preemption, as “plainly preempted.”³²⁴

While this position allowed the Court to render a generically broad preemption ruling, similar to the Court saying, in the words of Justice Stewart, that they “know it when [they] see it,”³²⁵ it did not require them to do more than affirm the Fourth Circuit holding on appeal. The Justices split 7-1-1 as to what type of preemption was present, but all agreed that states may not directly or indirectly implement energy regulation which affects or determines the outcome of wholesale power market rates:

One Justice, in his concurring opinion, found express preemption;³²⁶

one, in her concurring opinion, found only conflict preemption;³²⁷ and

323 Supplemental Brief for Respondents at 5-6, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

324 *Id.* at 6 (emphasis added). Respondents also added an argument that the states had no authority to regulate wholesale transactions under the guise of local power when the Federal Power Act was passed. Brief for Respondent at 25-27, *Hughes*, 136 S. Ct. 1288 (No. 14-614).

325 *Jacobellis v. Ohio*, 378 U.S. 184, 197 (1963) (Stewart, J., concurring). The Fourth Circuit reinforced Supreme Court “bright line” energy preemption precedent in *Mississippi Power & Light Co. v. Mississippi ex rel. Moore*, 487 U.S. 354, 371 (1988).

326 See *Hughes*, 136 S. Ct. 1288, 1301 (2016) (Thomas, J., concurring).

327 See *id.* at 1300 (Sotomayor, J., concurring at 3).

the majority opinion of 7 Justices discussed both field and conflict preemption, upholding the decision of the Fourth Circuit which found both.³²⁸

It is important to note that the cases relied on by the Fourth Circuit are the same cases that the Supreme Court relied on for upholding the Circuit's field preemption decision.³²⁹ The Fourth Circuit reinforced the Supreme Court "bright line" energy preemption precedent as stated in *Mississippi Power & Light Co. v. Mississippi ex rel. Moore*.³³⁰ Given how the petitioners/appellees presented the matter to the Court, when the Supreme Court upheld the Fourth Circuit opinion, relying on without qualifying the applicability of the *Mississippi Power* or *Nantahala* decisions (the two classic "bright line" energy precedents), it found no merit to appellant Maryland's position that there was no field preemption and this case was somehow distinct in time. In refusing to distinguish its prior "bright line" jurisdictional precedent, the Court also reinforced the long-established "bright line" in these two key Supreme Court precedents separating federal wholesale and state retail jurisdiction.³³¹ Field preemption is the "bright line" which prevents both state and federal energy regulators from intruding across the line into the jurisdiction exclusively reserved for the other level of government.

V. Constitutional Implications of Field Preemption for the Most Regulated Industry

The *Hughes* decision critically re-etches and reinforces a unique "bright line" separating government authority over power in U.S. law. The positions articulated by the primary opinion and two concurring Supreme Court opinions frame the enduring legacy and impact of this opinion going forward. The *Hughes* opinion operates as a jurisprudential inflection point between field preemption and conflict preemption pursuant to the Supremacy Clause of the Constitution. The state of Maryland lost every legal issue it presented before every one of the thirteen federal judges that heard the case at

328 See *supra* Section III A.

329 Compare *Hughes*, 136 S. Ct. 1288 with *Mississippi Power*, 487 U.S. 354 and *Nantahala Power and Light Co. v. Thornburg*, 476 U.S. 953 (2010).

330 PPL EnergyPlus, L.L.C. v. Nazarian, 753 F.3d 467, 475–77 (4th Cir. 2014), aff'd sub nom. *Hughes v. Talen Energy Marketing, L.L.C.*, 136 S. Ct. 1288 (2016) (citing *Mississippi Power*, 487 U.S. 354 (1988)).

331 See *Hughes*, 136 S. Ct. 1288.

all three levels of the federal court system. The Supreme Court chose to unanimously ratify the Fourth Circuit opinion which reset the “bright line” separating exclusive state and federal authority over power:

[T]he Supreme Court has expressly rejected the proposition that the “scope of [FERC’s] jurisdiction ... is to be determined by a case-by-case analysis of the impact of state regulation upon the national interest.” Instead, “Congress meant to draw a bright line easily ascertained, between state and federal jurisdiction... This was done in the [FPA] by making [FERC] jurisdiction plenary and extending it to all wholesale sales in interstate commerce except those which Congress has made explicitly subject to regulation by the States.” The federal scheme thus “leaves no room either for direct state regulation of the prices of interstate wholesales of [energy], or for state regulations which would indirectly achieve the same result.” “Even where state regulation operates within its own field, it may not intrude indirectly on areas of exclusive federal authority.” As a result, states are barred from relying on mere formal distinctions in “an attempt” to evade preemption and “regulate matters within FERC’s exclusive jurisdiction.”³³²

The Supreme Court did not grant any deference to the state of Maryland or state agencies. The Court granted expanded deference only to FERC at the federal level. The implications of this important decision, although concerning only the state of Maryland as defendant, extend to 47 of the 50 states in two important regards. First, the Supreme Court reinforced the “bright line” excluding all state regulation affecting directly or indirectly transactions in interstate wholesale power. This decision disables state authority.

This Supreme Court decision in *Hughes* is important because the Justices had a choice to find and apply either field preemption or conflict preemption. Field preemption erects an impenetrable jurisdictional wall that bars states in perpetuity from direct or indirect incursion in such energy matters, while conflict preemption bars only

332 *Nazarian*, 753 F.3d at 475–476 (citations omitted).

the specifically challenged regulation. This distinction is important. When carefully analyzed, neither the Supreme Court, nor either of the lower federal courts in this matter, chose the much narrower single-regulation conflict preemption to arrest only a particular Maryland regulation. The Supreme Court unanimously upheld and reset a “bright line” field-preempted barrier. *Hughes* preempts the entire field of regulatory jurisdiction of 47 of the 50 contiguous states that are interconnected by copper wire transmission lines.

Second, this Supreme Court decision exerts important policy implications for other states that participate in the nation’s interstate independent system operator wholesale power markets. Maryland chose to participate in the PJM interstate wholesale power market system, a large ISO in the United States.³³³ In recent times, the role of federal energy regulation “has become increasingly prominent.”³³⁴ The Circuit Court stressed that Maryland chose to abandon its prior state model “and throw in its lot with the federal interstate markets . . . effectively compelling” participation in and adherence to the exclusively federally regulated interstate wholesale market.³³⁵

States and their retail utilities have options in terms of the power markets in which they participate. Instead, if Maryland had elected a different regulated electricity structure, it could have imposed requirements as to its power generators’ location and operation of in-state power generation plants. However, the state cannot commit half-way to participation in an interstate wholesale market: Here, the state chose to participate in the PJM ISO interstate wholesale electricity market. Despite this wholesale market, the state provided incentives for in-state generation participating in that PJM wholesale market. When the final cost received for wholesale transactions by private in-state electric generators are “topped off” by the state, such wholesale markets are exclusively within FERC interstate and wholesale power jurisdiction. Therefore, this state regulation, which interferes even indirectly with FERC’s federally-approved PJM market tariffs, crosses the FPA “bright line” separating state and federal regulatory jurisdiction, which demarcates the preempted wholesale field, and was outlawed.

This decision will cause some states, lamenting their loss of control over a variety of power regulatory issues, to reexamine electric deregulation, re-regulation, and power supply options in

333 See *Hughes*, 136 S. Ct. at 1293–95.

334 *Nazarian*, 753 F.3d at 472 (citing *New York v. FERC*, 535 U.S. 1, 7 (2002)).

335 *Id.* at 473.

order to regain some energy authority post-*Hughes*. There are new challenges as the technologies and power markets evolve in the 21st century.³³⁶ Notwithstanding that the technologies and power markets are changing, the law's "bright line" separating the lines of exclusive state and federal authority over power was reinforced by the Supreme Court in 2016.

The decision interprets the Supremacy Clause and specifically declares supremacy of federal law as applied to the electricity market. The *Hughes* opinion construes the Supremacy Clause of the Constitution separating power in the American economy.³³⁷ As stated by another federal court in 2016:

The discovery of fire was a significant event, creating for mankind warmth against the cold and light in the darkness. We do not know which man or woman first noticed that a burning bundle of sticks produced those useful results of warmth and light, which in modern times are the products of alternative forms of energy. Electrical energy is one of these. The concept of electricity was first deduced by William Gilbert, a physician in the service of Elizabeth I of England (1533–1603). In 1752, Benjamin Franklin demonstrated the practical application of electricity by flying a kite carrying a key into a lightening storm. Today, electricity is a principal source of light and heat for the world and its people. As the importance of electricity has increased exponentially in human affairs, politicians and governments inevitably stepped up regulation of the generation and marketing of electrical energy.³³⁸

Under *Hughes*, the law applies differently to the electricity market than to any other thing in the U.S. economy. *Hughes* reset one of the most important constitutional federalist lines in American law. With an approximate order of magnitude increase in the percentage

³³⁶ See generally Steven Ferrey, *California Challenges and Vulnerabilities of the New Business Model Design for Power*, 6 SAN DIEGO J. OF CLIMATE & ENERGY L. 1 (2014-2015).

³³⁷ See *Hughes*, 136 S. Ct. 1288.

³³⁸ Allco Finance, Ltd. v. Klee, Nos. 3:15-cv-608 and 3:16-cv-508 at *2 (D. Conn Aug. 18, 2016).

of wholesale transactions, there is a fundamental shift inserting the federal government, rather than the states, into exclusive control of these transactions and their regulation.³³⁹

There are financial implications for states continuing to leap before they look, in terms of energy regulation in the 21st century. The U.S. Code allows challengers to collect attorney fees when they successfully challenge constitutionally preempted state regulation of energy markets.³⁴⁰ The plaintiffs in a related case, *Entergy Nuclear Vt. Yankee, L.L.C. v. Shumlin*,³⁴¹ successfully challenged state energy regulation on constitutional claims. After they prevailed, they made a claim for such fees, and were awarded those fees by the federal trial court in an amount just through the trial stage in excess of \$4 million.³⁴²

Legal separation of state and federal legal authority is at the core of the U.S. Constitution, incorporated through field preemption and conflict preemption pursuant to the Supremacy Clause of the Constitution. Electricity is the foundation of the modern economy.³⁴³ In *Hughes v. Talen Energy Marketing*, the Supreme Court re-etched in high relief one of the most prominent and distinctive Supremacy Clause "bright lines" in American law.³⁴⁴ Careful analysis of the

339 In 2014, nearly forty percent of U.S. electricity was generated by what the U.S. Information Administration terms "independent power producers," increased almost 400% from ten percent two decades earlier. See U.S. ENERGY INFO. ADMIN., *Electric Power Monthly*, at Tbls. 1.2–1.5, <http://www.eia.gov/electricity/monthly/archive/august2015.pdf>; Fed. Energy Reg. Comm'n Rep., Order No. 888, 61 Fed. Reg. 21, 540, 21, 541, 21, 549–50 (1996).

340 42 U.S.C. § 1983 ("Every person who...subjects, or causes to be subjected, any citizen of the United States or other person...to the deprivation of any rights, privileges, or immunities secured by the Constitution and laws, shall be liable to the party injured in an action at law, suit in equity, or other proper proceeding for redress...unless a declaratory decree was violated or declaratory relief was unavailable."). In any action enforcing Section 1983, 42 U.S.C. §§ 1987–88 permits a court to "allow the prevailing party, other than the United States, reasonable attorney's fee as part of the costs." Section 1988 also allows for expert fees to be included as part of the attorney's fee.

341 *Entergy Nuclear Vermont Yankee L.L.C. v. Shumlin*, 838 F. Supp. 2d 183 (D. Vt. 2012), *aff'd in part, rev'd in part*, 733 F.3d 393 (2d Cir. 2013).

342 Vermont appealed that case on the merits and lost its appeal in the Second Circuit; the generator then sought to recover the over \$4 million incurred in legal fees at the appellate stage. Sandy Levine, *The Dicey Economics of Hosting a Nuclear Plant*, CONSERVATION LAW FOUNDATION (Jan. 16, 2013) <https://www.clf.org/blog/the-dicey-economics-of-hosting-a-nuclear-plant/>.

343 See FERREY, *supra* note 7.

344 See *Hughes v. Talen Energy Mktg.*, L.L.C., 136 S. Ct. 1288 (2016).

opinion and the position of the parties on brief, underscores that this important unanimous Supreme Court opinion reinforces the “bright line” excluding state regulation of electricity impacting the field of the increasingly dominant wholesale power market. This is reinforced field preemption.

* * *

**The Cost of Justice: The Importance of a Criminal
Defendant's Ability to Pay
in the Era of *Commonwealth v. Henry***

*Cristina Rodrigues**

* Cristina Rodrigues works as a public defender for the Commonwealth of Massachusetts. In that job, she represents poor individuals who are accused of crimes in Boston's Municipal Courts. She was born and raised in Massachusetts. She received her law degree from Harvard Law School in 2013 and her undergraduate degree from Brown University in 2010.

Table of Contents

I.	Origins: The Policies Underpinning Restitution.....	18
A.	<i>Policy Justifications for Restitution</i>	18
B.	<i>Criticisms of Restitution</i>	23
II.	Structure: The Legal Framework Around Restitution	27
A.	<i>The Legal Basis for Court Authority to Order Restitution .</i>	27
B.	<i>The Broad Scope of Losses Recoverable Through Restitution.....</i>	30
C.	<i>Procedural Requirements for a Restitution Order.....</i>	34
D.	<i>Lower Court Practices Regarding Ability to Pay Prior to Henry.....</i>	36
III.	A Close Examination: Commonwealth v. Henry	43
A.	<i>The Facts of Henry</i>	43
B.	<i>The Reasoning of Henry</i>	45
IV.	The Impact of Henry	60
A.	<i>Henry's Impact on the Policy Underpinnings of Restitution</i>	61
B.	<i>Henry's Impact on Other Payments Imposed by Criminal Courts.....</i>	65
1.	<i>A Survey of Other Court-Imposed Payments</i>	65
2.	<i>Practical Consequences of the Imposition and Collection of Payments</i>	72
a.	<i>The Undeniable Reality of Debt-Based Jail Time in Massachusetts</i>	73
b.	<i>The Financial Burden on the State's Marginalized Communities.....</i>	76
3.	<i>The Potential Application of Henry Beyond Restitution</i>	82
V.	Conclusion.....	86

In recent years, growing attention has focused on just how expensive even minimal contact with the criminal legal system is for the accused. Regardless of guilt, the mere issuance of criminal charges pulls an individual into a web of court-imposed costs and fees.¹ In Massachusetts, criminal defendants face the prospect of having to pay bail, victim witness fees, legal counsel fees, default removal fees, arrest warrant fees, probation supervision fees, and restitution.² Defendants can be punished with jail time for failure to pay these fees.³ The collection of fees under the threat of incarceration is particularly troubling when considered alongside the reality that most defendants in Massachusetts' state criminal courts are poor.⁴ Approximately 75 percent of criminal defendants in Massachusetts qualify as indigent.⁵ It is estimated that more than

1 MASS. TRIAL CT. DIST. CT. DEP'T, POTENTIAL MONEY ASSESSMENTS IN CRIMINAL CASES (2015), <http://www.mass.gov/courts/docs/courts-and-judges/courts/district-court/potential-moneyassessment-criminalcases.pdf>. For a more general discussion, see NAT'L CONSUMER LAW CTR., CONFRONTING CRIMINAL JUSTICE DEBT: A GUIDE FOR LITIGATION 18 (2016), <http://cjpp.law.harvard.edu/assets/CJD-LitigationGuide.pdf> (noting that “in today’s criminal justice system, a broad array of fees – beyond merely indigent defense recoupment – purport to refund the courts and the justice system for the costs of administering the criminal process”).

2 See MASS. TRIAL CT. DIST. CT. DEP'T, *supra* note 1; MASS. GEN. LAWS ch. 276, § 57 (2014); MASS. GEN. LAWS ch. 276, § 31 (2000).

3 See, e.g., MASS. GEN. LAWS ch. 127, § 144 (2017) (providing for a defendant to receive a \$30 credit towards his court debt for every day he spends in jail for nonpayment of “fine[s] and expenses” and allowing for release of the defendant only once the debt is paid). Indeed, the Massachusetts Senate Post Audit and Oversight Committee released a report in November 2016 entitled “Fine Time Massachusetts.” MASS. S. POST AUDIT AND OVERSIGHT COMM., JUDGES, POOR PEOPLE, AND DEBTORS’ PRISON IN THE 21ST CENTURY 25 (2016), <http://senatormikebarrett.com/wp-content/uploads/2016/11/Fine-Time-the-Report-Nov.-7-2016..pdf>. The report concluded, *inter alia*, that “[f]ine time is alive and well in Massachusetts district courts, a function of unawareness at the top, unchallenged and sometimes arbitrary power at the bottom, and pressure for revenue throughout.” *Id.*

4 Ralph D. Gants, *What’s New In the State Court System*, COMMONWEALTH MAG. (Oct. 23, 2015), <http://commonwealthmagazine.org/criminal-justice/whats-new-in-the-state-court-system/>.

5 *Id.* Note that the Supreme Judicial Court defines indigency as “receiving an annual income, after taxes, of one hundred twenty-five percent or less of the current poverty guidelines,” receiving one of several enumerated public benefits, residing in one of several enumerated treatment facilities, or being a juvenile. See MASS. SUP. JUD. CT. R. 3:10 (h). National estimates place the rate at an even higher 80% of state criminal defendants qualifying for a

40 percent of the individuals who appear in Boston's criminal courts for minor offenses are either homeless or without stable homes.⁶ In the pointed words of Ralph Gants, the Chief Justice of the Supreme Judicial Court, amongst the individuals paying criminal court fees “[m]ost are dead broke, or nearly broke.”⁷

Even where poor individuals manage to find the money to make these payments and avoid debt-based jail time, the consequences of the payments are still severe. In 2016, Massachusetts State Senator Michael Barrett (D-Lexington) expressed concern over his “susp[icion]” that defendants desperate to avoid jail time for their failure to pay “lean very heavily on their equally poor families and somehow come up with the money.”⁸ Barrett has described the current practice of court fee collection as amounting to “a hugely regressive tax in instances where it’s paid.”⁹ According to a report

government-provided lawyer. See John Pfaff, Opinion, *A Mockery of Justice for the Poor*, N.Y TIMES (Apr. 30, 2016), <https://www.nytimes.com/2016/04/30/opinion/a-mockery-of-justice-for-the-poor.html>.

- 6 James H. Burnett III, *Boston's Homeless Court helps defendants get a fresh start*, Bos. GLOBE (Aug. 12, 2012), <https://www.bostonglobe.com/lifestyle/2012/08/11/finding-order-court-different-person-now-brian-campbell-said-used-comfortable-for-jail-now-think-what-kind-life-was-that/Ge8LqeMUVvR1X4hem4EJoN/story.html>.
- 7 Gants, *supra* note 4.
- 8 Shira Schoenberg, *Report Sheds Light on Poor Defendants Serving ‘Fine Time’ in Massachusetts*, MASSLIVE.COM (Nov. 7, 2016), http://www.masslive.com/politics/index.ssf/2016/11/report_sheds_light_on_poor_inm.html. For a scholarly analysis of the communal cost of punishment, see generally Darryl K. Brown, *Street Crime, Corporate Crime, and the Contingency of Criminal Liability*, 149 U. PA. L. REV. 1295, 1305-1308 (2001) (“Criminal sanctions impose costs on others beyond the defendant, and they impose costs on defendants beyond the formal components of punishment.”).
- 9 See Schoenberg, *supra* note 8. Barrett’s statements are supported by a 2016 report. Wendy Sawyer, *Punishing Poverty: The High Cost of Probation Fees in Massachusetts*, PRISON POL’Y INITIATIVE (Dec. 8, 2016), https://www.prisonpolicy.org/probation/ma_report.html (“Probation fees function as a punishing regressive tax, wherein the state raises revenue by charging the poorest communities the most.”). See also Michael Jonas, *Probation Fees Hit Poor the Hardest, Says Report*, COMMONWEALTH MAG. (Dec. 7, 2016), <https://commonwealthmagazine.org/criminal-justice/probation-fees-hit-poor-the-hardest-says-report/> (“Monthly fees charged to those on probation in Massachusetts disproportionately hit poorer communities, a costly obstacle to rehabilitation for those least able to afford it, according to a new study.”). Daniel Medwed, a professor at Northeastern University School of Law, appeared to agree with the sentiment, writing that “[i]n effect, the poor are subsidizing our government budgetary shortfall through a regressive tax cloaked in the guise of probation fees.” Daniel Medwed, *Poverty, Punishment,*

prepared by multiple state court judges and administrators, the trial courts of Massachusetts collected \$99,905,176.05 in fines, fees and court costs in fiscal year 2016.¹⁰ Much of that amount comes from poor residents, their poor families, and in the aggregate poor neighborhoods, as those groups are overrepresented in the criminal courts.¹¹

In Massachusetts, public interest in criminal justice reform appears to be growing.¹² The ongoing public conversation regarding the role of fees in the criminal legal system is not, however, limited to Massachusetts.¹³ The national dialogue was sparked perhaps

-
- and Probation: A Toxic Brew*, WGBH News (Dec. 8, 2016), <http://news.wgbh.org/2016/12/08/local-news/poverty-punishment-and-probation-toxic-brew>.
- 10 MASS. TRIAL CT. FINES AND FEES WORKING GROUP, REPORT TO TRIAL CT. CHIEF JUSTICE PAULA M. CAREY 5 (2016), <http://www.mass.gov/courts/docs/trial-court/report-of-the-fines-and-fees-working-group.pdf>.
- 11 See, e.g., Sawyer, *supra* note 9. For data regarding the general overrepresentation of Black and Latino residents in the Massachusetts criminal system, see generally RHIANA KOHL ET. AL., A TEN-YEAR TREND ANALYSIS OF RACE/ETHNICITY (2012), <http://www.mass.gov/eopss/docs/doc/research-reports/pop-trends/a-ten-year-trend-analysis-of-raceethnicity-2002-2011.pdf>; MASS. S. POST AUDIT AND OVERSIGHT COMM., FINE TIME MASSACHUSETTS: JUDGES, POOR PEOPLE AND DEBTORS' PRISON IN THE 21ST CENTURY (2016).
- 12 For instance, Governor Baker is currently considering a bill that would grant judges more flexibility in waiving court costs. See Jan Ransom, *Bill offered by Baker would put an end to 'fine time'*, Bos. GLOBE (Apr. 12, 2017), <https://www.bostonglobe.com/metro/2017/04/11/baker-bill-gives-poor-defendants-community-service-option-pay-fines/fS4MjzREEdj36q6I6bbDGI/story.html>; Medwed, *supra* note 9; Editorial Board, *Our view: Giving Criminals a Second Chance*, TAUNTON DAILY GAZETTE (Apr. 25, 2017), <http://www.tauntongazette.com/opinion/20170425/our-view-giving-criminals-second-chance>; Editorial, *No Place in Mass. For Prisoners of Poverty*, Bos. GLOBE (Nov. 25, 2016), <http://www.bostonglobe.com/opinion/editorials/2016/11/25/place-mass-for-prisoners-poverty/NCxcteur4Wc2JR9iLoYFLN/story.html>. Over sixty civil rights groups in Massachusetts have called for criminal justice reform. *Nearly 70 Massachusetts civil rights organizations issue call for criminal justice reform*, AM. CIVIL LIBERTIES UNION MASS. (Jan. 17, 2017), <https://aclu.org/uncategorized/nearly-70-massachusetts-civil-rights-organizations-issue-call-criminal-justice-reform/>.
- 13 See, e.g., Garrett Epps, *Can States Make People Pay Even When Their Convictions are Overturned?*, THE ATLANTIC (Jan. 6, 2017), <https://www.theatlantic.com/politics/archive/2017/01/can-states-make-people-pay-even-when-their-convictions-are-overturned/512360/> (discussing a Supreme Court case challenging Colorado's decision to deny repayment of restitution and fine payments to two individuals whose convictions had been overturned and arguing that "[m]ost states have increasingly turned to forcing criminal defendants to pay the costs of their own prosecution.").

most prominently by the Department of Justice's 2015 investigation into police and court practices in Ferguson, Missouri.¹⁴ The report concluded that law enforcement efforts in Ferguson were improperly focused on generating revenue rather than promoting public safety and that the practices "impose[d] a particular hardship" on Ferguson's residents who live at or near the poverty line.¹⁵ The Department of Justice released resources to assist other local municipalities in reforming fee and fine practices.¹⁶ In announcing those resources, then-Attorney General Loretta E. Lynch stated that "[t]he consequences of the criminalization of poverty are not only harmful – they are far-reaching."¹⁷ Since then, advocacy groups across the country have initiated lawsuits challenging local municipalities for improper fee collection from poor defendants.¹⁸

14 See, e.g., Sophie Quinton, *After Ferguson, States Struggle to Crack Down on Court Debt*, THE PEW CHARITABLE TR. (Aug. 26, 2015), <http://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2015/08/26/after-ferguson-states-struggle-to-crack-down-on-court-debt>; Andrea Marsh & Emily Gerrick, *Why Motive Matters: Designing Effective Policy Responses to Modern Debtors' Prison*, 34 YALE L. & POL'Y REV. 93, 98 (2015), http://ylpr.yale.edu/sites/default/files/YLPR/marshgerrick_finalprod.pdf.

15 See U.S. DEPT' OF JUST. CIVIL RIGHTS DIV., INVESTIGATION OF THE FERGUSON POLICE DEPARTMENT 4, 9-15, 52-54 (2015), https://www.justice.gov/sites/default/files/opa/press-releases/attachments/2015/03/04/ferguson_police_department_report.pdf.

16 *Justice Department Announces Resources to Assist State and Local Reform of Fine and Fee Practices*, U.S. DEP'T OF JUST. (Mar. 14, 2016), <https://www.justice.gov/opa/pr/justice-department-announces-resources-assist-state-and-local-reform-fine-and-fee-practices>.

17 *Id.*

18 Litigation on this issue of court costs imposed on poor defendants has been seen in *inter alia* Alabama, Colorado, Georgia, Louisiana, Mississippi, Missouri, South Carolina, Tennessee, Texas and Washington. See e.g., *Class Action Lawsuit Accuses 13 Missouri Cities of Running 'Debtors Prisons'*, NAT'L PUB. RADIO (Aug. 12, 2016, 4:35 PM), <http://www.npr.org/2016/08/12/489816672/class-action-lawsuit-accuses-13-missouri-cities-of-running-debtors-prisons>; Tom Kutsch, *St. Louis-area Municipalities Sued Over Alleged Discriminatory 'Debtors' Prisons'*, THE GUARDIAN (Aug. 10, 2016, 9:29 AM), <https://www.theguardian.com/us-news/2016/aug/10/st-louis-civil-rights-lawsuit-debtors-prisons-african-americans>; Meg Kinnard, *ACLU: South Carolina County is Operating Debtors Prison*, U.S. NEWS (June 1, 2017, 12:06 PM), <https://www.usnews.com/news/best-states/south-carolina/articles/2017-06-01/aclu-south-carolina-county-of-operating-debtors-prison>; Catherine Thorbecke, *Colorado Springs Will Re-Pay Those Sent to Debtor's Prison in Landmark Settlement*, ABC NEWS (May 5, 2016, 6:33 PM), <http://abcnews.go.com/US/colorado-springs-pay-debtors-prison-landmark-settlement/story?id=38905750>; *Probation Firm Extorted Money from Poor in Alabama, Suit Charges*, NBC NEWS (Mar. 13,

In August 2016, in the midst of this heightened local and national dialogue regarding the criminal system's collection of payments from poor individuals, the Supreme Judicial Court issued its bold, unanimous decision in *Commonwealth v. Henry*.¹⁹ In *Henry*, the Supreme Judicial Court made critical advances in the legal doctrine regarding restitution—that is, the payments that criminal defendants are ordered to pay to victims as a component of their punishment. The court ruled that lower court judges must consider a defendant's ability to pay when first determining the restitution amount.²⁰ The court barred judges from imposing restitution that causes the defendant a substantial financial hardship.²¹ The court explained that its decision was compelled by a long-standing and fundamental prohibition on the practice of punishing defendants more harshly based on their impoverishment.²² In the court's view, conditioning a defendant's probation on payments that he has no ability to make violates that prohibition.²³ Accordingly, the court also barred lower court judges from imposing a longer probation period on a defendant or extending a defendant's probation for the sole purpose of collecting more payments.²⁴

This Article critically examines the *Henry* decision by situating the case within existing academic discussions and case law on restitution and by exploring the case's potentially massive impact on the collection of court-imposed payments from poor defendants. Part I reviews how restitution fits into the criminal legal system, laying out the policy justifications for restitution and the most prominent criticisms lobbied in response. Part II then turns to the legal framework for restitution in Massachusetts. I set out the origins of criminal court authority to order restitution and the

2015, 1:30 PM), <http://www.nbcnews.com/news/us-news/probation-firm-extorted-money-poor-alabama-suit-charges-n323036>; Rebecca Santana, *Lawsuit Charges that Louisiana Judge Runs a Debtors' Prison*, SAN DIEGO UNION-TRIBUNE (June 21, 2016, 9:10 AM), <http://www.sandiegouniontribune.com/sdut-lawsuit-alleges-louisiana-judge-runs-debtors-2016jun21-story.html>. The issue of collection of court fees from poor defendants has also garnered academic attention. See, e.g., Note, *State Bans on Debtors' Prisons and Criminal Justice Debt*, 129 HARV. L. REV. 1024 (2016), <http://harvardlawreview.org/wp-content/uploads/2016/02/1024-1045-Online2-2.pdf>.

19 *Commonwealth v. Henry*, 55 N.E.3d 943 (Mass. 2016).

20 *Id.* at 950.

21 *Id.* at 954.

22 *Id.* at 952, 954.

23 *Id.* at 950.

24 *Id.* at 952.

substantive and procedural limits on restitution that existed before *Henry*. Part III examines *Henry* itself, closely inspecting the legal reasoning the Supreme Judicial Court employed and highlighting the most extraordinary parts of the decision. Finally, Part IV considers the impact of *Henry*. I argue that the rules set out in *Henry* render restitution orders both more constitutionally sound and practically effective. I also consider the many other payments imposed on poor defendants in Massachusetts, the disproportionate burden of those payments on the state's marginalized communities, and the multiple reasons *Henry*'s principles should be applied to all monetary sanctions.

I. Origins: The Policies Underpinning Restitution

A. Policy Justifications for Restitution

Restitution is the sum a person guilty of a criminal offense pays to the victim of the criminal offense to compensate for the financial losses experienced as a result of the criminal conduct.²⁵ Importantly, those payments are collected “under the supervision of the criminal justice system.”²⁶ Restitution seeks “to bring the victim as close as possible to the position that the victim was in before the crime occurred.”²⁷ The basic principle of compensation that underlies restitution is well settled in Massachusetts law.²⁸ Victim compensation has been recognized by Massachusetts courts since as early as 1834.²⁹ However, substantial questions have lingered regarding the extent to which that compensation should be achieved through restitution payments ordered as punishment in criminal cases.

Understanding the debate regarding restitution requires considering how restitution fits into the structure and goals of

25 Commonwealth v. Rotonda, 747 N.E.2d 1199, 1207 (Mass. 2001) (explaining that “[t]he purpose of restitution . . . is to compensate the injured party for losses incurred as a result of the defendant's criminal conduct”); Commonwealth v. Casanova, 843 N.E.2d 699, 704 (Mass. App. Ct. 2006) (stating that “[r]estitution . . . serves the ancillary purpose of compensating the victim for economic losses”).

26 Thomas M. Kelly, *Where Offenders Pay for Their Crimes: Victim Restitution and Its Constitutionality*, 59 NOTRE DAME L. REV. 685, 688 (1984).

27 Kelse Moen, Note, *Choice in Criminal Law: Victims, Defendants, and the Option of Restitution*, 22 CORNELL J. L. & PUB. POL'Y 733, 740 (2013).

28 See, e.g., Swift v. Barnes, 33 Mass. (1 Pick.) 194, 197 (1834).

29 *Id.*

the criminal legal system as a whole. The American criminal legal system in its current form views defendants as having committed offenses against society as a whole and, consequently, owing a debt to society.³⁰ In this view, whenever a crime is committed, members of the public constitute indirect victims and suffer “decreased personal and material security and diminished incentives toward socially acceptable ways of acquiring wealth.”³¹ The theories justifying our current criminal legal system are primarily retribution—punishment inflicted based on the idea that the person is morally culpable and deserving³²—and utilitarianism—punishment designed to deter future crime and rehabilitate defendants.³³ Taking these theories together, criminal law aims to 1) impose punishment on criminal wrongdoers based on culpability; 2) impose punishment in a manner that promotes the rehabilitation of wrongdoers; and 3) impose punishment in an effort to deter future crime.³⁴

To promote a uniform commitment to those goals, the state maintains overwhelming control of prosecutorial power.³⁵ Thus,

30 See, e.g., Albert Egash, *Creative Restitution -- A Broader Meaning for an Old Term*, 48 J. CRIM. L. & CRIMINOLOGY 619 (1958); Richard Adelstein, *Victims as Cost Bearers*, 3 BUFF. CRIM. L. REV. 131, 137 (1999); Joshua Kleinfeld, *Reconstructivism: The Place of Criminal Law in Ethical Life*, 129 HARV. L. REV. 1485, 1489 (2016) (“[C]riminal law has a distinctive role to play in the social world . . . because criminal law is the primary legal institution by which a community reconstructs the moral basis of its social order, its ethical life, in the wake of an attack on that ethical life.”); James William Casson III, *Restitution: An Economically and Socially Desirable Approach to Sentencing*, 9 NEW ENG. J. OF CRIM. & CIV. CONFINEMENT 349, 353 (1983), <http://offerofproof.net/wp-content/uploads/9.2.Casson.pdf> (discussing “the concept of harm to society” and describing it as “the state’s philosophical justification for punishing the criminal”).

31 Adelstein, *supra* note 30, at 156.

32 See, e.g., Mary Ellen Gale, *Retribution, Punishment, and Death*, 18 U.C. DAVIS L. REV. 973, 1003-04 (1985) (“[R]etribution begins with the recognition that we punish criminals at least partly because we believe that they ‘deserve’ punishment.”); Mike C. Materni, *Criminal Punishment and the Pursuit of Justice*, 2 BRIT. J. AM. LEGAL STUD. 263, 266 (2013).

33 See, e.g., JEREMY BENTHAM, THE RATIONALE OF PUNISHMENT, bk. 1, ch. 3 (Robert Heward, Wellington Street, the Strand, London 1830), <http://www.laits.utexas.edu/poltheory/bentham/rp/rp.b01.c03.html> (contending that the value and legitimacy of punishment is measured by whether it serves some greater good); Jean Hampton, *The Moral Education Theory of Punishment*, 13 PHIL. & PUB. AFF. 208, 212 (1984).

34 See Note, *Victim Restitution in the Criminal Process: A Procedural Analysis*, 97 HARV. L. REV. 931, 937 (1984).

35 See Roger A. Fairfax, Jr., *Delegation of the Criminal Prosecution Function to Private*

criminal cases in Massachusetts are brought by the state.³⁶ While the victims of crime may serve as essential witnesses in the criminal cases and may attempt to (and sometimes successfully) influence the decisions of individual prosecutors, victims have neither the capacity to file a criminal lawsuit nor the authority to control the direction a criminal case takes once a prosecutor has elected to bring charges against the offender.³⁷

The United States Supreme Court has recognized that “[t]he criminal justice system is not operated primarily for the benefit of victims, but for the benefit of society as a whole.”³⁸ The criminal law’s focus on the harm done to society is what distinguishes it from civil law, a distinction well-articulated by legal scholar Richard Nagareda:

A criminal sentence discharges the defendant’s figurative “debt to society” rather than a debt to a particular individual in the form of a civil judgment. The victims of crime get the psychological satisfaction of knowing that the law will punish the person who harmed them, or who attempted to do so, whereas the victims of torts get cash.³⁹

Given that the criminal legal system is designed from the

Actors, 43 U.C. DAVIS L. REV. 411, 435 (2009), https://lawreview.law.ucdavis.edu/issues/43/2/articles/43-2_Fairfax.pdf (“Regardless of its foggy origins in the United States, the prosecution of crime is now firmly entrenched as a public function.”). See also CESARE BONESANA DI BECCARIA, AN ESSAY ON CRIMES AND PUNISHMENTS (W.C. Little & Co. 1872) (1764) (arguing that “[t]he laws only can determine the punishment of crimes; and the authority of making penal laws can only reside with the legislator, who represents the whole society, united by the social compact”); Mary Margaret Gianni, Note, *The Swinging Pendulum of Victim’s Rights: The Enforceability of Indiana’s Victims’ Rights Laws*, 34 IND. L. REV. 1157, 1160 (2001), <https://mckinneylaw.iu.edu/ilr/pdf/vol34p1157.pdf> (“America’s criminal justice system has been almost entirely dominated by the public prosecution model.”).

36 Casson III, *supra* note 30, at 349 (“Victims of crime can neither compel the commencement of a prosecution nor insist on its termination. The case, if prosecuted, will be brought not in the name of the victim but in the name of ‘The People,’ ‘The United States of America,’ or of some similar abstraction.”).

37 *See id.*

38 Kelly v. Robinson, 479 U.S. 36, 52 (1986).

39 Richard A. Nagareda, *Outrageous Fortune and the Criminalization of Mass Torts*, 96 MICH. L. REV. 1121, 1175 (1998).

viewpoint of society as a whole,⁴⁰ restitution—with its singular focus on the individual victim—represents a significant departure from those principles.⁴¹ Nonetheless, in the view of its proponents, restitution properly fits within criminal law.⁴² According to proponents, even though criminal law focuses on protecting the public, that aim is not exclusive: criminal law also maintains a concern for the individual victims of crime.⁴³ Moreover, consistent with the goals of criminal law, restitution may promote personal responsibility, self-respect, and ultimately the rehabilitation of the defendant.⁴⁴ As a psychological matter, restitution forces the

40 See Richard A. Epstein, *Crime and Tort: Old Wine in Old Bottles*, in ASSESSING THE CRIMINAL: RESTITUTION, RETRIBUTION, AND THE LEGAL PROCESS 231, 233 (Randy E. Barnett & John Hagel III eds., 1977) (defining the prevailing approach to crime); Kleinfeld, *supra* note 30, at 1512 (“To make theoretical sense of criminal law’s status as public law, crime *has* to be regarded as an attack on society.”).

41 For an overview of the historical development of the criminal legal system, particularly the shift from victim-based responses to crime to sovereign-focused responses over multiple centuries, see generally Richard C. Boldt, *Restitution, Criminal Law, and the Ideology of Individuality*, 77 J. CRIM. L. & CRIMINOLOGY 969, 981-1007 (1986).

42 *Id.* at 975 (citing Casson III, *supra* note 30, at 353-54; Paul Keve, *Therapeutic Uses of Restitution*, in OFFENDER RESTITUTION IN THEORY AND ACTION 59, 60-64 (1978); Albert Eglash, *Creative Restitution: Some Suggestions for Prison Rehabilitation Programs*, 20 AM. J. CORRECTIONS 20, 20-22 (1958); Stephen Schafer, *Compensation of Victims of Criminal Offenses*, 10 CRIM. L. BULL. 605, 605, 618, 622 (1974)).

43 See e.g., *Snyder v. Massachusetts*, 291 U.S. 97, 122 (1934), *overruled on other grounds by Malloy v. Hogan*, 378 U.S. 1 (1964) (stating that “justice, though due to the accused, is due to the accuser also”); Mary Margare Giannini, *Equal Rights for Equal Rites?: Victim Allocation, Defendant Allocation, and the Crime Victims’ Rights Act*, 26 YALE L. & POL’Y REV. 431, 433 (2007); Anne M. Morgan, *Criminal Law – Victim Rights: Remembering the ‘Forgotten Person’ in the Criminal Justice System*, 70 MARQ. L. REV. 572, 574 (1987) (explaining that “[d]espite the distinction now drawn in this area between criminal and civil proceedings, the view that the state should assist victims of crime has historical support”); Note, *Victim Restitution in the Criminal Process: A Procedural Analysis*, *supra* note 34, at 935-36; Richard E. Laster, *Criminal Restitution: A Survey of its Past History and an Analysis of its Present Usefulness*, 5 U. RICH. L. REV. 71, 97 (1970) (“Although the victim’s present position is a shadow compared to his primitive forefathers, consideration for the harm suffered by the victim is still a part of the present system of criminal justice.”).

44 *Kelly v. Robinson*, 479 U.S. 36, 52 (1986) (explaining that restitution furthers the penal goal of rehabilitating offenders); Boldt, *supra* note 41, at 975-76 (explaining that restitution furthers the offender’s “reintegration into the community”); Anne Larason Schneider & Jean Shumway Warner, *The Role of Restitution in Juvenile Justice Systems*, 5 YALE L. & POL’Y REV. 382, 383 (1986);

defendant to take responsibility for his actions and acknowledge the victim's loss.⁴⁵ As a practical matter, restitution incentivizes the offender to maintain employment.⁴⁶ Being employed is credited with occupying an offender's time in a constructive manner, breaking his relationship with negative influences and opening the door to more constructive relationships.⁴⁷ In addition, offenders benefit from the kind of structured supervision that often accompanies restitution.⁴⁸

Proponents have argued that restitution has the added benefit of reducing the likelihood that victims seek personal revenge outside of the criminal system.⁴⁹ Prior to the development of modern criminal law, victims of wrongful conduct simply retaliated against the offender or had their families or tribes retaliate.⁵⁰ This often led to an endless cycle of violence, as each retaliation gave rise to a counter-retaliation.⁵¹ Modern criminal law exists in part to prevent a full return to that state of affairs.⁵² It is thought that victims who are

Note, *Victim Restitution in the Criminal Process: A Procedural Analysis*, *supra* note 34, at 938-39.

45 See, e.g., Kleinfeld, *supra* note 30, at 1528-29 ("Retributivism's concern for dignity is oriented to offenders: retributivism holds that offenders' dignity requires that they be treated as agents with the capacity for freedom and choice, which also means holding them responsible for their wrongs."); Commonwealth v. Avram A., 982 N.E.2d 548, 553 (Mass. App. Ct. 2013) (noting that "[t]he judge sought to instill a 'very important moral lesson' through the restitution order . . . the judge properly sought to teach the juvenile one of life's primary lessons: he is responsible for the actions he takes. Such an order not only provided an opportunity to build the juvenile's character and integrity, but also to promote his life as a law-abiding citizen. The judge's order had the added benefit of instilling in the juvenile the important values of respect for others (as well as their property) and a basic understanding of the value of work and money.").

46 See Schneider & Warner, *supra* note 44, at 398 ("The work required to make enough money to meet the restitution payments occupies much of the juvenile's time.").

47 *Id.*

48 *Id.* at 383.

49 Laster, *supra* note 43, at 71-72.

50 *Id.* ("At an early era in the history of mankind, it would have been difficult to distinguish victim from criminal because a person injured by a violent attack would take personal revenge against his attacker, and their original positions would then be reversed . . . [r]evenge was a personal or familiar matter, and not until societal attempts were made to limit its harmful effects did the concept of order in a structured community commence.").

51 *Id.*

52 See *id.* at 82 (noting that "to prevent the victim from attempting to circumvent the system of criminal justice, statutes have been passed to prevent him from compounding a crime").

receiving monetary compensation are less likely to pursue private revenge.⁵³ In this way, restitution promotes an orderly society.⁵⁴

B. Criticisms of Restitution

For their part, critics have raised multiple challenges to the legitimacy of criminal restitution orders. Critics raise these arguments without necessarily disputing the fundamental premise that victims should receive compensation for their losses: “To say that fairness or mercy demands that victims be compensated for their losses is not to say that this compensation must necessarily come from the cost imposer himself.”⁵⁵ The thrust of the criticism is that criminal punishment is not the appropriate vehicle for compensating victims.⁵⁶

First, critics have argued that victim compensation traditionally belongs in civil courts and is a hazardous preoccupation for criminal courts.⁵⁷ The division between criminal law and civil

53 Kelly, *supra* note 26, at 693. For instance, the Massachusetts Suffolk County District Attorney’s Office released a news announcement celebrating a \$30,000 restitution order against a defendant pleading guilty to a child pornography charge. *Suffolk DA Wins \$30K Restitution Order against Child Pornography Defendant*, SUFFOLK COUNTY DISTRICT ATTORNEY (Apr. 14, 2014), <http://www.suffolkdistrictattorney.com/suffolk-da-wins-30k-restitution-order-against-child-pornography-defendant/>. The District Attorney believed the order to be “the largest restitution order ever imposed on a child pornography defendant by a Massachusetts court.” *Id.* The District Attorney justified the order in this way: “Research shows that just knowing images of abuse are still circulating can make it almost impossible for victims to bring closure to the crimes against them. An order like this one can help identified victims make progress in their recovery that’s crucial to living a productive life.” *Id.*

54 Adelstein, *supra* note 30, at 167 (“[P]unishment provides a public good to the dispersed class of indirect victims, the satisfaction of their individual desire for vengeance, that roughly compensates them ‘in kind’ for the moral costs they have borne.”).

55 *Id.* at 172.

56 *Id.*

57 See Randy E. Barnett, *The Justice of Restitution*, 25 AM. J. JURIS. 117, 119 (1980), <http://scholarship.law.georgetown.edu/cgi/viewcontent.cgi?article=2569&context=facpub> (summarizing a criticism of the restitutive approach to punishment in this way: “The differences in the *prima facie* cases in each system and the remedies imposed can only be explained by their different functions. The purpose of the (civil) tort law is to compensate individuals who have been harmed by a rights invasion; the purpose of the criminal law is to punish individuals who have revealed a *mens rea* or bad intention to invade rights by an overt act which attempts (successfully or not) to carry out this intention . . . the function of the civil law is to restore the victim.”) (internal citation omitted); Katherine Beckett & Naomi Murakawa, *Mapping the Shadow*

law is the fundamental organizing principle of the American legal system.⁵⁸ Criminal law exists to address violations of public safety through the imposition of punishment, particularly state-imposed restrictions on liberty; civil law exists to coordinate the orderly compensation of monetary losses suffered by private individuals and entities.⁵⁹ Bringing the awesome weight of the state's incarceration powers into matters of debt collection erodes that boundary.⁶⁰

According to critics, the concerns regarding the use of the state's carceral power to compel payments are particularly alarming because most defendants have no ability to pay.⁶¹ Thus, requiring

Carceral State: Toward An Institutionally Capacious Approach to Punishment, 16 THEORETICAL CRIMINOLOGY 221, 238 (2012), https://www.princeton.edu/~murakawa/publications/Beckett_Murakawa_Mapping-the-Shadow-Carceral-State.pdf (arguing that criminal court debt has improperly expanded the carceral state and that "the carceral state ensnares and sanctions millions, by, for example . . . imposing debt and sanctioning nonpayment"); Note, *Victim Restitution in the Criminal Process: A Procedural Analysis*, *supra* note 34, at 935; Alexes Harris et al., *Drawing Blood from Stones: Legal Debt and Social Inequality in the Contemporary United States*, 115 AM. J. SOC. 1753, 1757 (2010), <http://faculty.washington.edu/kbeckett/articles/AJS.pdf> (noting that "monetary sanctions other than fines are neither intended as punishment nor legitimated in traditional penological terms and are therefore better understood as civil penalties than as monetary 'sanctions'").

58 Note, *Victim Restitution in the Criminal Process: A Procedural Analysis*, *supra* note 34, at 932.

59 *Id.* at 931-32.

60 Cortney E. Lollar, *Child Pornography and the Restitution Revolution*, 103 J. CRIM. L. & CRIMINOLOGY 343, 347 (2013), <http://scholarlycommons.law.northwestern.edu/jclc/vol103/iss2/1> ("Allowing personal vindication to be a goal of restitution undermines the traditional distinction between civil and criminal law, positioning the criminal justice system as a tool for personal retribution rather than societal protection."). For an example of a court acknowledging certain concerns regarding a judge conditioning the resolution of a case on the defendant making payments to the complaining witnesses, although still concluding that financial penalties are nonetheless appropriate when issued based on proper considerations, see Commonwealth v. Rotonda, 747 N.E.2d 1199, 1206-07 (Mass. 2001) ("Private payments exchanged for releases from criminal responsibility erode, if not completely erase, the demarcation between the criminal and civil systems of justice and 'benefit the individual at the expense of defeating the course of public justice.' Moreover, such payments create the perception that a class-based criminal justice system exists and that those with resources may buy their way out of criminal liability.").

61 Adelstein, *supra* note 30, at 171 ("Offenders are frequently without the pecuniary resources to compensate their direct victims . . ."); David Cole, *No Equal Justice*, 1 CONN. PUB. INT'L L. J. 19, 23 (2001), <http://scholarship.law.georgetown.edu/cgi/viewcontent.cgi?article=1347&context=facpub> ("The

defendants to pay to resolve their cases raises constitutional problems.⁶² First, equal protection concerns: poor defendants are more harshly punished by monetary sanctions.⁶³ A poor defendant may find himself lingering in the criminal system simply because he is unable to pay.⁶⁴ Critics also argue that restitution raises due process concerns, because *inter alia* ordering financial compensation as a criminal sanction denies the defendant all of the procedural protections that would apply if the compensation were sought through a civil trial for damages.⁶⁵ Concerns regarding the exercise of criminal court power to coerce poor defendants into paying are not theoretical: certain proponents of restitution have openly argued that victim compensation should remain in criminal courts because “criminal courts can use probation officers and the threat of revocation to induce probationers and parolees to pay their

vast majority of those behind bars are poor; 40 percent of state prisoners can’t even read.”); Helen A. Anderson, *Penalizing Poverty: Making Criminal Defendants Pay for Their Court-Appointed Counsel through Recoupment and Contribution*, 42 U. MICH. J. L. REFORM 323, 332 (2009) (“Defendants, whose job prospects are dim and who are often laboring under massive accrued child-support debt, generally do not have the resources to pay restitution, fines, fees, and other court costs, let alone recoupment debt.”).

62 See *Policing and Profit*, 128 HARV. L. REV. 1723, 1737 (2015), <http://harvardlawreview.org/wp-content/uploads/2015/04/Policing-and-Profit.pdf> (“[P]rofit motives violate the requirements of neutrality, fairness, equality, and proportionality that constitutional criminal procedure — the law of policing — imposes on law enforcement.”); State v. Garner, 566 P.2d 1055, 1057 (Ariz. Ct. App. 1977) (noting that “[g]reat constitutional problems develop if the amount of reparations is an amount larger than the defendant can pay. When this occurs, and the defendant is later incarcerated for his failure to pay, we have what may be an imprisonment for debt problem.”).

63 R. Barry Ruback, *The Benefits and Costs of Economic Sanctions: Considering the Victim, the Offender and Society*, 99 MINN. L. REV. 1779, 1781 (2015), http://www.minnesotalawreview.org/wp-content/uploads/2015/09/Ruback_5fmt_PDF.pdf [hereinafter Ruback, *Benefits and Costs*] (“The primary criticism of monetary sanctions is that they ‘are inherently inequitable’ in that they are unfair to many offenders who have little or no money but who face charges of hundreds or thousands of dollars, which they are unlikely to ever be able to pay.”).

64 *Id.*

65 Note, *Victim Restitution in the Criminal Process: A Procedural Analysis*, *supra* note 34, at 931-946 (explaining critics’ argument that “[u]sing sentencing procedures to determine the amount of restitution . . . violates due process by depriving the defendant of various procedural safeguards”). For a general statement on due process, see Halbert v. Michigan, 545 U.S. 605, 610-11 (2005) (explaining that “[t]he due process concern homes in on the essential fairness of the state-ordered proceedings”).

restitution orders.”⁶⁶

Critical literature has also drawn attention to the historical relationship between monetary sanctions and racism.⁶⁷ Legal scholars Alexes Harris, Heather Evans, and Katherine Beckett have traced monetary sanctions in the criminal process back to the “aftermath of slavery” and, more specifically, the convict lease system that existed in the United States through the 1940s.⁶⁸ Under that system, local jurisdictions “leased” prisoners who were saddled by fines and fees to corporations who exploited the prisoners’ free labor in various settings including farm plantations.⁶⁹ Imposing fines and fees on criminal defendants created a reliable source of extremely cheap labor.⁷⁰ Harris, Evans, and Beckett argued that, “[m]onetary sanctions were thus integral to systems of criminal justice, debt bondage, and racial domination.”⁷¹ While public narratives at the time suggested that the individuals being saddled with criminal debt and then leased were hardened criminals, “the records demonstrate the capture and imprisonment of thousands of random indigent citizens” and “the original records of county jails indicated thousands of arrests for inconsequential charges.”⁷² Critics contend that criminal debt continues to disproportionately affect poor people of color, particularly Black people, because they “face a far greater risk of being fined, arrested, and even incarcerated for minor offenses than other Americans.”⁷³ In this view, criminal debt is intimately tied to the nation’s history of institutionalized racism.

Overall in this debate, proponents of restitution have

66 Bradford C. Mank, *The Scope of Criminal Restitution: Awarding Unliquidated Damages in Sentencing Hearings*, 17 CAP. U. L. REV. 55 (1987), http://scholarship.law.uc.edu/cgi/viewcontent.cgi?article=1257&context=fac_pubs; Ruback, *Benefits and Costs*, *supra* note 63, at 1808 (writing that “offenders will make payments if they are threatened with probation revocation for nonpayment . . . [t]hreats can work if they are real, that is, if the threat is followed with action”).

67 Harris et al., *supra* note 57, at 1758.

68 *Id.*

69 *Id.*

70 *Id.*

71 *Id.*

72 Douglas A. Blackman, *Excerpt: Slavery by Another Name*, NPR (2008), <http://www.npr.org/templates/story/story.php?storyId=89051115>.

73 KAREN DOLAN & JODI L. CARR, INST. FOR POLICY STUDIES, THE POOR GET PRISON: THE ALARMING SPREAD OF THE CRIMINALIZATION OF POVERTY 6 (2015), <http://www.ips-dc.org/wp-content/uploads/2015/03/IPS-The-Poor-Get-Prison-Final.pdf>.

prevailed. Restitution has long been ordered by criminal courts.⁷⁴ The 1960s and 1970s, in particular, saw a substantial increase in the frequency with which state and federal judges ordered restitution.⁷⁵ During those years, the victims' rights movement grew⁷⁶ and included an increased demand for victim compensation.⁷⁷ In keeping with those trends, in 1961 the Supreme Judicial Court proclaimed that “[r]éparation or restitution in criminal cases is consonant with the public policy of the Commonwealth.”⁷⁸

II. Structure: The Legal Framework Around Restitution

A. The Legal Basis for Court Authority to Order Restitution

In Massachusetts, there does not appear to be a generally applicable statute expressly authorizing courts to order restitution.⁷⁹ Instead, in identifying the source of legal authority to order restitution,⁸⁰ the Supreme Judicial Court has pointed first to the “great latitude” judges are granted in sentencing.⁸¹ It is well established

74 For example, see the Pretrial Services Act, 18 U.S.C. § 3152 (1982), which authorized federal courts to impose restitution as a condition of criminal probation.

75 See Lollar, *supra* note 60, at 351; Schneider & Warner, *supra* note 44, at 389-90 (noting that the National Juvenile Restitution Program sponsored by an office of the Department of Justice was announced in 1976). *See also* STEPHEN SCHAFER, COMPENSATION AND RESTITUTION TO VICTIMS OF CRIME 12 (2d ed. 1970) (stating that “[h]istory suggests that growing interest in the reformation of the criminal is matched by decreasing care for the victim”).

76 Lollar, *supra* note 60, at 351. *See also* Edward McIntyre, *Scores of Victims, Years of Neglect: A View of the Massachusetts Compensation for Victims of Violent Crimes Statute, c. 258A*, 16 NEW ENG. L. REV. 45, 70 (1980) (advocating for revamping of system for compensating victims).

77 Lollar, *supra* note 60, at 351.

78 Novelty Bias Binding Co. v. Solomon Shevrin, 175 N.E.2d 374, 376 (Mass. 1961).

79 There do exist crime-specific statutes permitting restitution in cases involving that particular crime. *See, e.g.*, MASS. GEN. LAWS ch. 266, § 108 (2017) (providing for restitution where a person is convicted of destroying a ship or vessel); MASS. GEN. LAWS ch. 276, § 92A (2017) (providing for restitution where a person is convicted of motor vehicle theft); MASS. GEN. LAWS ch. 119, § 62 (2017) (providing for juvenile courts to order restitution as a condition of probation on a juvenile who has been found guilty of committing “an act involving liability in a civil action”).

80 Commonwealth v. Denehy, 2 N.E.3d 161, 174 n.20 (Mass. 2014) (stating that “[a]lthough there are no explicit statutory limitations guiding the imposition of restitution, we recognize several judicially imposed restrictions”).

81 Commonwealth v. McIntyre, 767 N.E.2d 578, 582 (Mass. 2002) (quoting

that “as long as the sentence imposed is within the limits provided by the statute under which the defendant is convicted” and is within constitutional limits, trial court judges maintain great discretion in crafting a sentence.⁸² Judges are granted this latitude so that sentences are individually-tailored to the particular defendant and “reflect the judge’s careful assessment of several goals: punishment, deterrence, protection of the public, and rehabilitation.”⁸³

Second, the Supreme Judicial Court has identified three statutes that provide courts the authority to impose probation.⁸⁴ Most prominently, Mass. Gen. L. c. 276, § 87 grants trial court judges the authority to place a defendant on probation.⁸⁵ Probation can be imposed either with the defendant’s permission pretrial or on the court’s own accord after a verdict of guilt.⁸⁶ Probation is a punishment that allows the defendant to avoid incarceration by instead being placed on conditional release back in the community for a period of time set by the court.⁸⁷ The defendant’s release is conditioned on compliance with whatever conditions the sentencing judge deems necessary to 1) ensure the community’s safety and 2) promote the defendant’s rehabilitation.⁸⁸ The judge may revoke probation if she determines by a preponderance of the evidence⁸⁹ that the defendant

Commonwealth v. Power, 650 N.E.2d 87, 89 (Mass. 1995)).

82 *Power*, 650 N.E.2d at 89 (citing Commonwealth v. Goodwin, 605 N.E.2d 827, 831 (Mass. 1993); Commonwealth v. Burke, 467 N.E.2d 846, 850–851 (Mass. 1987)). *See also* Commonwealth v. Ferguson, 571 N.E.2d 411, 415 (Mass. App. Ct. 1991) (citing United States v. Tucker, 404 U.S. 443, 446–47 (1972); Commonwealth v. Celeste, 264 N.E.2d 683, 685 (Mass. 1970)).

83 *Goodwin*, 605 N.E.2d at 831 (citing Cepulonis v. Commonwealth, 427 N.E.2d 17, 21 (Mass. 1981)).

84 *McIntyre*, 767 N.E.2d at 582 (pointing to MASS. GEN. LAWS ch. 276, §§ 87, 87A and MASS. GEN. LAWS ch. 279, § 1 as the statutes providing courts the authority to order conditions of probation, including restitution).

85 MASS. GEN. LAWS ch. 276, § 87 (2017) (“The superior court, any district court and any juvenile court may place on probation in the care of its probation officer any person before it charged with an offense or a crime . . .”).

86 *Id.*

87 *See Commonwealth v. Goodwin*, 933 N.E.2d 925, 930 (Mass. 2010) (citing Commonwealth v. Durling, 551 N.E.2d 1193, 1195 (Mass. 1990)).

88 *Goodwin*, 933 N.E.2d at 931 (“The only limitation is that, where a probation condition infringes on a defendant’s constitutional rights, it must be ‘reasonably related’ to the goals of sentencing and probation.”).

89 Commonwealth v. Holmgren, 656 N.E.2d 577, 578 (Mass. 1995) (“In a criminal case, of course, the Commonwealth must prove the elements of each crime charged beyond a reasonable doubt. In a probation revocation hearing the Commonwealth bears a lesser burden. Although the burden of proof in a revocation of probation proceeding has been variously described, it is proof

violated any of the conditions of his release.⁹⁰ Once a violation has been found, the judge has great discretion in deciding whether to incarcerate the defendant or return the defendant to probation with additional or revised conditions.⁹¹ If the judge decides to incarcerate the defendant, the incarceration can be for any amount of time up to the maximum sentence on the underlying conviction.⁹² Thus, individuals on probation live under the very real possibility of incarceration.⁹³

Sentencing judges are granted broad discretion in determining the conditions of probation for any particular defendant.⁹⁴ The Supreme Judicial Court has interpreted that discretion to include the authority to require defendants to make restitution payments while on probation.⁹⁵ The court has stated explicitly that “[r]estitution is . . . an appropriate condition of probation imposed by the judge within his sentencing discretion.”⁹⁶ In addition, by passing statutes that support victims’ rights to collect information on restitution orders, the Massachusetts Legislature appears to have embraced

by a preponderance of the evidence . . .”).

90 *Id.* at 579.

91 See Commonwealth v. Pena, 967 N.E.2d 603, 608 (Mass. 2012) (explaining that “a decision to revoke probation involves two distinct components. [T]he judge must determine, as a factual matter, whether the defendant has violated the conditions of his probation. If the judge determines that the defendant is in violation, he can either revoke the probation and sentence the defendant or, if appropriate, modify the terms of his probation. How best to deal with the probationer is within the judge’s discretion.” (quoting *Durling*, 551 N.E.2d at 1195)).

92 See Commonwealth v. Bruzzese, 773 N.E.2d 921, 929 (Mass. 2002) (explaining that “[i]f a defendant’s straight probation is revoked, whether it be on a single charge or on multiple charges, he is subject to sentencing on those charges in essentially the same light that existed at the time straight probation was originally imposed . . . [h]e may receive the maximum sentence on each conviction”).

93 See Commonwealth v. Wilcox, 841 N.E.2d 1240, 1245 (Mass. 2006) (“A probationer has only a conditional liberty interest.”).

94 MASS. GEN. LAWS ch. 276, § 87 (2017) (providing that probation shall be “for such time and upon such conditions as [the court] deems proper”). See also MASS. GEN. LAWS ch. 276, § 87A (2017) (identifying participation in rehabilitative programs and completion of community service as permissible conditions of probation; section 87A makes clear that trial judges are not limited to those terms of conditions).

95 Commonwealth v. McIntyre, 767 N.E.2d 578, 582 (Mass. 2002).

96 Commonwealth v. Nawn, 474 N.E.2d 545, 551 (Mass. 1985) (citing United States v. Satterfield, 743 F.2d 827, 836 (11th Cir. 1984)).

restitution.⁹⁷

B. The Broad Scope of Losses Recoverable Through Restitution

Restitution orders are limited by certain substantive requirements.⁹⁸ Not all losses are recoverable through restitution ordered as a criminal sanction.⁹⁹ Whether a loss is recoverable depends on two factors: (1) the nature of the loss¹⁰⁰ and (2) the degree to which the loss is connected to the crime.¹⁰¹ “[R]estitution is limited to . . . economic losses.”¹⁰² Specifically, it is limited to those “economic losses” that are “caused by the conduct of the defendant and documented by the victim.”¹⁰³ Massachusetts courts have reasoned that pain and suffering is not economic loss.¹⁰⁴ The category of losses that have expressly been recognized as proper to include in a restitution order, however, is wide.¹⁰⁵ Medical expenses, damage to property, lost property, lost pay, and even court-related travel expenses are recoverable in restitution orders.¹⁰⁶

An economic loss must be sufficiently connected to the defendant’s criminal conduct to be recoverable through restitution.¹⁰⁷ As a general matter, a criminal sanction of any kind must be tailored to the crime of conviction.¹⁰⁸ And “[l]ike any other criminal sanction, restitution best serves penal objectives when it bears a proper relationship to the crime of conviction, both in kind and proportion.”¹⁰⁹

97 See MASS. GEN. LAWS ch. 258B, § 3(o) (2012) (authorizing victim to inter alia request restitution as a component of resolution and receive information regarding the schedule of payments).

98 See Commonwealth v. Rotonda, 747 N.E.2d 1199, 1207 (Mass. 2001).

99 *Id.*

100 See Commonwealth v. Hastings, 756 N.E.2d 1188, 1189 (Mass. App. Ct. 2001).

101 Commonwealth v. McIntyre, 767 N.E.2d 578, 582 (Mass. 2002).

102 *Hastings*, 756 N.E.2d at 1189.

103 *Id.*

104 Commonwealth v. Buckley, 57 N.E.3d 1051, 1053 n.2 (Mass. App. Ct. 2016), *review denied*, 63 N.E.3d 386 n.2 (Mass. 2016).

105 Commonwealth v. Rotonda, 747 N.E.2d 1199, 1207 (Mass. 2001).

106 *Id.* (“[I]tems [such] as medical expenses, court-related travel expenses, property loss and damage, lost pay, or even lost paid vacation days required to be used to attend court proceedings might all be included in an appropriately documented restitution order.”).

107 Commonwealth v. McIntyre, 767 N.E.2d 578, 583-84 (Mass. 2002).

108 *Id.* at 582.

109 *Id.* (citing Richard Lowell Nygaard, *On the Philosophy of Sentencing: Or, Why*

However, as the following cases will show, Massachusetts courts have interpreted the connection requirement with some leniency. In determining whether the loss is causally connected to the crime, Massachusetts courts are permitted to look beyond the crime of conviction—that is, the crime the defendant was actually found guilty of committing.¹¹⁰ A loss may be included in the restitution order so long as the loss resulted from the underlying facts that are before the court.¹¹¹ This means that restitution orders can include losses arising from conduct that the defendant was not specifically charged with and conduct on which the jury did not necessarily return a verdict.¹¹²

Commonwealth v. McIntyre well demonstrates the broadness of this rule. In *McIntyre*, the Supreme Judicial Court held that a restitution order could include damage to the victim’s car even though the defendant was not found guilty of the crime of destruction of property.¹¹³ The defendant, who was convicted of assault and battery by means of a dangerous weapon, argued that the court was only permitted to order him to pay restitution for those economic losses that directly flowed from the elements of the assault and battery charge.¹¹⁴ The court rejected that argument, reasoning that even though the damage to the car door was not a part of the crime of conviction, there was a sufficient connection because the defendant

Punish?, 5 WIDENER J. PUB. L. 237, 265-266 (1996).

110 *McIntyre*, 767 N.E.2d at 584 (indicating that courts must “look to the underlying facts of the charged offense, not the name of the crime [of which the defendant was convicted, or] to which the defendant entered a plea” (citing *State v. Landrum*, 832 P.2d 1359, 1364 (Wash. Ct. App. 1992))).

111 See, e.g., *Commonwealth v. Palmer* P, 808 N.E.2d 848, 850 (Mass. 2004) (ordering restitution to cover the costs of items allegedly taken during a breaking and entering even where the juvenile defendant was acquitted of larceny and explaining that “the breaking and entering by the defendant into the [victim’s] home bore a causal relationship to the taking of the personal property”); *McIntyre*, 767 N.E.2d at 584 (concluding that it was proper to include damage to victim’s car in restitution order where defendant was convicted of assault and battery by means of a dangerous weapon because defendant caused damage by kicking the car as victim attempted to flee the assault); *Commonwealth v. DeGaetano*, No. 05-P-32, 2006 WL 473807, at *1, *5 (Mass. App. Ct. Feb. 28, 2006) (affirming order of restitution for medical costs paid to the defendant after he falsified injuries, although the defendant was never charged with falsifying injuries and was convicted only of larceny and fraud).

112 See *McIntyre*, 767 N.E.2d at 584.

113 *Id.*

114 *Id.* at 581.

caused the damage to the car by kicking the car as the victim attempted to flee the assault.¹¹⁵ The fact that the damage to the door occurred during the assault was sufficient.¹¹⁶

Commonwealth v. Buckley is also illustrative of the leniency with which the “causal connection” requirement has been read. In *Buckley*, the Appeals Court made clear that restitution can include even losses that are the indirect result or the “foreseeable consequence[]” of the defendant’s conduct.¹¹⁷ That is, even where intervening third party negligence contributed to the victim’s economic loss, the defendant can still be ordered to pay restitution for the full amount of loss.¹¹⁸ The defendant in *Buckley* was convicted of larceny of a motor vehicle.¹¹⁹ When he was arrested, the defendant told the officers where the stolen car was and the officers recovered it.¹²⁰ However, for several months, this information was not shared with the victim.¹²¹ By the time the victim learned that the car was being held in an auto-body shop, the car had accumulated \$3,036 in storage fees and the victim had already purchased a replacement car.¹²² Although the victim could have recovered his car earlier and avoided purchasing a replacement car if the police officers had shared his car’s location, the trial judge set the restitution to the estimated book value of the replacement car.¹²³ The defendant appealed, arguing that the failure of the officers and the auto-body shop to inform the victim that the car had been recovered was an intervening act.¹²⁴ The court rejected that argument, reasoning that the restitution order could include the full cost of the car because third party acts of “ordinary negligence” are foreseeable consequences of a defendant’s underlying criminal act and do not break the chain of causation between the defendant and the loss.¹²⁵ The court was satisfied that “the defendant set in motion a chain of events that resulted in the loss of the victim’s

115 *Id.* at 584.

116 *Id.*

117 *Commonwealth v. Buckley*, 57 N.E.3d 1051, 1056-1058 (Mass. App. Ct. 2016), *review denied*, 63 N.E.3d 386 n.2 (Mass. 2016).

118 *Id.* at 1056-1058.

119 *Id.* at 1054.

120 *Id.*

121 *Id.*

122 *Id.*

123 *Id.* at 1055.

124 *Id.* at 1054-55.

125 *Id.* at 1058.

vehicle.”¹²⁶

The relative leniency of the Massachusetts standard becomes apparent when compared to the standards other jurisdictions have adopted. Indeed, while reaching its decision in *McIntyre*, the Supreme Judicial Court acknowledged that other jurisdictions had resolved the issue differently.¹²⁷ Courts such as the United States Supreme Court and the Indiana Court of Appeals have adopted a more strict approach, allowing recovery through restitution for only those losses that flow from the specific crime the defendant was convicted of.¹²⁸ While acknowledging that the rule might lead to victims receiving less than full compensation for their losses, the Supreme Court’s reasoning suggested that any under-compensation reflected the many bargains, risk-calculations and concessions that characterize the entire criminal system.¹²⁹ Another cluster of courts allow recovery for whatever losses the defendant’s conduct proximately caused.¹³⁰ The Supreme Court of Arizona, one of the courts employing this approach, explained that “losses [that] would not have occurred without the concurrence of a second causal event . . . constitute indirect damages and cannot qualify for restitution.”¹³¹ Allowing restitution orders to include losses that are the indirect result of the defendant’s conduct “confuses damage causation with access to a particular source of recovery.”¹³²

This practice of awarding restitution for losses that go beyond the crime the defendant was convicted of has certainly drawn criticism. Professor of Law Richard C. Boldt, for example, has contended that “the limited justifications for the use of restitution do not form an adequate basis for ordering repayment beyond an offense of conviction.”¹³³ Boldt cautioned:

The primary aim of the system is to create and reinforce

126 *Id.* at 1056.

127 Commonwealth v. McIntyre, 767 N.E.2d 578, 583 n.3 (Mass. 2002).

128 *Id.* (citing *Hughey v. United States*, 495 U.S. 411, 413 (1990); *State v. Aubert*, 811 P.2d 44 (Idaho 1991); *Polen v. State*, 578 N.E.2d 755, 757-58 (Ind. Ct. App. 1991)).

129 *Hughey*, 495 U.S. at 421.

130 *State v. Wilkinson*, 39 P.3d 1131, 1133 (Ariz. 2002) (stating standard stricter than “but for” analysis); *State v. Bradley*, 637 N.W.2d 206, 215 (Iowa Ct. App. 2001); *People v. Flanagan*, 478 N.E.2d 666 (Ill. Ct. App. 1985).

131 *Wilkinson*, 39 P.3d at 1133.

132 *Id.* at 1134-35.

133 Boldt, *supra* note 41, at 1019.

an ideology of individuality, not to insure that all victims receive a full recovery. Both are important, perhaps necessary, functions of some institution or institutions in society. It is unreasonable, however, to expect the criminal justice system to do anything particularly well if it is asked to do too much simultaneously.¹³⁴

C. Procedural Requirements for a Restitution Order

Massachusetts courts have also laid out the procedural requirements for a restitution order. At base, “[t]he procedure used to determine the amount of restitution or reparation must be reasonable and fair.”¹³⁵ This mandate arises from the fundamental principle that “[p]ersons forced to settle their claims of right and duty through the judicial process must be given a meaningful opportunity to be heard.”¹³⁶ Accordingly, “[i]f the record reveals an arbitrary method of determining the amount of restitution, the order cannot stand.”¹³⁷

Defendants are entitled to a hearing at which the restitution amount is determined.¹³⁸ At the hearing, the Commonwealth carries the burden of proving the sum of the victim’s loss.¹³⁹ That loss must be proven by a preponderance of the evidence—not by the far more demanding beyond a reasonable doubt standard that is required in criminal trials adjudicating guilt.¹⁴⁰ The Commonwealth may introduce testimony from the victim and relevant experts.¹⁴¹ While the restitution order “must be supported by facts and evidence,”¹⁴² the rules of evidence are not strictly applicable to restitution hearings.¹⁴³

134 *Id.* at 1020.

135 Commonwealth v. Nawn, 474 N.E.2d 545, 550 (Mass. 1985).

136 *Id.* (citing *Boddie v. Connecticut*, 401 U.S. 371, 377 (1971)).

137 *Nawn*, 474 N.E.2d at 550 (“Because the defendant was not afforded any meaningful opportunity to challenge the amount of money ordered to be repaid, we vacate the sentence and remand for a new sentencing proceeding.”).

138 *See id.*

139 *Id.* (citing Victim and Witness Protection Act of 1982, Pub. L. 97-291, § 5(a), 96 Stat. 1253–1255 (1982)).

140 *Nawn*, 474 N.E.2d at 550.

141 *Id.* at 551 n.5.

142 Commonwealth v. McIntrye, 767 N.E.2d 578, 583 (Mass. 2002) (citing Commonwealth v. Rotonda, 747 N.E.2d 1199, 1207-08 (Mass. 2001)).

143 *See Commonwealth v. Morris M.*, 876 N.E.2d 462, 469 (Mass. App. Ct. 2007) (citing Commonwealth v. Casanova, 843 N.E.2d 699, 705 (Mass. App. Ct. 2006)). *See also Commonwealth v. Pimentel*, No. 09-P-557, 2010

It is permissible to admit hearsay so long as it is “reliable.”¹⁴⁴

At the hearing, the defendant is entitled to dispute the Commonwealth’s claim regarding the victim’s financial loss.¹⁴⁵ The defendant has a right to cross examine the victim regarding losses.¹⁴⁶ Additionally, the defendant must be granted the opportunity to present evidence or experts of his own.¹⁴⁷ To deny the defendant any of these basic procedural protections would render the restitution procedure “arbitrary” and the resulting order objectionable.¹⁴⁸ Still, there are undoubtedly limitations to a defendant’s rights regarding restitution: “[t]he hearing need not be elaborate,”¹⁴⁹ a defendant’s right to cross examine the victim during a restitution hearing is not absolute,¹⁵⁰ and a defendant is not entitled to a jury trial on restitution.¹⁵¹

WL 2268917, at *2 (Mass. App. Ct. June 8, 2010) (rejecting the defendant’s argument on appeal that the restitution order should be reversed because the Commonwealth had not complied with the rules of evidence at the restitution hearing).

144 *Morris M.*, 876 N.E.2d at 469. For an example of a judge expressly considering hearsay at a restitution hearing, see *Commonwealth v. Gomes*, No. 15-P-1315, 2016 WL 5681822, at *3 (Mass. App. Ct. Oct. 3, 2016), *review denied*, 65 N.E.3d 660 (Table) (Mass. 2016) (explaining that “[a]s to the second question, the judge did not abuse his discretion in determining that the hearsay statement of the victim’s estimate of the jewelry’s value was sufficiently reliable. The victim gave a detailed description of the missing jewelry at trial, including the facts that the earrings contained gemstones and the bracelets were twenty-two karat gold. The jury’s verdict demonstrates that they found the victim to be credible, a conclusion the judge was also entitled to reach.”).

145 *Nawn*, 474 N.E.2d at 550.

146 *Id.*

147 *Id.* (citing MASS. GEN. LAWS ch. 279, § 4B ¶ 4 (2017)).

148 See *Nawn*, 474 N.E.2d at 550 (vacating sentence where restitution was ordered without allowing the defendant an opportunity to challenge the amount of restitution sought).

149 *Id.*

150 There are exceptions to this “presumptive right” to cross examine the victim in the restitution setting. See *Commonwealth v. Molina*, 71 N.E.3d 117, 135–36 (Mass. App. Ct. 2017). The judge is permitted to deny the defendant the opportunity to call the victim as a witness in a restitution hearing where the victim’s compelled participation would cause trauma that “overcomes” the defendant’s presumptive right. *Id.* “If there is ‘good cause’ for the Commonwealth not to call a witness with personal knowledge to testify but to offer instead reliable hearsay or other evidence to establish the basis for its request for restitution, the requirements of due process are likely to be satisfied.” *Id.* (citations omitted).

151 *Commonwealth v. Denehy*, 2 N.E.3d 161, 174–75 (Mass. 2014) (describing “restitution as an entirely judicially determined penalty” and holding that the

D. Lower Court Practices Regarding Ability to Pay Prior to Henry

As noted in the opening paragraphs of this Article, *Henry* clarifies the importance of considering a defendant's ability to pay. However, that proclamation was not the first of its kind in Massachusetts jurisprudence. In 1985, in *Nawn v. Commonwealth*, the Supreme Judicial Court stated that a defendant's ability to pay was relevant to restitution.¹⁵²

A central question posed to the *Nawn* court was whether defendants are entitled to a hearing to challenge the amount of financial loss claimed in restitution.¹⁵³ As discussed, the court declared that defendants are entitled to a hearing on the amount of restitution.¹⁵⁴ Yet, in the same stroke the court held that the right to a jury does not attach to the restitution hearing.¹⁵⁵ Tension arises from the denial of a jury at the restitution hearing because Article 12 of the Massachusetts Declaration of Rights provides that defendants have a right to a jury trial in the proceedings adjudicating whether they are guilty of the charged crime¹⁵⁶ and Article 15 of the Declaration of Rights provides for a jury trial “[i]n all controversies concerning property, and in all suits between two or more persons, except in cases in which it has heretofore been otherways used and practiced.”¹⁵⁷ Given that restitution is part of a criminal case and it is a controversy regarding financial losses, it is unsettling to some that juries do not resolve restitution.¹⁵⁸

In *Nawn*, the Supreme Judicial Court explained this lack of a right to a jury on restitution.¹⁵⁹ The court pointed to the fact

lack of a jury trial does not violate the Sixth Amendment); *Nawn*, 474 N.E.2d at 551 (no jury trial on restitution).

152 *Nawn*, 474 N.E.2d at 550–51.

153 *Id.* at 550.

154 *Id.*

155 *Id.* at 551.

156 Commonwealth v. Osborne, 840 N.E.2d 544, 547 (Mass. 2006) (quoting MASS. CONST., DECLARATION OF RIGHTS, art. 12).

157 *Nawn*, 474 N.E.2d at 551.

158 Courtney E. Lollar, *What is Criminal Restitution?*, 100 IOWA L. REV. 93, 149 (2015), <https://ilr.law.uiowa.edu/assets/Uploads/ILR-100-1-Lollar.pdf> (“Now that the Supreme Court again has acknowledged criminal restitution’s role as punishment, it should take the next step and recognize the constitutional protections that must adhere to criminal restitution [sic] light of its punitive character . . . Criminal restitution should be subject to the jury trial right of the Sixth Amendment, just as any other criminal punishment is.”).

159 *Nawn*, 474 N.E.2d at 551.

that restitution is a matter of sentencing.¹⁶⁰ Because judges rather than juries decide sentencing, the defendant is not entitled to a jury on restitution.¹⁶¹ In addition, the court stressed that restitution is “not a civil adjudication of damages.”¹⁶² In discussing a restitution order, the Supreme Judicial Court explained that the “controversy . . . is at heart a criminal one, focusing on adjudicating the guilt or innocence of the defendant.”¹⁶³ Unlike civil damages, “[t]he amount of restitution is not merely the measure of the value of the goods and money stolen from the victim by the defendant.”¹⁶⁴ The court instructed that “in a criminal case, the judge must . . . decide the amount that the defendant is able to pay and how such payment is to be made.”¹⁶⁵ Moreover, in setting restitution, lower court judges are permitted to consider “the defendant’s employment history and financial prospects.”¹⁶⁶ The court noted that such factors are not “central” in a civil suit for damages.¹⁶⁷

Through these statements, the Supreme Judicial Court implied that consideration of a defendant’s ability to pay is essential to the very definition of restitution.¹⁶⁸ Pursuant to the court’s reasoning, consideration of ability to pay is a fundamental distinction between restitution and civil damages.¹⁶⁹

However, the *Nawn* Court did not provide clear instruction as to how lower courts were to incorporate ability to pay into the restitution order.¹⁷⁰ The court neither provided a formula for measuring ability to pay nor laid out how much weight to give to ability to pay.¹⁷¹ The court did not expressly prohibit restitution orders that exceeded a defendant’s ability to pay.¹⁷² Indeed, in 2006, Justice Peter Agnes, a judge of the Massachusetts Superior Court, drew attention to the lack of clarity: “Neither state nor federal law contains a formula to guide courts in balancing the various factors

160 *Id.*

161 *Id.*

162 *Id.*

163 *Id.*

164 *Id.*

165 *Id.*

166 *Id.*

167 *Id.*

168 *See id.*

169 *Id.*

170 *See id.* at 550-51.

171 *See id.*

172 *See id.*

that come into play in order to determine a just order for the payment of restitution.”¹⁷³

In the years following *Nawn*, indigent criminal defendants attempted to rely on *Nawn* to challenge restitution orders that failed to properly account for the defendants’ ability to pay. Where judges affirmatively refused to consider ability to pay after defendants raised it, defendants tended to find appellate success in disrupting the restitution order.¹⁷⁴ For instance, in *Commonwealth v. Desouza*, the lower court failed to enter a finding that the defendant was able to pay the \$13,000 restitution amount ordered.¹⁷⁵ The lower court judge stated that “[t]he question of whether or not [the defendant] has the ability at this time to pay the money is not the issue.”¹⁷⁶ The Appeals Court disapproved of that approach: “Although the record does not indicate that the defendant is unable to make reasonable payments towards restitution, our cases require that the judge make specific findings on this issue [of ability to pay].”¹⁷⁷ Accordingly, the Appeals Court vacated the restitution order and remanded.¹⁷⁸

But success in challenging restitution orders proved limited. Many appellate courts read *Nawn* narrowly; restitution orders that were not supported by an express finding by the judge that the defendant had the ability to pay the restitution were oftentimes upheld.¹⁷⁹ After *Nawn*, lower court judges repeatedly failed to enter an express finding regarding ability to pay and on appeal the courts often excused that failure by faulting the defendant.¹⁸⁰ *Commonwealth*

173 *Commonwealth v. Stevanovich*, No. 2006-0086, 2006 WL 6493485, at *4 (Mass. Super. Ct. May 25, 2006).

174 See, e.g., *Commonwealth v. Desouza*, No. 07-P-1884, 2009 WL 7708 (Mass. App. Ct. Jan. 2, 2009).

175 *Id.* at *2.

176 *Id.* at *1.

177 *Id.* at *2.

178 *Id.*

179 See e.g., *Commonwealth v. Norman*, No. 10-P-492, 2011 WL 11345, at *1 (Mass. App. Ct. Jan. 4, 2011).

180 See, e.g., *Commonwealth v. Avram A.*, 982 N.E.2d 548, 552 (Mass. App. Ct. 2013) (“To satisfy its burden to prove the dollar amount of the damage caused by the juvenile, the Commonwealth offered the testimony of two witnesses and four exhibits. Although given the opportunity, as the judge found, the juvenile did not offer any evidence on his ability to pay.”); *Commonwealth v. Spaulding*, No. 08-P-217, 2009 WL 395375, at *1 (Mass. App. Ct. Feb. 19, 2009); *Commonwealth v. Orin O.*, No. 04-P-1126, 2005 WL 2649169, at *2 (Mass. App. Ct. Oct. 17, 2005) (rejecting argument that judge’s failure to consider ability to pay rendered the restitution order defective and explaining that “[n]otwithstanding the opportunity to present such

v. *Spaulding* is illustrative of this dynamic. In *Spaulding*, the Appeals Court rejected the defendant's argument that the lower court's failure to consider the defendant's ability to pay rendered the restitution order defective.¹⁸¹ According to the Appeals Court, the defendant had "waived her right to challenge the 'ability to pay' aspect of the order of restitution" because the defendant had not affirmatively raised ability to pay at the restitution hearing or raised any objection when she received notice of the condition of probation.¹⁸² The restitution order was not vacated.¹⁸³

Lower courts also reasoned that where a defendant had pleaded guilty to the offense giving rise to restitution, a defendant's ability to pay the restitution was implicit in the acceptance of the plea bargain.¹⁸⁴ In such cases, appellate courts upheld restitution orders even though the lower court judges had not inquired into the defendants' abilities to pay the restitution amount included in their pleas.¹⁸⁵ For instance, the Appeals Court embraced this reasoning in *Commonwealth v. Payne*.¹⁸⁶ There, the defendant failed to pay restitution as ordered and, without holding a hearing on ability to pay, the judge jailed the defendant.¹⁸⁷ In affirming, the Appeals Court declared that "the plea bargain itself was tantamount to a representation by [the defendant] that he could pay the restitution."¹⁸⁸ The court record reflected that "[b]ut for the plea bargain, the . . . judge would have

information concerning his alleged inability to pay, the juvenile presented no evidence."). Indeed, courts even faulted the defendant where both a hearing on restitution and a finding on the defendant's ability to pay were lacking. *See, e.g.*, Commonwealth v. Casserly, 501 N.E.2d 540, 541 (Mass. App. Ct. 1986) (stating that "what occurred amounted to an informed waiver of the opportunity to be heard further" where the defendant "did not request a hearing on the amount of restitution").

181 *Spaulding*, 2009 WL 395375 at *1.

182 *Id.*

183 *Id.*

184 *See, e.g.*, Commonwealth v. Giarrusso, No. 12-P-490, 2013 WL 967547, at *1 (Mass. App. Ct. Mar. 14, 2013) (citing to *Payne* and affirming revocation of probation and sentence of incarceration for failure to pay restitution because defendant's acceptance of the plea constituted a representation that she had the ability to pay the restitution); *Norman*, 2011 WL 11345, at *1 ("While a judge must consider a defendant's ability to pay when imposing a restitution scheme, the defendant here accepted the \$202,459 restitution as part of a plea bargain in 2005.") (citations omitted).

185 *See, e.g.*, *Giarrusso*, 2013 WL 967547, at *1; *Norman*, 2011 WL 11345, at *1.

186 Commonwealth v. Payne, 602 N.E.2d 594, 597 (Mass. App. Ct. 1992).

187 *Id.* at 595.

188 *Id.* at 597.

opted for some imprisonment.”¹⁸⁹ In the Appeals Court’s view, the defendant “was offered the privilege of making restitution as part of a plea bargain, with the express condition that, should he not hold up his end of the bargain, a sentence of incarceration would be imposed.”¹⁹⁰

In addition, lower courts focused much attention regarding ability to pay on the payment schedule attached to restitution orders, rather than on the overall amount of restitution ordered.¹⁹¹ Many lower courts reasoned that to the extent they were required to consider a defendant’s ability to pay, that responsibility could be delegated to the probation officer and satisfied by a lengthier payment schedule.¹⁹² For instance, in *Commonwealth v. Yeshulas*, the Appeals Court rejected the defendant’s argument that the restitution order was improper because it failed to consider the defendant’s ability to pay.¹⁹³ The court explained: “contrary to the defendant’s argument, the probation department does have discretion to ‘consider the defendant’s financial resources in setting the [restitution] payment schedule.’”¹⁹⁴ The court pointed to the fact that “[t]he order as entered on the docket reads as follows: ‘Restitution in the amount of \$2,500 payable at a rate to be determined by [the] Probation [Department].’”¹⁹⁵

Still more, lower courts repeatedly extended defendants’ probation for the sole purpose of collecting more restitution.¹⁹⁶

189 *Id.*

190 *Id.*

191 See, e.g., *Commonwealth v. Faisal F.*, No. 13-P-1298, 2014 WL 1343288, at *1 (Mass. App. Ct. Apr. 7, 2014) (rejecting defendant’s argument that the judge failed to consider ability to pay and noting that the restitution order allowed for restitution to be “payable at a rate to be determined by the Probation Department”); *Commonwealth v. Yeshulas*, 746 N.E.2d 587, 592-93 (Mass. App. Ct. 2001); *Commonwealth v. Rescia*, 688 N.E.2d 1026, 1028 (Mass. App. Ct 1998).

192 See, e.g., *Faisal F.*, 2014 WL 1343288, at *1; *Yeshulas*, 746 N.E.2d at 592-93.

193 *Yeshulas*, 746 N.E.2d at 592.

194 *Id.* at 593 (alteration in original).

195 *Id.* (alteration in original).

196 See, e.g., *Commonwealth v. Avram A.*, 982 N.E.2d 548, 551-54 (Mass. App. Ct. 2013) (extending probation to allow 12-year-old more time to pay \$1,313.78 in restitution); *Commonwealth v. Norman*, No. 10-P-492, 2011 WL 11345, at *1 (Mass. App. Ct. Jan. 4, 2011) (describing an extension of probation so that restitution could be collected as a reasonable alternative to incarceration). There are also instances where courts indicated that they were open to extending probation for the purpose of allowing the defendant to pay more restitution. See, e.g., *Commonwealth v. LeBlanc*, No. 11-P-147, 2012

Commonwealth v. Norman is emblematic of this practice.¹⁹⁷ In *Norman*, the lower court judge found the defendant in violation of probation for not having made full payment on outstanding restitution; as a result, the judge extended the defendant's probation three years and ordered the defendant to pay the restitution in full.¹⁹⁸ The defendant appealed, arguing *inter alia* that she was indigent, was unable to pay the restitution, and had not willfully violated any term of her probation.¹⁹⁹ The court record included the defendant's testimony that she was unemployed and did not have savings, credit cards, or investments.²⁰⁰ The question before the Appeals Court was strictly whether the lower court committed error by extending the defendant's probation to permit her to pay more restitution.²⁰¹ The Appeals Court found no error.²⁰² The Appeals Court explained there may have been error had the judge incarcerated the defendant based on her inability to pay, but the extension of probation was permissible.²⁰³

In resolving challenges based on *Nawn*, lower courts also suggested that if a defendant was unable to pay the restitution, he could raise his inability as a defense at any final surrender hearing that alleged failure to pay restitution.²⁰⁴ In effect, a defendant's ability to pay was relegated to a defense available to the indigent rather than a limitation on the judge's discretion in setting the restitution order.²⁰⁵ In *Commonwealth v. Morris M.*, for instance, the defendant argued that the restitution order could not stand because his ability to pay had not been considered when the restitution order was initially issued.²⁰⁶ The Appeals Court disagreed.²⁰⁷ In addition to pointing to the fact that the defendant failed to present any evidence

WL 761958, at *1 (Mass. App. Ct. Mar. 12, 2012) (rejecting the defendant's challenge to the restitution order and noting that "[i]f it is deemed necessary, the probation may be extended to a point where [the defendant] would be able to satisfy his obligations[]").

197 *Norman*, 2011 WL 11345, at *1.

198 *Id.*

199 *Id.*

200 *Id.*

201 *Id.* at *2.

202 *Id.*

203 *Id.*

204 See, e.g., *Commonwealth v. Chase*, 877 N.E.2d 945, 955 (Mass. App. Ct. 2007); *Commonwealth v. Morris M.*, 876 N.E.2d 462, 470 (Mass. App. Ct. 2007).

205 See, e.g., *Chase*, 877 N.E.2d at 955.

206 *Morris M.*, 876 N.E.2d at 469–70.

207 *Id.*

regarding his ability to pay, the Appeals Court also highlighted that “[t]he restitution judge did not order the juvenile to pay the amount forthwith or in installments, either of which would have some bearing on his present ability to pay.”²⁰⁸ The Appeals Court “determine[d] that there was no risk that justice miscarried because the payment of restitution was a condition of the juvenile’s probation, and he may raise his ability to pay at any future probation revocation hearing that should take place based on his nonpayment.”²⁰⁹

Similarly, in *Commonwealth v. Chase*, while upholding a restitution order, the Appeals Court noted that “in the event the defendant faces revocation of his probation for failure to pay the restitution, ‘he may raise his ability to pay at any future probation revocation hearing that should take place based on his nonpayment.’”²¹⁰ On appeal, the defendant had argued that a motion for a new trial as to the restitution order should be granted because “the judge arrived at the amount without a consideration of the defendant’s ability to pay.”²¹¹ According to the Appeals Court, however, since there had been a hearing regarding restitution and the defendant had not raised ability to pay during that hearing, there was no abuse of discretion.²¹² The Appeals Court summarized the proceedings below in the following way:

At the restitution hearing, the judge stated that there were two issues to consider: “I think we need to make a determination first what the restitution is and then we will have to deal with the secondary issue of the ability to pay.” He then added, “But let’s deal with the issue of restitution first and then obviously at some point, you know, if he can’t pay it then some other judge is going to have to make a determination on the ability to pay or what to do about that.”²¹³

The Appeals Court did not criticize the way the lower court judge articulated the role of the defendant’s ability to pay.²¹⁴ All

208 *Id.* (citing *Commonwealth v. Nawn*, 474 N.E.2d 545, 552 (Mass. 1985)).

209 *Morris M.*, 876 N.E.2d at 470.

210 *Chase*, 877 N.E.2d at 955 (quoting *Morris M.*, 876 N.E.2d at 470).

211 *Chase*, 877 N.E.2d at 955.

212 *Id.*

213 *Id.* at 945.

214 *See id.*

of these cases demonstrate that while *Nawn* expressly stated that ability to pay was relevant, the generality of that instruction allowed its impact to be hollowed.

III. A Close Examination: Commonwealth v. Henry

A. The Facts of Henry

In *Henry*, the defendant (Kim Henry) was charged in the Salem Division of the District Court Department with larceny over a value of more than \$250 in violation of Mass. Gen. L. c. 266, § 30(1).²¹⁵ Ms. Henry worked as a cashier at a Walmart store.²¹⁶ It was alleged that Ms. Henry assisted certain customers in taking store items without paying for those items.²¹⁷ Walmart had surveillance camera footage that showed Ms. Henry placing store items into shopping bags without first scanning the items.²¹⁸ The complaint that issued charged Ms. Henry with stealing the property of Walmart pursuant to a “single larcenous scheme” running from July 20 to September 4, 2013.²¹⁹

The criminal case against Ms. Henry was resolved through a plea.²²⁰ In April 2014, Ms. Henry admitted to facts sufficient to warrant a finding of guilt.²²¹ The judge then continued Ms. Henry’s case without a finding of guilt for eighteen months; she was placed on probation for those eighteen months.²²² The judge ordered two conditions of probation: Ms. Henry stay away from Walmart and pay restitution to Walmart.²²³ The judge reserved for a later date decision on the amount of restitution Ms. Henry would be ordered to pay.²²⁴

The restitution hearing was then held in September 2014.²²⁵ At the hearing, Ms. Henry stipulated that Walmart’s losses were \$5,256.10.²²⁶ The judge ordered Ms. Henry to pay restitution in

215 Commonwealth v. Henry, 55 N.E.3d 943, 947-48 (Mass. 2016).

216 *Id.* at 947.

217 *Id.*

218 *Id.*

219 *Id.* at 947-48.

220 *Id.* at 948.

221 *Id.*

222 *Id.*

223 *Id.*

224 *Id.*

225 *Id.*

226 *Id.*

that amount.²²⁷ A month later, however, Ms. Henry filed a motion to revise and revoke the restitution order; the judge allowed the motion.²²⁸ In November 2014, a new restitution hearing was held.²²⁹ At this restitution hearing, Ms. Henry did not agree to the amount of Walmart's loss.²³⁰ The loss protection manager at Walmart testified that the retail sales price of the items stolen amounted to \$5,256.10.²³¹

Ms. Henry, in turn, testified regarding her poverty.²³² She testified that she had been discharged from Walmart after working there nearly twelve years as a cashier.²³³ Since her firing, she had been unable to secure new employment.²³⁴ She was unemployed, had no income and was receiving no government benefits.²³⁵ For the first three months following her discharge from Walmart, she had received unemployment benefits.²³⁶ The Department of Unemployment Assistance, however, ultimately found Ms. Henry ineligible and ordered her to pay back the benefits.²³⁷ She testified that she had been evicted from her apartment.²³⁸

Ms. Henry argued that she should not be ordered to pay any restitution because of her impoverishment.²³⁹ Ms. Henry contended that a restitution order in the amount the Commonwealth sought would inevitably place her in violation of probation.²⁴⁰ The Commonwealth argued that the restitution order should not be reduced on account of Ms. Henry's inability to pay because Ms. Henry had created that inability by stealing and losing her job.²⁴¹ The court resolved that Ms. Henry had to pay \$5,256 in restitution.²⁴² The court ordered the probation department to set the payment

227 *Id.*

228 *Id.*

229 *Id.*

230 *See id.*

231 *Id.*

232 *See id.*

233 *Id.*

234 *See id.*

235 *Id.*

236 *Id.*

237 *Id.*

238 *Id.*

239 *Id.*

240 *Id.*

241 *Id.*

242 *Id.*

schedule.²⁴³ The Supreme Judicial Court accepted the case on direct appellate review to address “whether a defendant’s ability to pay should be considered by a judge in deciding whether to order restitution as a condition of probation and in deciding the amount of any such restitution.”²⁴⁴

B. The Reasoning of Henry

At the outset of the discussion in its unanimous decision, the Supreme Judicial Court stated matter-of-factly that a judge must indeed consider a defendant’s ability to pay in ordering restitution.²⁴⁵ For that proposition, the Court directly quoted from *Nawn*.²⁴⁶ The court resolved the question of whether ability to pay had to be considered in a single paragraph, the first paragraph of its analysis, and with the single citation to *Nawn*.²⁴⁷ The hasty resolution of that question is not surprising. As discussed,²⁴⁸ Massachusetts courts already recognized the general instruction in *Nawn* that ability to pay had to be considered.²⁴⁹

The Supreme Judicial Court declared that two findings must be made at every restitution hearing.²⁵⁰ First, lower courts must decide the amount of economic loss attributable to the defendant’s wrongful conduct.²⁵¹ Second, lower courts must enter a finding regarding the amount the defendant is able to pay.²⁵² The Commonwealth bears the burden as to the amount of economic loss.²⁵³ Defendants bear the burden as to ability to pay.²⁵⁴

The Supreme Judicial Court then turned to clarifying why

243 *Id.*

244 *Id.* at 947, 949.

245 *Id.* at 949.

246 *Id.* (“In deciding whether to order restitution and, if so ordered, the amount, the judge should ‘consider whether the defendant is financially able to pay the amount ordered.’”) (quoting Commonwealth v. *Nawn*, 474 N.E.2d 545, 550 (Mass. 1985)).

247 *Henry*, 55 N.E.3d at 949.

248 See Section II(C)(1).

249 See, e.g., Commonwealth v. Desouza, No. 07-P-1884, 2009 WL 7708, at *2 (Mass. App. Ct. Jan. 2, 2009) (vacating restitution order because the lower court failed to make specific findings as to the defendant’s ability to pay (citing *Nawn*, 474 N.E.2d at 551)).

250 *Henry*, 55 N.E.3d at 949.

251 *Id.* (citing Commonwealth v. McIntyre, 767 N.E.2d 578, 583 (Mass. 2002)).

252 *Henry*, 55 N.E.3d at 949 (citing *Nawn*, 474 N.E.2d at 545).

253 *Henry*, 55 N.E.3d at 949.

254 *Id.* (citing United States v. Fuentes, 107 F.3d 1515, 1532 (11th Cir.1997); Commonwealth v. Porter, 971 N.E.2d 291, 297 (Mass. 2012)).

ability to pay must be considered.²⁵⁵ In this analysis, the court offered bold and broad language regarding the interaction between poverty and criminal punishments. The court marshaled available case law to support this powerful conclusion: “a criminal defendant should not face additional punishment solely because of his or her poverty.”²⁵⁶ The court made reference to a string of constitutionally significant concepts to justify its instruction that lower court judges must consider ability to pay at the outset of determining the restitution order.²⁵⁷ That consideration is essential to the twin constitutional promises that the criminal system provides equal treatment of defendants regardless of their financial status and grants all defendants a fundamentally fair process.

In reaching its conclusion, the court began with this point: a judge must consider a defendant’s ability to pay precisely because restitution is ordered as a condition of probation.²⁵⁸ Since restitution is a part of probation, like all other conditions of probation, “the collection of restitution is enforced by the threat or imposition of a criminal sanction for violation of a probation condition.”²⁵⁹ In support of this line of reasoning, the court cited to *Commonwealth v. Denehy*,²⁶⁰ which states that “[t]he power to order restitution in criminal cases ‘derives from the judge’s power to order conditions of probation’”²⁶¹ and *Commonwealth v. Goodwin*,²⁶² which sets out that “for a defendant, any condition of probation poses the risk that its violation may result in the revocation of probation and the end of his conditional release from imprisonment.”²⁶³

In this way, the Supreme Judicial Court clarified that probation is a criminal disposition, a form of punishment with very real liberty-related consequences. The Supreme Judicial Court placed restitution, as a condition of probation, firmly within this

255 *Henry*, 55 N.E.3d at 950.

256 *Id.* (citing *Bearden v. Georgia*, 461 U.S. 660, 669 n.10 (1983); *Commonwealth v. Canadyan*, 944 N.E.2d 93, 96 (Mass. 2010); *Commonwealth v. Gomes*, 552 N.E.2d 101, 105–06 (Mass. 1990)).

257 *Henry*, 55 N.E.3d at 951–52 (using phrases such as “fundamental fairness,” “equal justice,” “violat[ion of] due process” and “constitutional prerequisite”).

258 *Id.* at 950.

259 *Id.* (citing *Commonwealth v. Denehy*, 2 N.E.3d 161, 174 (Mass. 2014); *Commonwealth v. Goodwin*, 933 N.E.2d 925, 929–30 (Mass. 2010)).

260 *Henry*, 55 N.E.3d at 950.

261 *Denehy*, 2 N.E.3d at 174.

262 *Henry*, 55 N.E.3d at 950 (citing *Goodwin*, 933 N.E.2d at 930).

263 *Goodwin*, 933 N.E.2d at 930.

criminal apparatus. The court then turned to a body of preexisting case law establishing that violations of probation must be based on willful conduct.²⁶⁴ The court used these cases to support its claim that “[a] defendant can be found in violation of a probationary condition only where the violation was wilful, and the failure to make a restitution payment that the probationer is unable to pay is not a wilful violation of probation.”²⁶⁵ More specifically, the court cited to three cases: *Bearden v. Georgia*, *Commonwealth v. Gomes*, and *Commonwealth v. Canadyan*.²⁶⁶ Each of these cases addressed the propriety of imposing punishment on a defendant whose allegedly wrongful conduct resulted from his impoverishment.²⁶⁷ A careful review of *Bearden*, *Gomes*, and *Canadyan* is necessary to understand the nature and scope of the Supreme Judicial Court’s reliance.

In *Bearden v. Georgia*, a 1983 case, the United States Supreme Court held that it is unconstitutional for any state to revoke probation and incarcerate an individual for failure to make court-ordered payments if the defendant has made all reasonable efforts to pay and adequate alternative methods are available for punishing the individual.²⁶⁸ In *Bearden*, the defendant pled guilty to burglary and theft.²⁶⁹ The lower court sentenced the defendant to probation and ordered the defendant to pay a \$500 fine and \$250 in restitution as a condition of probation.²⁷⁰ The defendant made an initial payment; he later alerted the probation office that he was unable to make the final payment on schedule.²⁷¹ When the deadline passed without the defendant paying the sum due, the probation office moved to revoke his probation.²⁷² The defendant argued that imprisoning him for his inability to pay violated the Equal Protection Clause of the Fourteenth Amendment because his impoverishment rendered him unable to make the payment.²⁷³ The state court rejected that

264 *Henry*, 55 N.E.3d at 950.

265 *Id.*

266 *Id.*

267 *Id.* (citing *Bearden v. Georgia*, 461 U.S. 660, 669 n.10 (1983); *Commonwealth v. Canadyan*, 944 N.E.2d 93 (Mass. 2010); *Commonwealth v. Gomes*, 552 N.E.2d 101 (Mass. 1990)).

268 *Bearden*, 461 U.S. at 668–69.

269 *Id.* at 662.

270 *Id.*

271 *Id.* at 662–63.

272 *Id.* at 663.

273 *Id.*

argument and sentenced him to incarceration.²⁷⁴ The Supreme Court reversed.²⁷⁵ The Supreme Court reasoned that “a probationer who has made sufficient bona fide efforts to pay his fine and restitution, and who has complied with the other conditions of probation, has demonstrated a willingness to pay his debt to society and an ability to conform his conduct to social norms.”²⁷⁶ To rule otherwise, the Court explained, “would deprive the probationer of his conditional freedom simply because, through no fault of his own, he cannot pay the fine.”²⁷⁷

The Court agreed with the defendant that “[s]uch a deprivation would be contrary to the fundamental fairness required by the Fourteenth Amendment.”²⁷⁸ Thus, incarcerating a defendant for failure to pay despite his bona fide efforts to do so and where there exist reasonable adequate alternatives to incarceration is barred by the equal protection principles of the federal constitution.²⁷⁹ Notably, the Court compared those equal protection concerns to the potential due process concerns.²⁸⁰ The Court described the equal protection concerns with incarcerating a poor defendant for inability to pay as “substantially similar to asking directly the due process question of whether and when it is fundamentally unfair or arbitrary for the State to revoke probation when an indigent is unable to pay the fine.”²⁸¹ The Court stated that “[d]ue process and equal protection principles converge in the Court’s analysis in these cases.”²⁸² Thus, while the Court did not declare that the incarceration of a defendant for inability to pay where there existed reasonable alternatives to incarceration violated the due process clause of the Fifth Amendment, its language suggested that it was open to the possibility that the due process guarantees of the federal constitution are also strained by the practice of incarcerating defendants for their inability to pay.²⁸³ Commentators nationwide have treated *Bearden* as a seminal case regarding debt-based incarceration.²⁸⁴

274 *Id.*

275 *Id.* at 674.

276 *Id.* at 670.

277 *Id.* at 672-73.

278 *Id.* at 673.

279 *Id.*

280 *Id.*

281 *Id.* at 666.

282 *Id.* at 665.

283 *See id.*

284 Ann Wagner, *The Conflict over Bearden v Georgia in State Court: Plea-Bargaining*

In 1990, in *Commonwealth v. Gomes*, the Supreme Judicial Court relied extensively on federal case law to rule that incarcerating an indigent defendant based solely on his inability to pay court fees is unconstitutional.²⁸⁵ The lower court judge in *Gomes* jailed the defendant for his failure to pay a fee the defendant had been charged for missing a court date.²⁸⁶ The underlying offense: possessing an open liquor container in public.²⁸⁷ The lower court judge made no determination regarding whether the defendant was financially able to pay the fee before jailing him.²⁸⁸ On appeal, the defendant argued that because he was indigent, the court could not incarcerate him for missed payment.²⁸⁹ The Supreme Judicial Court agreed that

Probation Terms and the Spector of Debtors Prisons, 2010 U. CHI. LEGAL F. 383, 385 (2010), <http://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1466&context=uclf> (“By requiring a finding of willfulness, *Bearden* prevented a revival of debtors’ prison for criminal defendants. No longer would mere inability to pay a debt to the state be sufficient grounds for imprisonment.”); Colin Reingold, *Pretextual Sanctions, Contempt, and the Practical Limits of Bearden-Based Debtors’ Prison Litigation*, 21 MICH. J. RACE & L. 361, 362, 374 (2016) (explaining that “[t]oday, *Bearden* is invoked in courtrooms throughout America to protest when judges attempt to jail a defendant for reasons that directly or indirectly stem from poverty” and “*Bearden*’s importance to the quest for justice cannot be overstated.”); Sally Yates, *Poverty is Not a Crime*, THE UNITED STATES DEPT OF JUST. ARCHIVES (Dec. 2, 2015), <https://www.justice.gov/archives/opa/blog/poverty-not-crime> (last visited Dec. 27, 2017) (“Again, in 1983, in a case called *Bearden v. Georgia*, the Supreme Court reaffirmed that the Constitution does not permit “punishing a person for his poverty.”); Michael Williams, *How to End Debtors Prisons in America*, HUFFPOST (Aug. 2, 2016, 5:39 PM), http://www.huffingtonpost.com/entry/how-to-end-debtors-prisons-in-america_us_57a0c201e4b004301c51e6bf; Whitney Bents & Blake Strode, *Debtors’ Prisons in 21st-Century America*, THE ATLANTIC (Feb. 23, 2016), <https://www.theatlantic.com/business/archive/2016/02/debtors-prison/462378/>; Whitney Scherck, *Battering the Poor: How Georgia’s Mandatory Family Violence Classes Deny Indigent Defendants Equal Protection of the Law*, 5 WIDENER J. L. ECON. & RACE 169, 170 (2014), <http://blogs.law.widener.edu/wjler/files/2014/06/7Scherck.pdf>; Note, *State Bans on Debtors’ Prisons and Criminal Justice Debt*, 129 HARV. L. REV. 1024, 1026 (2016), <https://harvardlawreview.org/2016/02/state-bans-on-debtors-prisons-and-criminal-justice-debt/>.

285 *Commonwealth v. Gomes*, 552 N.E.2d 101, 105 (Mass. 1990) (citing *Santiago v. United States*, 889 F.2d 371, 373 (1st Cir. 1989); *United States v. Grose*, 687 F.2d 1298, 1301 (10th Cir. 1982); *United States ex rel. Meyer v. Weil*, 458 F.2d 1068, 1070 (7th Cir. 1972), *cert. denied*, 409 U.S. 1060 (1972)).

286 *Gomes*, 552 N.E.2d at 102-03.

287 *Id.* at 102.

288 *Id.* 103.

289 *Id.* at 105.

such practices were constitutionally intolerable.²⁹⁰ The court cited to multiple federal cases and quoted the United States Supreme Court's declaration that "the Constitution prohibits the State from imposing a fine as a sentence and then automatically converting it into a jail term solely because the defendant is indigent and cannot forthwith pay the fine in full."²⁹¹ By contrast, the incarceration of a defendant who willfully "refuses" or "neglects" to pay "pose[s]" "[n]o constitutional difficulty."²⁹² Thus, for a defendant to be incarcerated based on his failure to pay, there must first be a hearing at which there is a finding that the failure to pay was willful and the defendant must have been offered less restrictive alternatives.²⁹³ Only where those requirements are satisfied is incarceration constitutional, because it then constitutes "punishment, not coercion."²⁹⁴

Commonwealth v. Canadyan, a 2010 Supreme Judicial Court case, reapplied the willfulness requirement to probation violations.²⁹⁵ *Canadyan* provides support for the proposition that where the defendant's impoverishment renders him incapable of satisfying a probation condition, that noncompliance is not willful and therefore cannot constitute a violation of probation.²⁹⁶ The defendant in *Canadyan* was ordered to maintain a global positioning system monitoring device ("GPS device"); the probation officer moved to violate him when the GPS device was not maintained.²⁹⁷ The defendant, however, presented undisputed evidence that he resided at a homeless shelter and the shelter could not accommodate the technology to maintain the GPS device.²⁹⁸ On appeal, the Supreme Judicial Court concluded that the defendant's failure to comply with the GPS condition was "through no fault of his own."²⁹⁹ The court recognized "a tension between [the probation condition at issue] and the practical reality of homelessness."³⁰⁰ Given that

290 *Id.*

291 *Id.* (quoting *Tate v. Short*, 401 U.S. 395, 398 (1971)).

292 *Gomes*, 552 N.E.2d at 105.

293 *Id.* at 105–06 (citing *Bearden v. Georgia*, 461 U.S. 660, 672 (1983)).

294 *Gomes*, 552 N.E.2d at 105.

295 *Commonwealth v. Canadyan*, 944 N.E.2d 93, 96 (Mass. 2010).

296 *Id.* at 93–96.

297 *Id.* at 94–95

298 *Id.* at 95.

299 *Id.* at 96.

300 *Id.* at 95.

there was “no evidence of wilful noncompliance,”³⁰¹ punishing the defendant for the inoperability of the GPS amounted to punishing the defendant for being homeless—an outcome the Court implied constituted a violation of “basic fairness.”³⁰² Although the court did not explicitly constitutionalize its holding, it quoted *Bearden*’s declaration that “basic fairness forbids the revocation of probation when the probationer is without fault in his failure to [comply].”³⁰³

In *Henry*, the Supreme Judicial Court combined the principles regarding impoverishment-based punishment from *Bearden*, *Gomes*, and *Canadyan* and leveraged these principles in the restitution context. Relying on these cases, the *Henry* court was able to firmly conclude that where the defendant does not have the ability to make a restitution payment the defendant’s failure to pay is not willful.³⁰⁴ And where the missed payment is not willful, it cannot constitute a probation violation.³⁰⁵ To hold otherwise would violate constitutional norms.

The *Henry* court could have ended its inquiry there, having resolved the question immediately before it of whether ability to pay was relevant and having provided poor defendants a robust defense if they were violated for failure to pay restitution. But the court took these principles one step further. The court took an expansive, forward-looking view of the principle that a defendant cannot be violated for conduct that is not willful. In the court’s view, the prohibition on punishing non-willful conduct also compelled the conclusion that the judge’s decision regarding the defendant’s ability to pay must be entered at the original restitution hearing.³⁰⁶ The court “require[d] the judge to consider the defendant’s ability to pay when initially setting the restitution amount.”³⁰⁷ The court barred judges from delaying consideration of ability to pay until the time at which the defendant missed a payment.³⁰⁸ Any delay in considering the defendant’s ability to pay gave rise to the “unlawful result” in which the conditions of probation “simply doom[] the defendant to noncompliance.”³⁰⁹ Although the court did not expressly invoke

301 *Id.* at 96.

302 *Id.* (citing *Bearden v. Georgia*, 461 U.S. 660, 669 n.10 (1983)).

303 *Id.*

304 *Commonwealth v. Henry*, 55 N.E.3d 943, 950 (Mass. 2016).

305 *Id.*

306 *Id.*

307 *Id.*

308 *Id.* at 951.

309 *Id.* at 950.

either the Fifth or Fourteenth Amendment of the federal constitution, the court quoted the Supreme Court of Iowa's statement that "[a] court's assessment of a defendant's reasonable ability to pay is a constitutional prerequisite for a criminal restitution order."³¹⁰

Where probation is doomed to fail, the Supreme Judicial Court explained, the consequences for the indigent defendant are tangible and constitutionally intolerable.³¹¹ Allowing a judge to issue a restitution order that a defendant has no ability to pay invites upon the "blameless" defendant a host of punishments.³¹² A defendant who is accused of failing to pay restitution faces a series of criminal sanctions: the risk of arrest on a probation warrant, the incurred cost of a warrant fee and possibility of jail if the judge fails to recognize the prohibition on jailing a defendant for inability to pay.³¹³ Moreover, if the judge failed to realize that ability to pay is a defense to a probation violation, there is the possibility that the defendant would be jailed for the time between the hearing at which probation first moves to violate the defendant and the separate hearing at which a judge determines if there had in fact been a violation. There is also the possibility of alterations to the defendant's probation status, including the addition of new, more restrictive conditions of probation, an extension of probation or revocation of probation followed by incarceration.³¹⁴ The imposition of such sanctions is impermissible, the court reasoned, based on the principles outlined in *Bearden*, *Gomes*, and *Canadyan*.³¹⁵ In addition to noting that these sanctions waste lower courts' resources,³¹⁶ the court declared that "[b]urdening a defendant with these risks by imposing restitution that the defendant will be unable to pay violates the fundamental principle that a criminal defendant should not face additional punishment solely because of his or her poverty."³¹⁷ That phrasing goes beyond the holdings of *Bearden*³¹⁸ and *Gomes*,³¹⁹ which constrained judges only in incarcerating; *Henry* is substantially more

310 *Id.* at 950-51 (citing *State v. Blank*, 570 N.W.2d 924, 927 (Iowa 1997)).

311 *Henry*, 55 N.E.3d at 950.

312 *Id.*

313 *Id.*

314 *Id.* at 951 (citing *Commonwealth v. Goodwin*, 933 N.E.2d 925, 930 (Mass. 2010)).

315 *Henry*, 55 N.E.3d at 950.

316 *Id.*

317 *Id.*

318 *Bearden v. Georgia*, 461 U.S. 660, 672 (1983).

319 *Commonwealth v. Gomes*, 552 N.E.2d 101, 105 (Mass. 1990).

protective as it applies to punishment and the threat of punishment more broadly.

The court read the principle against punishing a defendant more severely based on his impoverishment as barring the practice of placing poor defendants on probation for longer periods for the sole purpose of collecting more restitution.³²⁰ Thus, courts can neither initially determine the length of probation based on a defendant's limited financial resources nor extend probation in response to a defendant's limited resources.³²¹ To justify this conclusion, the court returned to the proposition that probation "is not a civil program or sanction."³²² Instead, "[p]robation 'serves as a disposition of and punishment for a *crime*'"³²³

The Supreme Judicial Court walked through the various ways in which probation is a punishment.³²⁴ First, a defendant on probation suffers severe restrictions to his liberty during the duration of the probation.³²⁵ While on probation, defendants may be ordered to check in regularly with their probation officers, submit to drug testing, remain in the state unless permission from the court or probation has been obtained, and comply with whatever other restrictions and obligations the judge imposes.³²⁶ Home confinement, GPS monitoring, and curfew hours may be ordered as conditions of probation.³²⁷ Probationers can also be ordered to participate in various programs.³²⁸ Defendants are subject to monthly probation fees.³²⁹

Second, an individual on probation remains in a state of diluted constitutional rights for the duration of his probation.³³⁰ While on probation, a defendant's home can be searched by a probation officer so long as there is reasonable suspicion of a crime (rather than the more rigorous requirement of probable clause that is applicable to home searches of individuals who are not on

320 *Henry*, 55 N.E.3d at 950-52.

321 *Id.* at 952.

322 *Id.* at 951 (quoting *Commonwealth v. Cory*, 911 N.E.2d 187 (Mass. 2009)).

323 *Henry*, 55 N.E.3d at 951 (emphasis in original).

324 *Id.*

325 *Id.*

326 *Id.*

327 *Id.*

328 *Id.*

329 *Id.*

330 *See id.*

probation).³³¹ In addition, the defendant retains fewer rights if accused of a new crime.³³² Whereas a crime must typically be proven beyond a reasonable doubt before punishment may be imposed, an individual on probation may face jail time so long as the new crime is proven by the much less rigorous standard of preponderance of the evidence.³³³ Based on an alleged new offense, the defendant may have his probation revoked and be sentenced to incarceration on the underlying crime for which he was on probation.³³⁴ Because of the lesser standard of proof in the probation context, the defendant may face incarceration based on a judge's conclusion that a preponderance of the evidence shows that the defendant committed the new offense.³³⁵ This can occur even where the defendant is acquitted in the separate proceedings adjudicating his guilt on the newly alleged offense.³³⁶ In other words, a defendant on probation can be found not guilty of a newly charged crime yet still serve jail time triggered by that new charge.³³⁷

Thus, the court concluded that "extending the length of a probationary period because of a probationer's inability to pay subjects the probationer to additional punishment solely because of his or her poverty."³³⁸ The various liberty restrictions, dilution of constitutional rights, and financial costs probation imposes render

331 *Id.* (citing Commonwealth v. LaFrance, 525 N.E.2d 379, 381 (Mass. 1988)).

332 *Henry*, 55 N.E.3d at 951.

333 *Id.* (citing Commonwealth v. Hartfield, 51 N.E.3d 465, 473–74 (Mass. 2016)).

334 *Henry*, 55 N.E.3d at 951.

335 Commonwealth v. Wilcox, 841 N.E.2d 1240, 1246 (Mass. 2006) ("The probationer escaped total loss of liberty only as a result of the trial judge's exercise of discretion. It follows that, at a revocation proceeding, 'a probationer need not be provided with the full panoply of constitutional protections applicable at a criminal trial.' The finding of a violation is not by a jury but by a judge, and is based only on a preponderance of the evidence, not proof beyond a reasonable doubt.") (internal citations omitted) (citing Commonwealth v. Durling, 551 N.E.2d 1193, 1197 (Mass. 1990)).

336 *Wilcox*, 841 N.E.2d at 1246 (citing Commonwealth v. Holmgren, 656 N.E.2d 577, 578-79 (Mass. 1995)).

337 *Holmgren*, 656 N.E.2d at 578 ("Principles of collateral estoppel do not bar the Commonwealth from revoking probation based on evidence of a violation of law of which a probationer has been found not guilty. The reason for this result lies in the difference in the burden of proof in the two proceedings. In a criminal case, of course, the Commonwealth must prove the elements of each crime charged beyond a reasonable doubt. In a probation revocation hearing the Commonwealth bears a lesser burden.").

338 *Henry*, 55 N.E.3d at 952.

probation itself a punishment.³³⁹ Forcing an indigent defendant to remain on probation for longer than necessary to rehabilitate the defendant and protect the public is a practice the court unequivocally rejected.³⁴⁰ The court expressly declined to constitutionalize this prohibition.³⁴¹ Instead of resolving whether extending probation based on a defendant's inability to pay violates the Massachusetts Declaration of Rights, the court invoked its superintendence power to prohibit lower court judges from extending an individual's probation due to inability to pay.³⁴²

The court's prohibition on extending probation as a response to a defendant's inability to pay is significant because the United States Supreme Court has resolved that granting indigent defendants extensions of probation to satisfy court-ordered payments is consistent with the federal constitution.³⁴³ In *Bearden*, the Supreme Court invited judges to extend probation to allow indigent defendants more time to make their payments: "'[t]he State is not powerless to enforce judgments against those financially unable to pay a fine.' For example, the sentencing court could extend the time for making payments . . .'"³⁴⁴ In *Gomes*, the Supreme Judicial Court had embraced that viewpoint.³⁴⁵ The *Henry* court recognized its departure: "We acknowledge that extending the length of probation in such circumstances has not been recognized to be in violation of Federal constitutional law."³⁴⁶

As mentioned, the Supreme Judicial Court also prohibited lower court judges from taking into account a defendant's ability to pay when initially determining the length of probation.³⁴⁷ The court stated that "equal justice" means that the length of probation cannot be affected by a defendant's ability to pay.³⁴⁸ The court endorsed this language from the Superior Court Working Group on Best Practices: "[a]n extended period of supervision for the purpose of collecting money can be particularly troublesome since

339 *Id.* at 951.

340 *Id.* at 952.

341 *Id.*

342 *Id.*

343 *Bearden v. Georgia*, 461 U.S. 660, 674 (1983).

344 *Id.* at 672 (quoting *Tate v. Short*, 401 U.S. 395, 399 (1971)).

345 *Commonwealth v. Gomes*, 552 N.E.2d 101, 106 (1990) (citing *Bearden*, 461 U.S. at 672).

346 *Henry*, 55 N.E.3d at 952 n.7 (citing *Bearden*, 461 U.S. at 674).

347 *Henry*, 55 N.E.3d at 952 n.7.

348 *Id.*

it necessarily means that greater burdens are imposed on poor offenders³⁴⁹ Extending probation would mean that courts are permitted to impose the substantial burdens tied to being on probation based on a defendant's financial status rather than what is necessary to achieve the legitimate goals of the criminal justice system.³⁵⁰ A poor defendant convicted of the very same crime as a wealthy defendant would be entangled with the criminal courts longer simply because of his impoverishment.³⁵¹ For this point, the court relied on the following reasoning from the Supreme Court of Montana: to the extent that a longer suspended sentence was imposed on the defendant due to his indigent status and in an effort to collect restitution, the defendant's rights to due process and fundamental fairness were violated.³⁵² Thus, though it did not explicitly invoke the Fifth or Fourteenth Amendments of the federal constitution, the Supreme Judicial Court in *Henry* relied on case law invoking both and employed language that referenced both equality ("equal justice") and fairness ("fundamental fairness").³⁵³ Thus, the *Henry* court appeared at a minimum aware of the constitutional undertones of the issue and willing to rely on constitutional concepts in its analysis.

"To ensure that a defendant does not face a longer probationary period because of his or her limited means,"³⁵⁴ the court provided very clear instructions. Judges must first determine the length of probation compelled by the legitimate penal goals of protecting the public and rehabilitating the defendant.³⁵⁵ Only after judges decide the appropriate length of probation warranted for the particular defendant are judges then permitted to turn to restitution.³⁵⁶ Judges must then determine the defendant's ability to pay and judges may

349 *Id.* (citing SUPERIOR COURT WORKING GROUP ON SENTENCING BEST PRACTICES, CRIMINAL SENTENCING IN THE SUPERIOR COURT: BEST PRACTICES FOR INDIVIDUALIZED EVIDENCE-BASED SENTENCING 15 (Mar. 2016)).

350 For example, in *Commonwealth v. Malick*, the court described "the four fundamental purposes of sentencing: incapacitation, deterrence, retribution, and rehabilitation." *Commonwealth v. Malick*, 14 N.E.3d 338, 344 (Mass. App. Ct. 2014) (citing *Commonwealth v. McIntyre*, 767 N.E.2d 578, 582 (Mass. 2002)).

351 *Henry*, 55 N.E.3d at 952.

352 *Id.* (citing *State v. Farrell*, 676 P.2d 168, 176–77 (Mont. 1984)).

353 *Henry*, 55 N.E.3d at 952.

354 *Id.*

355 *Id.* (citing *Commonwealth v. Cory*, 911 N.E.2d 187, 194 (Mass. 2009)).

356 *Henry*, 55 N.E.3d at 952.

order the defendant to pay a monthly amount of restitution that does not exceed the defendant's ability to pay.³⁵⁷ Judges cannot delegate determination of the monthly amount of restitution to the probation department.³⁵⁸ If there are any material changes to a defendant's ability to pay during the probation period, the judge may be asked to increase or decrease the restitution payments.³⁵⁹

The court recognized that restricting restitution to what the defendant is able to pay might lead to victims receiving less than full compensation.³⁶⁰ Where there is a difference between the sum of restitution ordered and the victim's losses, the victim may seek recovery through a civil action.³⁶¹ Once the defendant has completed his probation period, he is entitled to be released from the grip of the criminal courts.³⁶² At that point, the criminal court's intervention in the defendant's life is not justified. It must end. What remains is a debt. The victim must turn to the civil court system if he or she seeks further compensation.³⁶³

The court seized the opportunity to set the legal standard for measuring ability to pay.³⁶⁴ The court acknowledged that it was articulating this standard for the first time.³⁶⁵ The court declared that meaningful consideration of ability to pay means that restitution cannot impose "substantial financial hardship."³⁶⁶ Lower court judges must consider not only the defendant's financial resources but also the defendant's financial obligations.³⁶⁷ A defendant's financial resources include his income and net assets.³⁶⁸ A defendant's financial obligations include "the amount necessary to meet minimum basic human needs such as food, shelter, and clothing for the defendant and his or her dependents."³⁶⁹ From the defendant's resources and

357 *Id.* Moreover, the entirety of the restitution ordered over the course of the probation cannot exceed the victim's actual loss. *Id.*

358 *Id.*

359 *Id.* at 953.

360 *See id.* at 950 n.4.

361 *Id.*

362 *See id.* at 952 n.7.

363 *Id.*

364 *Id.* at 953.

365 *Id.*

366 *Id.* at 954.

367 *Id.* at 953.

368 *Id.* Note that "a judge may attribute potential income to the defendant only after specifically finding that the defendant is earning less than he or she could through reasonable effort." *Id.* at 954.

369 *Id.* at 953–54 (citing MASS. GEN. LAWS ch. 261, § 27A (2000); United States

obligations, judges must determine the sum the defendant has available to pay restitution.³⁷⁰ The court stated that “[r]estitution payments that would deprive the defendant or his or her dependents of minimum basic human needs would cause substantial financial hardship.”³⁷¹ The court openly cautioned that “[w]here a defendant has been found indigent by the court for purposes of the appointment of counsel, a judge should consider carefully whether restitution can be ordered without causing substantial financial hardship.”³⁷² Although the court did not expressly acknowledge it, it would seem to follow from its reasoning that defendants whose sole income comes from government benefits that are set specifically to meet the defendant and his dependents’ basic needs are not resources available for restitution.

Finally, the Supreme Judicial Court applied these newly announced principles to Ms. Henry’s case.³⁷³ Where the Supreme Judicial Court was declaring that judges could not order restitution without considering ability to pay, the lower court judge had ordered restitution in the full amount of the Walmart’s losses, while refusing to consider Ms. Henry’s ability to pay.³⁷⁴ The lower court judge ordered the full amount of losses even though Ms. Henry presented apparently uncontested evidence of her impoverishment.³⁷⁵ The Supreme Judicial Court noted that the lower court judge seemed to have recognized Ms. Henry’s inability to pay when he said at the hearing that “you can’t get blood out of a stone” and that he was “not feeling great about this.”³⁷⁶ Moreover, where the Supreme Judicial Court was ruling that judges cannot delegate consideration of a defendant’s ability to pay to probation, the lower court initially delegated to probation the decision regarding the amount of the monthly restitution payments.³⁷⁷ Six months after Ms. Henry’s restitution order was imposed, a probation violation notice issued charging that Ms. Henry had failed to pay.³⁷⁸ Ms. Henry did not

v. McGiffen, 267 F.3d 581, 589 (7th Cir.2001); Museitef v. United States, 131 F.3d 714, 716 (8th Cir.1997)).

370 *Henry*, 55 N.E.3d at 953.

371 *Id.* at 954.

372 *Id.*

373 *Id.*

374 *Id.*

375 *Id.* at 948.

376 *Id.* at 954 n.8.

377 *Id.*

378 *Id.* at 955.

appear at the hearing set to resolve the violation notice.³⁷⁹ As such, a default was entered and a warrant was issued for her arrest – one of the very risks the Supreme Judicial Court highlighted in its analysis.³⁸⁰ Ms. Henry’s warrant was recalled two weeks later.³⁸¹ After Ms. Henry ultimately stipulated to the probation violation, the judge placed her back on probation and specifically ordered that she pay \$30 of restitution monthly moving forward.³⁸² Ms. Henry made the \$30 monthly payments.³⁸³ Yet, on the day her probation was set to expire, probation issued another violation notice charging that a balance of \$5,176 remained unpaid.³⁸⁴ There was no claim that she had violated any other condition of her probation, been charged with a new crime or otherwise failed to conform to the social standards probation is designed to promote.³⁸⁵

The Supreme Judicial Court concluded that this series of events constituted “prolonged punishment” and occurred “only because of her poverty.”³⁸⁶ As the court put it, “[t]he consequence of [the lower court’s] errors demonstrates why it is so important that ability to pay be considered in setting the amount of restitution.”³⁸⁷ Ms. Henry was subjected to a restitution order that “doom[ed her] to noncompliance,”³⁸⁸ a probation violation notice that would not have been issued but for her poverty and an arrest warrant that would not have issued but for her poverty.³⁸⁹ Had the hearing on her final probation notice not been continued pending the appeal to the Supreme Judicial Court, Ms. Henry would have faced the very real possibility of her probation being formally extended. And with it, the possibility of her constitutional rights remaining diluted for the sole purpose of collecting payments.

Still more, as to those payments Ms. Henry made under this restitution order that did not consider her ability to pay, there is the significant likelihood that those payments caused her the very kind of “substantial financial hardship” the Supreme Judicial Court was

379 *Id.*

380 *Id.* at 950, 955.

381 *Id.* at 955.

382 *Id.*

383 *Id.*

384 *Id.*

385 *See id.*

386 *Id.*

387 *Id.* at 954.

388 *Id.* at 950.

389 *Id.* at 955.

prohibiting in *Henry*.³⁹⁰ Based on the evidence presented, Ms. Henry had no income.³⁹¹ She was unemployed and receiving no government benefits.³⁹² She was not only ineligible to receive unemployment benefits but also indebted to the Department of Unemployment Assistance.³⁹³ She had been evicted and was effectively without stable housing.³⁹⁴

It is likely that Ms. Henry's ability to pay was zero under the standard the Supreme Judicial Court articulated. Indeed, it is plausible that the monies Ms. Henry paid towards restitution under the order issued without consideration of her ability to pay denied her the opportunity to "retain[] sufficient means for reasonable living expenses and existing family obligations."³⁹⁵ Possibly, those payments—issued under the threat of criminal sanction—deprived Ms. Henry of "minimum basic human needs such as food, shelter, and clothing."³⁹⁶ The facts detailed in the court's decision do not reveal whether Ms. Henry's eviction resulted in any way from missed rent payments; it remains unclear whether she was faced with choosing between being evicted if she failed to pay rent and facing criminal sanctions if she failed to pay restitution. In all of these ways, Ms. Henry's financial reality likely required "a judge . . . [to] consider carefully whether restitution [could] be ordered without causing substantial financial hardship."³⁹⁷ In the end, Ms. Henry faced prolonged punishment due to her impoverishment.³⁹⁸

IV. The Impact of Henry

The impact of *Henry* is profound.³⁹⁹ It is best examined in two parts: (1) the benefits *Henry* offers to both critics and proponents of

390 *Id.* at 954.

391 *Id.* at 948.

392 *Id.*

393 *Id.*

394 *Id.*

395 *Id.* at 954 (citing MODEL PENAL CODE § 6.04(2) (AM. LAW INST., Proposed Official Draft 2012)).

396 *Henry*, 55 N.E.3d at 953.

397 *Id.* at 954.

398 *Id.*

399 Lower courts have already applied *Henry* and overturned restitution orders that did not rely on an explicit finding of ability to pay. See, e.g., Commonwealth v. Amendola, No. 16-P-1073, 2017 WL 2261067, at *1-2 (Mass. App. Ct. May 23, 2017) (remanding case because "in accordance with the recent guidance provided in *Comm. v. Henry*, 421 Mass. at 121, the judge was also required to determine the defendant's ability to pay").

restitution and (2) the potential impact of *Henry* on other criminal court-ordered payments.

A. *Henry's Impact on the Policy Underpinnings of Restitution*

Henry provides reason for both critics and supporters of restitution to celebrate. For those who criticize the practice of imposing restitution as a condition of probation, one of the criticisms of restitution is that it unfairly punishes poor defendants.⁴⁰⁰ Given the protections *Henry* offers poor defendants against restitution orders that would impose substantial financial hardship, incarceration for missed payments and unduly long terms of probation,⁴⁰¹ critics of restitution have obvious reason to embrace *Henry*. To the extent that lower court judges adhere to it, *Henry* constitutes a significant step in the progressive direction. Criminal justice organizations in Massachusetts have celebrated the *Henry* decision.⁴⁰² The American Civil Liberties Union proclaimed:

Ordering a defendant to pay restitution she cannot possibly afford is not only unfair to the defendant, but also useless to the victim and costly to the Commonwealth. The Court's decision [in *Henry*] is a victory for everyone, as it ensures that judges will craft restitution orders that defendants can comply with and that the Commonwealth can enforce.⁴⁰³

Proponents of restitution also have reason to embrace *Henry*.⁴⁰⁴

400 Anderson, *supra* note 61, at 326.

401 See *Henry*, 55 N.E.3d at 954–955.

402 For example, the American Civil Liberties Union has described *Henry* as a “victory.” *State supreme court requires judges to consider defendants’ ability to pay restitution*, ACLU MASSACHUSETTS (Aug. 8, 2016), <https://aclum.org/uncategorized/state-supreme-court-requires-judges-consider-defendants-ability-pay-restitution/>. See also Letter from Carol S. Steiker et al., Harvard Law School, to Maura S. Doyle, Clerk, Suffolk Cty. Sup. Jud. Ct. (Aug. 15, 2016) (on file with Harvard Law School), <http://cjjp.law.harvard.edu/assets/SJC-Final-Letter.pdf> (two professors from Harvard Law’s Criminal Justice Policy Program relying on the *Henry* decision in an advocacy letter).

403 *State supreme court requires judges to consider defendants’ ability to pay restitution*, ACLU MASSACHUSETTS (Aug. 8, 2016), <https://aclum.org/uncategorized/state-supreme-court-requires-judges-consider-defendants-ability-pay-restitution/>.

404 That is not to say that they have embraced the decision. For instance, the Essex District Attorney Jonathan Blodgett made a statement describing the

The requirement that restitution orders must reflect an amount that the defendant can pay without experiencing substantial financial hardship places restitution on more sound constitutional ground. As discussed, critics of monetary sanctions have forcefully argued that the imposition of monetary sanctions raises due process⁴⁰⁵ and equal protection concerns.⁴⁰⁶ Even proponents of restitution have at times acknowledged that the proper administration of restitution should include consideration of a defendant's ability to pay.⁴⁰⁷

There is also reason to expect more practical benefits to restitution collection in the post-*Henry* format, as research suggests that setting restitution to ability to pay may actually increase collection rates.⁴⁰⁸ If restitution is to satisfy its purpose of compensating victims,

Henry decision as a “very difficult pill to swallow.” He further stated that the decision “really put[s] these victims at a disadvantage.” Julie Manganis, *DA Slams SJC Ruling on Restitution*, SALEM NEWS (Aug. 11, 2016), http://www.salemnews.com/news/local_news/da-slams-sjc-ruling-on-restitution/article_16e93871-56cc-53c8-a5fd-8906cd34cb8a.html.

- 405 See Charles Pengilly, *Restitution, Retribution and the Constitution*, 7 ALASKA L. REV. 333, 346 (1990), <http://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1271&context=alr> (explaining that “a restitution order involves a deprivation of property – which must be preceded by ‘the opportunity to be heard at a meaningful time and in a meaningful manner’”); Anderson, *supra* note 61, at 364 (noting that “[w]here jurisdictions do not require a pre-imposition determination of ability to pay, or notice and opportunity to be heard before the debt is imposed, basic due process is violated”).
- 406 Anderson, *supra* note 61, at 366 (explaining that “equal protection is violated when an indigent, partially indigent, or even formerly indigent defendant is threatened with imprisonment for nonpayment, while a non-indigent who refuses to pay retained counsel can never be sent to prison for failure to pay the civil debt”).
- 407 See Mank, *supra* note 66, at 73 (explaining that “[i]f sentencing judges consider a defendant's ability to pay when they make restitution orders that alone may be enough to defeat any Seventh Amendment challenge” and “as long as they consider the defendant's ability to pay, sentencing judges may be able to award any type of damages”). See also Frances L. Pierson, *Victim Compensation through Sentencing*, COLONIAL LAWYER, Winter 1984–85, at 30 (embracing restitution and explaining that “[t]o be effective, the assigned restitution should be clearly defined, measurable and achievable . . . [a]n achievable restitution figure can be set if the court considers the offender's ability to pay”).
- 408 Matthew Dickman, *Should Crime Pay?: A Critical Assessment of the Mandatory Victims Restitution Act of 1996*, 97 CAL. L. REV. 1687, 1695 (2009) (citing R. Barry Ruback et. al., *Evaluation of Best Practices in Restitution and Victim Compensation Orders and Payments*, PA. COMM'N ON SENTENCING (Apr. 2006)), <http://pcs.la.psu.edu/publications-and-research/research-and-evaluation-reports/special-reports/act-37-of-2007-economic-sanctions/special-report->

payments must actually be collected—not just imposed. Collection rates are a well-documented problem across jurisdictions.⁴⁰⁹ Legal scholars studying the federal courts' collection rates have found that "the foremost reason for the nonpayment of criminal economic sanctions is the offender's inability to pay them."⁴¹⁰ Defendants who are ordered to pay restitution amounts that exceed their ability to pay are often demoralized and left with less incentive to make any payments.⁴¹¹ In 2006, criminal defense attorney James Felman testified to the United States House of Representatives Judiciary Committee that "[a]mong the costs of ordering a defendant to pay what everyone recognizes he or she cannot is that there is little incentive for the defendant to try."⁴¹² Excessive restitution orders risk striking defendants as unfair.⁴¹³ Defendants' perceptions that their sentences are unfair⁴¹⁴ increase noncompliance and undermine rehabilitation.⁴¹⁵

There is also reason to believe that restitution orders that are

evaluation-of-best-practices-in-restitution-and-victim-compensation-orders-and-payments/view; R. Barry Ruback, *The Imposition of Economic Sanctions in Philadelphia: Costs, Fines, and Restitution*, 68 FED. PROB. 1, 21, 25 (2004) [hereinafter Ruback, *Economic Sanctions*]).

409 Ruback, *Benefits and Costs*, *supra* note 63, at 1780 (noting that "simply because these economic sanctions are imposed does not mean that they are collected . . . in fact enormous amounts of the monies that are owed are unpaid").

410 *Id.* at 1781.

411 See *id.* at 1805 (explaining that "if the offender perceives the economic sanctions as so high as to be impossible to pay, he or she is likely to perceive the system as unfair and illegitimate, and, as a result, to simply give up attempting to pay"). See also Anderson, *supra* note 61, at 341 (considering the practice of requiring poor defendants to pay towards the cost of their court-appointed lawyer and arguing that "the automatic imposition of the repayment obligation on those who will never be able to repay does not serve a legitimate state interest and demoralizes defendants").

412 Dickman, *supra* note 408, at 1696 (citing *Criminal Restitution Improvement Act of 2006: Hearing on H.R. 5673 Before the H. Subcomm. on Crime, Terrorism, and Homeland Security*, 109th Cong. 1 28-29 (2006) (testimony of James Felman, Partner, Kynes, Markman & Felman)).

413 Ruback, *Benefits and Costs*, *supra* note 63, at 1805. See also Ruback, *Economic Sanctions*, *supra* note 408, at 21, 25.

414 Ruback, *Economic Sanctions*, *supra* note 408, at 21, 25.

415 William M. Acker, Jr., *The Mandatory Restitution Act Is Unconstitutional. Will the Courts Say So After Southern Union v. United States*, 64 ALA. L. REV. 803, 834 (2013), <https://www.law.ua.edu/pubs/lrarticles/Volume%2064/Issue%204/2%20Acker%20803-843.pdf> ("When a defendant believes, often with good cause, that the restitution portion of his sentence is illegal or fundamentally unfair, and learns, as a practical matter, that it is vulnerable to attack, he will lose whatever incentive he might otherwise have had to pay it.").

set to a defendant's ability to pay may increase victim satisfaction.⁴¹⁶ A 1992 study supported by a grant from the State Justice Institute sought to measure which factors most affected victims' level of satisfaction with restitution.⁴¹⁷ The research found that victim satisfaction was most highly correlated with the percentage of the restitution order actually paid.⁴¹⁸ That correlation was higher than the correlation between whether the restitution paid actually covered the victim's losses and victim satisfaction.⁴¹⁹ That correlation was also higher than the correlation between the amount of time it took the defendant to pay and victim satisfaction.⁴²⁰ The conductors of this study concluded that "the single best predictor of victim satisfaction with restitution is offender compliance with awards."⁴²¹ Thus, scholars have posited that restitution orders that exceed a defendant's ability to pay only serve to give victims false hope.⁴²² This research supports the conclusion that "it is more important to victims that offenders comply with restitution orders than that judges order restitution in the full amount of their loss."⁴²³ Indeed, a representative of the Judicial Conference of the United States testified to the United States Congress that "imposition of the restitution without consideration of ability to pay would make many such orders unenforceable which in turn would erode respect for the justice system on the part of victims."⁴²⁴ While more recent data

416 See generally Robert C. Davis et al., *Restitution: The Victim's Viewpoint*, 15 JUST. Sys. J. 746, 752-53 (1992).

417 *Id.* at 750-53.

418 *Id.* at 754.

419 *Id.* (explaining that whether the entirety of the victim's losses had been paid had a correlation rate of .43 with victim satisfaction, whereas percentage of award paid had a correlation rate of .58).

420 *Id.*

421 *Id.* at 756.

422 Dickman, *supra* note 408, at 1698. See also Edna Erez & Pamela Tontodonato, *Victim Participation in Sentencing and Satisfaction with Justice*, 9. JUST. Q. 393, 410 (1992); CAROL SHAPIRO, *Is Restitution Legislation the Chameleon of the Victims' Movement?*, in CRIMINAL JUSTICE, RESTITUTION AND RECONCILIATION 73, 76 (Burt Galaway & Joe Hudson eds., 1990) ("It is quite possible that enabling legislation and victims' bills of rights that stress financial compensation and restitution by offenders may exacerbate victims' sense of powerlessness within the criminal justice system. Restitution orders may raise victim expectations and even prolong their initial anger and distress.").

423 Dickman, *supra* note 408, at 1699.

424 *A Bill to Provide for Restitution of Victims of Crimes, and for Other Purposes: Hearing on S. 173 Before the S. Comm. on the Judiciary*, 104th Cong. 1, 27 (1995) (statement of Judge Maryanne Trump Barry, Chair, Comm. on Criminal Law, Judicial

would be necessary for confirmation, this data suggests that *Henry's* mandate that restitution is set to ability to pay might improve victims' experience.

B. Henry's Impact on Other Payments Imposed by Criminal Courts

1. A Survey of Other Court-Imposed Payments

Restitution is only one of many potential payments imposed on a defendant in the process of resolving his criminal case.⁴²⁵ The rise of fines, fees, and payments in Massachusetts' criminal courts reflects a national trend.⁴²⁶ Since 2010, all but two states have increased criminal court fees.⁴²⁷ Commentators have pointed to increasing costs and a lack of revenue as among the reasons jurisdictions have moved towards charging offenders for the costs of operating the criminal system.⁴²⁸ Budding bipartisan alliances to reduce mass incarceration often emphasize reducing the cost of the criminal system.⁴²⁹ Whether intentional or not, that emphasis

-
- Conf. of the United States).
- 425 A substantial amount of literature has documented the trend of requiring payments of criminal defendants. See, e.g., Anderson, *supra* note 61, at 372 (noting that "recoupment and contribution are just one more line item on a growing list of defendant obligations creating a crushing financial burden. The trend in recent years has been to charge those convicted of felonies for numerous consequences of conviction.").
- 426 *Id.* Harris et al., *supra* note 57, at 4 (concluding that "monetary sanctions are now imposed by the courts on a substantial majority of the millions of U.S. residents convicted of felony and misdemeanor crimes each year").
- 427 Karen Dolan & Jodi L. Carr, *The Poor Get Prison: The Alarming Spread of the Criminalization of Poverty*, INSTITUTE FOR POLICY STUDIES 10 (2015), <http://www.ips-dc.org/wp-content/uploads/2015/03/IPS-The-Poor-Get-Prison-Final.pdf>.
- 428 Ruback, *Benefits and Costs*, *supra* note 63, at 1779-80. See also *Fines, Fees, and Bail: Payments in the Criminal Justice System that Disproportionately Impact the Poor*, COUNCIL OF ECONOMIC ADVISERS 2 (Dec. 2015), https://obamawhitehouse.archives.gov/sites/default/files/page/files/1215_cea_fine_fee_bail_issue_brief.pdf (noting that "[i]n the 1990s, policy makers began arguing that taxpayers should not bear responsibility for these increasing costs, but rather the individuals convicted of crimes. State and local governments, who pay many of the operational costs of the criminal justice system, have increasingly turned to monetary sanctions as a source of additional revenue").
- 429 Carl Takaesu, *From Mass Incarceration to Mass Control, and Back Again: How Bipartisan Criminal Justice Reform May Lead to a For-Profit Nightmare*, 20 U. PA. J. L. & SOC. CHANGE 125, 171 (2017). See *Policing and Profit*, *supra* note 62, at 1735 ("To be sure, some of these attempts to monetize law enforcement began with benign

has contributed to “billing criminal defendants for the operation of the system that prosecutes them” and “[f]unding . . . criminal justice on the backs of people entangled in the system.”⁴³⁰ Scholars have argued that nationwide, “penal institutions are increasingly imposing a particularly burdensome and consequential form of debt on a significant and growing share of the poor.”⁴³¹

It bears setting out the full array of payments that can be imposed on adult criminal defendants in Massachusetts. At the outset of a criminal case, anyone who is determined by the court to be too poor to afford to hire an attorney is subject to a minimum \$150 fee for a court-appointed attorney.⁴³² The fee does not take into account the amount of time an attorney spends on the case and so it may be imposed on even a poor person whose case is resolved in a single court date.⁴³³ Massachusetts maintains a cash bail system.⁴³⁴ Because defendants are not legally entitled to a bail amount they can afford, it is permissible for their judicially-determined bail amounts to exceed their financial resources.⁴³⁵ Preliminary data found median bail amounts to be significantly higher for racial minorities in multiple counties of Massachusetts; in one county the median bail amount for Black defendants was found to be five times higher than it was for White defendants.⁴³⁶ There are fees separate from and in

intentions. For example, opportunities to conditionally avoid prosecution or incarceration could be thought of as offers of leniency. But not only does charging money for these options necessarily discriminate against the poor, allowing these programs to generate revenue also tends to corrupt whatever value they might otherwise provide.”).

430 Takaei, *supra* note 429, at 153.

431 Harris et al., *supra* note 57, at 1756.

432 MASS. GEN. LAWS ch. 211D, § 2A(f) (2012) (providing that “[a] person provided counsel under this chapter shall be assessed a counsel fee of \$150.”) The statute provides for waiver of the fee only where certain specified conditions are met and allows for a judge to impose at her discretion community service in lieu of payment. *Id.*

433 See *id.* See also *The Bail Process: Arrest to Arraignment*, MASS.GOV, <https://www.mass.gov/service-details/the-bail-process-arrest-to-arraignment> (last visited Dec. 27, 2017) (“If the defendant qualifies for a court-appointed attorney, the minimum counsel fee is \$150 – regardless of the amount of time the lawyer spends defending that individual.”).

434 MASS. GEN. LAWS ch. 276, § 57 (2014). See also *Brangan v. Commonwealth*, 80 N.E.3d 949, 959 (Mass. 2017).

435 *Brangan*, 80 N.E.3d at 959.

436 A 2015 report from the MassINC, an independent Boston think tank, found that “[t]he median bail for minority inmates is significantly higher than the median bail for white inmates in every county for which data are available . . .

addition to the bail amount itself. If arrested outside of court hours, defendants may have their bail preliminarily set by a commissioner wherever the defendant is being detained, often a police station;⁴³⁷ the defendant may be charged up to \$40 for this preliminary review.⁴³⁸ According to a 2014 story from Fox 25 Boston, one commissioner “made \$19,060 in [these] bail fees, \$40 at a time.”⁴³⁹ In addition, certain jails in Massachusetts charge an “administrative fee” for the mere act of accepting bail money.⁴⁴⁰

During the pendency of the case, criminal courts can issue warrants against defendants who miss their court date.⁴⁴¹ Those warrants authorize law enforcement to arrest the defendant at any time and introduce multiple costs: a default fee of \$50,⁴⁴² an arrest fee of \$75,⁴⁴³ and “expenses incurred as a result of the entry of a default.”⁴⁴⁴ A warrant can lead to suspension of the defendant’s

. [m]edian bail amounts for African-Americans in Berkshire County are five times higher than for white defendants.” Alexander Jones & Benjamin Foreman, *Exploring the Potential for Pretrial Innovation in Massachusetts* 4 (MassINC Policy Brief Sept. 2015), https://massinc.org/wp-content/uploads/2015/09/bail.brief_.3.pdf. See also Jan Ransom, *Minorities more likely to be ordered held awaiting trial, study says*, Bos. GLOBE (Sept. 30, 2015), <https://www.bostonglobe.com/metro/2015/09/29/black-defendants-more-likely-held-before-trial-than-whites-some-counties-study-finds/BP2z3aZYHkvUhy5NNxHaxL/story.html>.

437 See Quinn v. State Ethics Comm’n, 516 N.E.2d 124, 125 (Mass. 1987) (“A bail commissioner . . . does not normally perform services in a courthouse or during normal court hours but rather holds a bail hearing in a police station or other place of detention where a person is being held under arrest at night or on weekends or holidays . . . Bail commissioners are entitled to charge a fee to prisoners whose bail hearings they conduct. G.L. c. 262, § 24. Almost always they receive the maximum authorized fee.”).

438 MASS. GEN. LAWS ch. 262, § 24(a) (2002).

439 Mike Beaudet, *FOX Undercover Investigates Mass. Bail System*, FOX 25 BOSTON (May 12, 2014), <http://www.fox25boston.com/news/fox-undercover-investigates-mass-bail-system-1/141886972>.

440 See, e.g., *Posting Bail*, WORCESTER CNTY. SHERIFF’S OFFICE, <https://worcestercountysheriff.com/facility/posting-bail/> (last visited Dec. 27, 2017) (on which the Worcester County Sheriff announces that “[i]n addition to the dollar amount of bail, individuals who are posting bail must pay the bail administrative fee of \$40.00 (must also be paid in cash only).”).

441 MASS. R. CRIM. P. 6(a)(2). This fee is applicable to defendants who are over the age of twelve.

442 MASS. GEN. LAWS ch. 276, § 30 ¶ 1 (2000).

443 *Id.* at ¶ 2; MASS. TRIAL CT. DIST. CT. DEP’T, POTENTIAL MONEY ASSESSMENTS IN CRIMINAL CASES (2015), <http://www.mass.gov/courts/docs/courts-and-judges/courts/district-court/potential-moneyassessment-criminalcases.pdf>.

444 MASS. R. CRIM. P. 6 (d)(1). See also MASS. GEN. LAWS ch. 280, § 6 (2017)

driver's license.⁴⁴⁵ The cost for reinstating a suspended license: \$100.⁴⁴⁶ Where the prosecution elects to dismiss a criminal case, dismissal can be conditioned on the defendant paying the "reasonable and actual expenses of the prosecution."⁴⁴⁷

It is worth stressing that many of these particular fees thus far discussed are applicable to even those defendants whose charges are dismissed and those defendants who are found not guilty.⁴⁴⁸ In the current set-up, there is nothing expressly barring a judge from imposing some of the statutory fees on someone whose conviction is ultimately vacated,⁴⁴⁹ or someone whom a jury determines is not guilty of the charged crime.

A defendant who is found guilty or who pleads guilty may then face an additional set of payments as a part of her punishment. Defendants found guilty of certain crimes are punished with a fine; the amount of the fine is crime-specific.⁴⁵⁰ Defendants who are placed

(“A justice may impose reasonable costs as a result of a default by a criminal defendant that was intentional or negligent and without good cause.”).

445 *Criminal Defaults*, MASS. REGISTRY OF MOTOR VEHICLES, <http://www.massrmv.com/SuspensionsandHearings/Suspensions/CriminalDefaults.aspx> (last visited Dec. 27, 2017).

446 *Id.*

447 MASS. GEN. LAWS ch. 280, § 6 (2017) (“A justice may, as a condition of the dismissal or placing on file of a complaint or indictment . . . order the defendant to pay the reasonable and actual expenses of the prosecution.”).

448 See, e.g., Kate Levine, *If You Cannot Afford a Lawyer: Assessing the Constitutionality of Massachusetts’s Reimbursement Statute*, 42 HARV. CIV. RTS. – CIV. LIBERTIES L. REV. 191, 191 (2007) (a law review article in which a law student representing a criminal defendant through a clinic recounts that the client was charged the \$150 legal counsel fee even though the charges were dismissed and “any competent lawyer could have dispensed with the case with about twenty minutes of work, including the court appearance.”) The client lived on welfare benefits and after paying her rent had \$50 of disposable income per month, yet the judge imposed the legal counsel fee. *Id.* at 191-92.

449 Indeed, the Supreme Judicial Court has expressly rejected a defendant’s argument that the \$780 he had paid in probation fees and \$50 he had paid as a victim-witness assessment should be returned to him after his drug conviction was vacated based on the fact that the substance officers seized from him had been tested by a state chemist, Annie Dookhan, who has admitted to falsifying evidence. See *Commonwealth v. Martin*, 63 N.E.3d 1107, 1108-09 (Mass. 2016). For more information on Ms. Dookhan, see Katie Mettler, *How a Lab Chemist Went from ‘Superwoman’ to Disgraced Saboteur of More than 20,000 Drug Cases*, WASH. POST: MORNING MIX (Apr. 21, 2017), https://www.washingtonpost.com/news/morning-mix/wp/2017/04/21/how-a-lab-chemist-went-from-superwoman-to-disgraced-saboteur-of-more-than-20000-drug-cases/?utm_term=.2f9071037d3c.

450 See MASS. GEN. LAWS ch. 279, §§ 1, 1A, 7, 9, 10, 11 (2017).

on probation must pay to be on probation; probation fees can be as much as \$65 per month.⁴⁵¹ Defendants must also pay separately for certain conditions of probation. For instance, if a defendant is ordered to complete a certified batterers' intervention program, he must pay a \$350 fee for the course.⁴⁵² Defendants who are ordered to wear a GPS device must pay a daily fee.⁴⁵³ Defendants convicted of certain drug-related offenses are charged a drug analysis fee that can run as high as \$500.⁴⁵⁴ In addition, defendants can be charged a victim/witness assessment fee of not less than \$90 if convicted of a felony and a \$50 fee if convicted of a misdemeanor.⁴⁵⁵ This fee can be imposed even where the crime did not involve a victim.⁴⁵⁶ Defendants can be charged for the cost of submitting their DNA for examination,⁴⁵⁷ DNA that may be used against them in the prosecution. A whole series of fees running as high as \$250 attaches specifically to driving related charges.⁴⁵⁸

Once incarcerated, inmates in Massachusetts can be charged for a variety of basic services, such as phone calls and medical copays.⁴⁵⁹ Without attempting to provide an exhaustive list of the many charges that can be imposed on inmates, it is illustrative to note that inmates in Massachusetts prisons can even be charged

451 MASS. GEN. LAWS ch. 276, § 87A ¶¶ 2-9 (2016).

452 MASS. GEN. LAWS ch. 209A, § 10 (2017).

453 MASS. GEN. LAWS ch. 265, § 47 (2006); MASS. TRIAL CT. DIST. CT. DEP'T, POTENTIAL MONEY ASSESSMENTS IN CRIMINAL CASES (2015), <http://www.mass.gov/courts/docs/courts-and-judges/courts/district-court/potential-moneyassessment-criminalcases.pdf>; MASS. S. POST AUDIT AND OVERSIGHT COMM., FINE TIME MASSACHUSETTS: JUDGES, POOR PEOPLE AND DEBTORS' PRISON IN THE 21ST CENTURY, S.2504, (2016).

454 MASS. GEN. LAWS ch. 280, § 6B (2017).

455 MASS. GEN. LAWS ch. 258B, § 8 (2014). The fees are only applicable to defendants seventeen years and older.

456 The reference here is to individual third party human victims. *See, e.g.*, Commonwealth v. Martin, 63 N.E.3d 1107, 1109 (Mass. 2016) (affirming imposition of victim witness fee in drug possession case).

457 MASS. GEN. LAWS ch. 22E, § 4(b) (2012).

458 For instance, MASS. GEN. LAWS ch. 90C, § 3(C)(3) (2016) provides for a civil motor vehicle assessment; MASS. GEN. LAWS ch. 90, § 24D (2006) provides for an OUI fee of \$250; § 24(1)(a)(1) ¶ 3 provides for a \$50 assessment to the OUI victims fund and § 24 (1)(a)(1) ¶ 2 provides for a \$250 assessment, the majority of which goes into a head injury treatment services fund.

459 *Inmate Fees as a Source of Revenue Review of Challenges*, EXEC. OFFICE OF PUB. SAFETY AND SEC. 3 (2011), <http://www.mass.gov/eopss/docs/eops/inmate-fee-final-7-1-11.pdf>; MASS. GEN. LAWS ch. 124, §1(r)-(s) (2003).

a fee for haircuts.⁴⁶⁰ As a practical matter, these costs are often passed onto the inmates' family members.⁴⁶¹ Indeed, a statutorily created Special Commission studying whether Massachusetts jails should charge inmates for room and board ultimately recommended against the adoption of those charges because *inter alia* "there are a number of fees . . . [that] may be imposed pre-incarceration, during incarceration, and post-incarceration, often increasing the debt-burden 'on a population uniquely unable to make payments.'"⁴⁶² The Special Commission noted that "[t]he entire judicial system relies heavily on fees as sources of revenue, such that increasing inmate fees puts Houses of Correction and the Department of Correction in competition for scarce resources."⁴⁶³

460 MASS. GEN. LAWS ch. 124, §1(r) (2003).

461 See SPECIAL COMM'N TO STUDY THE FEASIBILITY OF ESTABLISHING INMATE FEES, EXEC. OFFICE OF PUB. SAFETY AND SEC., INMATE FEES AS A SOURCE OF REVENUE: REVIEW OF CHALLENGES 3 (2011), <http://www.mass.gov/eopss/docs/eops/inmate-fee-final-7-1-11.pdf>; Lauren-Brooke Eisen, *Paying for Your Time: How Charging Inmates Fees Behind Bars May Violate the Excessive Fines Clause*, BRENNAN CTR. FOR JUST. AT N.Y. UNIV. SCH. OF L. (July 31, 2014), <https://www.brennancenter.org/analysis/paying-your-time-how-charging-inmates-fees-behind-bars-may-violate-excessive-fines-clause> (arguing against charging inmates for their incarceration and contending that "family members often pick up the tab for these fees, depositing funds in inmate commissary accounts so they do not leave jail or prison with criminal justice debt"). See also Milton J. Valencia, *Advocates, Families Fight Jail Phone Fees*, Bos. GLOBE (July 13, 2012), <https://www.bostonglobe.com/metro/2012/07/12/inmates-families-challenging-costly-phone-bills/SWxp4emcQXp05c0pL2Z2AO/story.html#comments> (describing the costs that accrue for the family members of inmates who accept their phone calls); Connor Lentz, *The High Toll of Inmate Phone Calls*, COMMONWEALTH MAG. (July 11, 2017), <https://commonwealthmagazine.org/criminal-justice/the-high-toll-of-inmate-phone-calls/> (quoting State Representative Chynah Tyler regarding the high cost of inmate phone services, "The communities that are most affected are low income and people of color . . . This is not a financial burden that should be falling on these families").

462 The Special Commission was composed of representatives from across law enforcement, including sheriffs and attorneys. The Commission released its findings in this 2011 report. See SPECIAL COMM'N TO STUDY THE FEASIBILITY OF ESTABLISHING INMATE FEES, EXEC. OFFICE OF PUB. SAFETY AND SEC., *supra* note 461, at 3-4 (citing REBEKAH DILLER ET AL., BRENNAN CTR. FOR JUST. AT N.Y. UNIV. SCH. OF L., CRIMINAL JUSTICE DEBT: A BARRIER TO REENTRY (2010)).

463 SPECIAL COMM'N TO STUDY THE FEASIBILITY OF ESTABLISHING INMATE FEES, EXEC. OFFICE OF PUB. SAFETY AND SEC., *supra* note 461, at 12. Notably, efforts are still being made to charge inmates for room and board. See, e.g., Elisha Machado, *Massachusetts Inmates May be Billed for Jail*

Ralph Gants, the Chief Justice of the Supreme Judicial Court, has summarized the collective impact of some of these fees: “[f]or an indigent defendant convicted of one felony and sentenced to one year of supervised probation, the fees total \$1,020, more if a GPS bracelet is a condition of probation . . . should we not stop and ask: who are we asking to pay these fees?”⁴⁶⁴ The answer: for a single individual living in poverty in Massachusetts, \$1,020 amounts to nearly the entirety of her monthly income or approximately seven percent of her annual income.⁴⁶⁵ Meanwhile, at the state level, the amount of monies collected through these payments is stunning: for example, \$21 million in probation supervision fees, \$7 million in indigent counsel fees, and \$2.4 million in victim-witness fees per year.⁴⁶⁶ Over \$30 million collected from just those three fees. While judges are permitted to waive some of these fees, “there does not exist a uniform legal standard for waiver of a fee due to inability to pay”⁴⁶⁷ and data suggests that judges allow waivers inconsistently⁴⁶⁸ and infrequently.⁴⁶⁹ According to Massachusetts Senator Mike

Time, 22 NEWS WWLP.COM (June 23, 2017), <http://wwlp.com/2017/06/23/massachusetts-inmates-may-be-billed-for-jail-time/>; Brian Fraga, *Sheriff Hodgson tries again for \$5 daily inmate fee*, THE HERALD NEWS (Jan. 8, 2017), <http://www.heraldnews.com/news/20170108/sheriff-hodgson-tries-again-for-5-daily-inmate-fee>.

- 464 Chief Justice Ralph D. Gants, Massachusetts Supreme Judicial Court, Annual Address: State of the Judiciary (Oct. 20, 2015), <http://www.mass.gov/courts/docs/sjc/docs/speeches/sjc-chief-justice-gants-state-of-judiciary-speech-2015.pdf>.
- 465 These calculations are based off the standard for indigency set out by the Supreme Judicial Court. According to the SJC, a family of 1 with an annual income of \$15,075 qualifies as indigent. See Supreme Judicial Court, *Poverty Guidelines for Affidavits of Indigency* 4 (Feb. 22, 2017), <http://www.mass.gov/courts/docs/sjc/docs/povertyguidelines.pdf>.
- 466 Gants, *supra* note 464, at 9-10.
- 467 MASS. TRIAL CT. FINES AND FEES WORKING GRP., REPORT TO TRIAL COURT CHIEF JUSTICE PAULA M. CAREY 6 (2016), <http://www.mass.gov/courts/docs/trial-court/report-of-the-fines-and-fees-working-group.pdf>. Moreover, “[t]he waiver must be made after a hearing, and the judge’s finding must be made not only on the court record, but also in writing. Even orders for partial payment require a hearing and written findings.” *Id.* at 7.
- 468 See COMMONWEALTH OF MASS. OFFICE OF THE STATE AUDITOR, TRIAL COURT – ADMINISTRATION AND OVERSIGHT OF PROBATION SUPERVISION FEE ASSESSMENTS (2016), <http://www.mass.gov/auditor/docs/audits/2016/201451603j.pdf>.
- 469 Emily Cutts, *Probation Comes at High Price for the Poor*, GREENFIELD RECORDER (Dec. 20, 2016), <http://www.recorder.com/Report-Poorest-pay-for-probation-6789877> (“Barrett said he was surprised to learn that judges

Barrett, “[m]ore waivers are issued in [the wealthy neighborhood of] Newton than in the poorest district court jurisdictions of Massachusetts.”⁴⁷⁰

2. *Practical Consequences of the Imposition and Collection of Payments*

When criminal courts require payment from a defendant, the threat of incarceration is ever present for failure to pay. While most state constitutions include a provision explicitly prohibiting debtors prisons,⁴⁷¹ the Massachusetts constitution contains no such prohibition.⁴⁷² Instead, Massachusetts statutory law expressly allows judges to jail defendants to “pay off” their court debt: every day in jail knocks \$30 off the defendant’s debt.⁴⁷³ The weight of that threat is tremendous.⁴⁷⁴ It leaves poor defendants with limited options: (1) put whatever little money they have towards court costs rather than necessary life-expenses such as shelter, food, clothing, insurance, child care, commuting, education, and job training;⁴⁷⁵ (2) borrow

were likely to use fee waivers disproportionately. There is already a system in place meant to help those unable to afford the fees. Massachusetts judges have the ability to waive probation fees but it is used infrequently and unevenly, according to the report.”). Massachusetts legislators have recently sought to improve waiver practice. *See Andy Metzger, Mass. Senate clears way for probation fee waiver*, THE PATRIOT LEDGER (May 26, 2016), <http://www.patriotledger.com/news/20160526/mass-senate-clears-way-for-probation-fee-waiver>

470 Cutts, *supra* note 469.

471 Wagner, *supra* note 284.

472 *State Bans on Debtors’ Prison and Criminal Justice Debt – Appendix*, 129 HARV. L. REV. F. 153 (2016), <https://harvardlawreview.org/2016/02/state-bans-on-debtors-prisons-and-criminal-justice-debt-appendix/> (identifying the states whose constitutions prohibit incarceration for debt). The League of Women Voters of Massachusetts has supported proposed legislation to minimize “fine time” and in a letter to the Joint Committee of the Judiciary, the League highlighted the lack of a constitutional prohibition on jailing people for debt in Massachusetts. *See Colleen Kirby, Specialist on Courts and Criminal Justice Reform, League of Women Voters of Mass., Testimony Submitted to the Joint Committee on the Judiciary in Support of S.777/H.2359 – An Act to Reduce the Criminalization of Poverty* (June 5, 2017), https://lwwma.org/wp-content/uploads/2017/06/S.777.Testimony-June-5-2017.final_.pdf.

473 MASS. GEN. LAWS ch. 127, § 144 (2017).

474 Wagner, *supra* note 284, at 403 (noting that “nothing about Bearden prevents probation officers from threatening to revoke probation for failure to pay, so the prospect of imprisonment remains a powerful stimulus to payment”).

475 Neil L. Sobol, *Charging the Poor: Criminal Justice Debt and Modern-Day Debtors’ Prisons*, 75 MD. L. REV. 486, 486-87 (2016), <http://digitalcommons.law.umaryland.edu/cgi/viewcontent.cgi?article=3701&context=mlr> (“The

money from family and friends who are often poor and struggle with basic financial obligations of their own;⁴⁷⁶ or (3) face debt-based jail time.⁴⁷⁷ According to the Prison Policy Initiative, the effect is that “the state uses its carceral power to pressure people to prioritize payment of fees over other needs and responsibilities.”⁴⁷⁸

a. The Undeniable Reality of Debt-Based Jail Time in Massachusetts

Stories regarding people threatened with jail for inability to pay have recently caught the attention of local news. In 2016, the Boston Globe told the story of a man who appeared in Dudley District Court to address outstanding fines from a drug arrest years earlier.⁴⁷⁹ He sought to address the fines in order to get his driver’s license and apply for a job.⁴⁸⁰ When he appeared in court, the judge jailed him for 36 days to pay off the \$1000 debt, even though he had no money and had recently stayed at a homeless shelter.⁴⁸¹ In another instance, the Community Newspaper Holdings Inc. told the 2015 story of a local Worcester man who was threatened with jail if he could not produce the \$600 he owed the court.⁴⁸² The man, who was unemployed and struggling to pay child support, was expressly

impact of criminal justice debt is especially severe on the poor and minorities... [o]ften they have to decide between paying criminal justice debt and buying family necessities.”).

476 See generally SANETA DEVUONO-POWELL ET AL., ELLA BAKER CTR. FOR HUMAN RIGHTS, WHO PAYS? THE TRUE COST OF INCARCERATION ON FAMILIES 12-14 (2015), <http://ellabakercenter.org/sites/default/files/downloads/who-pays.pdf> (noting, for example, that “[f]amilies are often forced to choose between supporting an incarcerated loved one and meeting basic needs for their families and themselves” and “[t]he weight of these fees, which can total nearly a year’s income for some families, pushed many of the family members surveyed to take out loans or fall into financial dire straits as a result”).

477 MASS. GEN. LAWS ch. 127, § 144 (2017).

478 Sawyer, *supra* note 9.

479 Milton J. Valencia, *Scores of Indebted Become ‘Fine-Time’ Inmates*, BOS. GLOBE (Nov. 7, 2016), <https://www.bostonglobe.com/metro/2016/11/06/study-finds-scores-inmates-who-serve-fine-time/B0eipuyXLzeDYLrwicfhUJ/story.html>.

480 *Id.*

481 *Id.*

482 Christian M. Wade, *Low-level offenders in Massachusetts struggle to cover probation fees*, CNHI (June 3, 2016), http://www.cnhi.com/featured_stories/low-level-offenders-in-massachusetts-struggle-to-cover-probation-fees/article_7908a2f4-29d0-11e6-9edf-7bf156c39029.html.

denied the opportunity to complete community service.⁴⁸³ Finally, according to the Metro West News, in 2015 a Framingham man was arrested on a warrant for not paying \$450 in court fees.⁴⁸⁴ Once in court, he told the judge that he had no money.⁴⁸⁵ The judge told the defendant that he could satisfy the fees with jail time and the judge said that he would only release the man if the man could offer “a creative idea.”⁴⁸⁶ The judge decided to release the defendant only after he offered his sneakers as bail.⁴⁸⁷ The sneakers were to be returned only after the defendant had paid the fees or completed community service.⁴⁸⁸

These are not isolated incidents. In November 2016, a Massachusetts Senate Post Audit and Oversight Committee (“Senate Committee”) released a report examining the practice of jailing poor defendants for their inability to pay.⁴⁸⁹ The Senate Committee concluded that “fine time is alive and well in Massachusetts district courts.”⁴⁹⁰ The Senate Committee reviewed one hundred and five cases from three counties in Massachusetts.⁴⁹¹ In each of the one hundred and five cases, defendants had been incarcerated for failure to pay fines.⁴⁹² Inmates refer to this practice as “fine time” and the mere existence of a nickname may speak to the frequency with which it occurs.⁴⁹³ The Senate Committee’s review not only confirmed that defendants in Massachusetts are still being jailed for failing to make court payments but also closely examined the process that led to the incarceration.⁴⁹⁴

The Senate Committee’s review offers support for the notion that the majority of defendants serving fine time are charged with minor offenses and many of them were not convicted of the crime at

483 *Id.*

484 Norman Miller, *In Lieu of Cash, Framingham Judge Accepts Pair of Nikes for Bail*, METROWEST DAILY NEWS (Jan. 2, 2015), <http://www.metrowestdailynews.com/article/20150102/NEWS/150109330>.

485 *Id.*

486 *Id.*

487 *Id.*

488 *Id.*

489 MASS. S. COMM. ON POST AUDIT AND OVERSIGHT, FINE TIME MASSACHUSETTS: JUDGES, POOR PEOPLE, AND DEBTORS’ PRISON IN THE 21ST CENTURY, S. 2504, 189TH Sess. (2016).

490 *Id.*

491 *Id.* at 11.

492 *Id.* at 10-13.

493 *Id.* at 5, 10.

494 *Id.* at 11.

issue.⁴⁹⁵ Of the one hundred and five cases, most of them involved charges that were relatively minor.⁴⁹⁶ Forty percent of the cases involved driving infractions that did not include operating under the influence.⁴⁹⁷ In sixteen percent of the cases the underlying crime was disorderly conduct, public drinking or trespassing.⁴⁹⁸ Fifty-nine percent of the cases were resolved without a conviction.⁴⁹⁹ One of the jailed individuals owed a mere thirty dollars.⁵⁰⁰

The procedural protections offered to these defendants before they were jailed fell short of the requirements outlined in *Bearden, Gomes, Canadyan* and now *Henry*. First, the Committee found that ability to pay was rarely considered.⁵⁰¹ Only six percent of the cases contained an indication that the judge made a finding regarding the defendant's ability to pay before imposing jail time.⁵⁰² Yet, the Committee found in these cases recurring suggestions in the record of difficult finances including express statements of homelessness and unemployment.⁵⁰³ Still more, the majority of the defendants in the cases reviewed were not represented by an attorney at the hearing where jail time was imposed.⁵⁰⁴ In no more than forty-six percent of the cases studied were defendants represented or offered representation.⁵⁰⁵ Finally, the Senate Committee concluded that the review of the cases also raised an “unexpected[]” and “fresh question” regarding whether the defendants had received notice on the prior court date that they would be incarcerated if they failed once more to make their payments.⁵⁰⁶ The Senate Committee described this notice question as a matter of “fundamental fairness” and “due process.”⁵⁰⁷

The Senate Committee offered its view of the reasons for this practice of incarcerating defendants unable to make payment.⁵⁰⁸ The Committee described fine time as “a function of unawareness at the top, unchallenged and sometimes arbitrary power at the bottom, and

495 *Id.*

496 *Id.*

497 *Id.*

498 *Id.*

499 *Id.*

500 *Id.* at 12.

501 *Id.* at 13.

502 *Id.*

503 *Id.*

504 *Id.* at 14.

505 *Id.*

506 *Id.*

507 *Id.*

508 *Id.* at 25.

pressure for revenue throughout.”⁵⁰⁹ The Report further concluded that “[i]mprisoning poor people for their inability to pay fines, fees and court costs – essentially, for not having money – is the logical consequence of the Legislature’s designating the ground-level criminal courts as sources of serious revenue for the state.”⁵¹⁰

b. The Financial Burden on the State’s Marginalized Communities

Where the fees are paid and jail time is avoided, the financial burden associated with criminal court involvement is severe not only for individual defendants but also for their families and neighborhoods.⁵¹¹ A 2015 report by the Ella Baker Center for Human Rights regarding the finances of incarceration found that interaction with the criminal system imposes “large sums of debt” on entire families and those families are overwhelmingly poor.⁵¹² Relying on data collected from fourteen states, the Ella Baker Center concluded that “[a]cross respondents of all income brackets, the average debt incurred for court-related fines and fees alone was \$13,607.”⁵¹³ The report explained: “[M]any of the extensive costs associated with legal defense, detention, sentencing and incarceration fall on family members. Families are often forced to choose between supporting an incarcerated loved one and meeting basic needs for their families and themselves.”⁵¹⁴ The report concluded that the “[c]ommunities most heavily impacted by incarceration are some of our nation’s most economically disadvantaged, and are disproportionately communities of color.”⁵¹⁵

These conclusions may ring particularly true in Massachusetts where data indicates that there still exists pronounced residential segregation—in terms of both income and race.⁵¹⁶ For instance,

509 *Id.*

510 *Id.*

511 See, e.g., SPECIAL COMM’N TO STUDY THE FEASIBILITY OF ESTABLISHING INMATE FEES, EXEC. OFFICE OF PUB. SAFETY AND SEC., *supra* note 461, at 3 (noting that “many inmates rely on family members for contributions . . . fees will not only impact inmates but also their family members”).

512 DEVUONO-POWELL ET AL., *supra* note 476, at 7, 9.

513 *Id.*

514 *Id.* at 12.

515 *Id.*

516 See generally Brian Steele, *Boston among most racially segregated U.S. metro cities; Hispanics affected most, report finds*, MASSLIVE (Aug. 28, 2015), http://www.masslive.com/news/boston/index.ssf/2015/08/boston_is_7th_most_racially_se.html.

the Boston-Cambridge-Newton metropolitan area of Eastern Massachusetts has been ranked as the nation's seventh most racially segregated area.⁵¹⁷ A sociologist at Brown University and the leader of the 2010 U.S. Census Project at Brown University concluded that of the 50 metro areas with the largest Black populations in 2010, Boston ranked 11th place for most extreme Black-White segregation.⁵¹⁸ The Boston Public Health Commission has acknowledged that "Boston was ranked among the top 20% of highly segregated metropolitan areas in the United States in 2010, alongside Cincinnati, Ohio and Birmingham, Alabama."⁵¹⁹ A 2017 report on Metro Boston by the Metropolitan Area Planning Council indicated that "the region's poorest households are becoming increasingly concentrated into low-income neighborhoods with little income diversity."⁵²⁰ That report further stated that "[e]ven as it grows more diverse, the region remains racially and economically segregated."⁵²¹ In metro Boston "[t]he average income for the highest-earning fifth of households (\$272,500) is 18 times higher than average income for the lowest-income fifth of households (\$14,900)."⁵²² Moreover, "income polarization is increasing [in metro Boston], and . . . disproportionately affects Black and Latino residents."⁵²³

The racial disparity in poverty rates in Massachusetts is stark. While approximately 19.4% of Black families, 28.2% of Latino families and 10.3% of Asian families in Massachusetts qualify as

517 *Id.*

518 Meghna Chakrabarti, *2010 Census Shows Boston Among Most Segregated Cities*, WBUR (Apr. 4, 2011), <http://radioboston.legacy.wbur.org/2011/04/04/boston-segregated> (writing that "[w]hen we last talked to sociologist John Logan, leader of the U.S. 2010 Census Project at Brown University . . . Logan has come up with some sobering observations about Greater Boston: Among the nation's big cities, Boston is in 11th place for the most extreme residential segregation between blacks and whites"). See also JOHN R. LOGAN & BRIAN J. STULTS, THE PERSISTENCE OF SEGREGATION IN THE METROPOLIS: NEW FINDINGS FROM THE 2010 CENSUS 6 (2011), <https://s4.ad.brown.edu/Projects/Diversity/Data/Report/report2.pdf>.

519 BOS. PUB. HEALTH COMM'N, 2014-2015 HEALTH OF BOSTON 50 (2015), http://www.bphc.org/healthdata/health-of-boston-report/Documents/HOB-2014-2015/1_Demographics_HOB%202014-2015.pdf (citing LOGAN & STULTS, *supra* note 518).

520 GENE FOSTER ET AL., METRO BOSTON REGIONAL INDICATORS: STATE OF EQUITY 2017 UPDATE 4 (2017), ftp://ftp.mapc.org/StateOfEquity/SOE_2017Update_ExecutiveSummary.pdf.

521 *Id.*

522 *Id.* at 7.

523 *Id.*

impoverished, only 4.8% of White families qualify as impoverished.⁵²⁴ The concentration of poverty in communities of color is also painfully reflected in childhood poverty rates. Researchers from Tufts University have found that “[t]he overwhelming majority (85.3%) of all young people in Boston who are 17 years and under, and impoverished, are Blacks and Latinos.”⁵²⁵ Data indicates that 42% of children in Dorchester, Roxbury and Mattapan lived in poverty in 2011.⁵²⁶ Of Massachusetts’ cities, Boston has the largest number of Black residents;⁵²⁷ Roxbury, Dorchester and Mattapan, in turn, have the largest numbers of Black residents in Boston’s neighborhoods.⁵²⁸

Research also confirms that the state’s communities of

-
- 524 JAMES JENNINGS ET AL., BLACKS IN MASSACHUSETTS: COMPARATIVE DEMOGRAPHIC, SOCIAL AND ECONOMIC EXPERIENCES WITH WHITES, LATINOS, AND ASIANS 61 (2015), <http://emerald.tufts.edu/~jjenni02/pdf/BlackComparativeExperience.pdf>. For poverty rates for the Boston metro area, see Steele, *supra* note 516 (stating that “[w]hile approximately 21.5 percent of black residents in the Boston metro area live in poverty and 26.1 percent of Hispanics do, only 7 percent of white residents in the Boston metro area live in poverty”).
- 525 JAMES JENNINGS, SOCIAL, DEMOGRAPHIC, AND ECONOMIC PROFILE OF YOUNG BLACK AND LATINO MALES BOSTON, MASSACHUSETTS 2010-2018 (2014), <https://static1.squarespace.com/static/5515d04fe4b0263cc20b3984/t/555ce9c8e4b0373c4038aadbd/1432152520097>Status+of+Black+and+Latino+Young+Males+April+2014.pdf>.
- 526 Meghan E. Irons, *Poverty Worsening in Hub, Study Says*, BOSTON.COM (Nov. 9, 2011), http://archive.boston.com/news/local/massachusetts/articles/2011/11/09/poverty_concentrated_deepening_in_mattapan_dorchester_and_roxbury/.
- 527 JENNINGS ET AL., *supra* note 524, at 15.
- 528 ALIX CANTAVE, WILLIAM MONROE TROTTER INST., CRIME IN THE AFRICAN-AMERICAN NEIGHBORHOOD 6 (2007), https://scholarworks.umb.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1000&context=trotter_pubs (“There are three neighborhoods in Boston in which blacks or African-Americans are the majority group. These neighborhoods are Mattapan, Roxbury, and Dorchester . . . Boston is a highly segregated city with 70 percent of the black population in Dorchester, Mattapan, and Roxbury.”); Wesley Lowery, *New Study: Only 24% of Population, Blacks in Boston Make Up 63% of Stop and Frisk Encounters*, WASH. POST (Oct. 8, 2014), https://www.washingtonpost.com/news/post-nation/wp/2014/10/08/new-study-only-24-of-population-blacks-in-boston-make-up-63-of-stop-and-frisk-encounters/?utm_term=.65b24c9dedd0 (referring to the “minority-rich neighborhoods of Roxbury, Dorchester and Mattapan”); George Walters-Sleyon, *Studies on Religion and Recidivism: Focus on Roxbury, Dorchester, and Mattapan*, 21 TROTTER REV. 22, 24 (2013), http://scholarworks.umb.edu/cgi/viewcontent.cgi?article=1340&context=trotter_review (“According to the 2010 Census, the largest numbers of blacks in Massachusetts are concentrated in Roxbury, Dorchester, and Mattapan.”).

color experience a disproportionate amount of interaction with the criminal system. According to an Associated Press review, in Boston “[a]t least 71 percent of all street level, police-civilian encounters from 2015 through early 2016 involved persons of color.”⁵²⁹ Arrests are concentrated in the neighborhoods with significant Black populations: “[a]rrests reported in Dorchester, Mattapan, and Roxbury surpassed all other residential neighborhoods in [Boston].”⁵³⁰ The ACLU found that in 2014 Black people in Massachusetts were arrested for marijuana sales at 7 times the rate White people were arrested.⁵³¹ In addition, the incarceration rate for Black people in Massachusetts is nearly eight times the rate for White people.⁵³² Taken together, Hispanic⁵³³ and Black people constitute approximately 55 percent of the incarcerated population in Massachusetts⁵³⁴ even though one recent estimate indicates that Massachusetts is approximately 74 percent (non-Hispanic) White.⁵³⁵

-
- 529 Philip Marcelo, *APNewsBreak: Boston Police Make Little Progress on Race Gap*, U.S. NEWS & WORLD REPORT (Apr. 26, 2017, 5:50 PM), <https://www.usnews.com/news/us/articles/2017-04-26/apnewsbreak-little-change-on-race-in-boston-police-stops>.
- 530 CANTAVE, *supra* note 528, at 15.
- 531 JON B. GETTMAN ET AL., AM. CIVIL LIBERTIES UNION, THE WAR ON MARIJUANA IN BLACK AND WHITE: A MASSACHUSETTS UPDATE 8 (2016), <https://aclum.org/wp-content/uploads/2016/10/TR-Report-10-2016-FINAL-with-cover.pdf>.
- 532 MASS. SENTENCING COMM’N, SELECTED RACE STATISTICS 2 (2016), <http://www.mass.gov/courts/docs/sentencing-commission/selected-race-statistics.pdf>.
- 533 The term “Hispanic” and “Latino” are used throughout this article where the cited source measured in terms of that racial classifier. To avoid any misstatement of statistics pulled from external sources, racial classifiers from the source are maintained and used with the same definition intended in the original source.
- 534 MASS. DEP’T OF CORR., A TEN-YEAR TREND ANALYSIS OF RACE/ETHNICITY (2002-2011) (2012), <http://www.mass.gov/eopss/docs/doc/research-reports/pop-trends/a-ten-year-trend-analysis-of-raceethnicity-2002-2011.pdf>
- 535 Shira Schoenberg, *Participants in Massachusetts’ Drug Courts Are Overwhelmingly White*, MASSLIVE (Apr. 20, 2016), http://www.masslive.com/politics/index.ssf/2016/04/participants_in_massachusetts.html (“In Massachusetts, 74 percent of the population is non-Hispanic white, according to 2014 census data, and minorities are more likely to be arrested and convicted of crimes.”). Note that Boston, as a city, is more diverse than Massachusetts, as a state. As of 2010, Boston’s population was 47% White, 22.4% Black, 17.5% Hispanic and 8.9% Asian. See BOS. REDEVELOPMENT AUTH., BOSTON’S RACIAL GROUPS 3 (2014), <http://www.bostonplans.org/getattachment/68068312-d112-4b02-8ec6-26727c7e9072>. Other estimates of the percentage of the White population are even higher. See, e.g., *Population of Massachusetts:*

The Business Insider has recognized that “[d]isparities between imprisonment rates for Hispanics and Whites are particularly high in Massachusetts (4.3:1).”⁵³⁶ A report by the Boston Indicators Project, MassINC, and Massachusetts Criminal Justice Reform Coalition found that residents of the Roxbury neighborhood of Boston are detained at twice the rate of Boston residents as a whole.⁵³⁷ In the Franklin Field area of Dorchester more than one out of every five male residents between the age of 25-29 was incarcerated during the six year period between 2009 and 2015.⁵³⁸ Overall, according to CommonWealth Magazine, “[l]arge swaths of mostly minority Boston neighborhoods are so heavily affected by the criminal justice system that nearly every street has a resident who has spent time in jail.”⁵³⁹ This increased contact with the criminal system leaves the state’s low-income communities of color particularly exposed to the minefield of fees and costs that attach to being personally involved with or having a loved one involved with criminal courts.

A 2016 report from the Prison Policy Initiative further supports this notion that criminal court costs are falling disproportionately on Massachusetts’ poorer communities.⁵⁴⁰ The report focused on probation fees.⁵⁴¹ At this point, in Massachusetts, of the individuals under criminal court supervision, more are on probation than

Census of 2010 and 2000 Interactive Map, Demographics, Statistics, Quick Facts, CENSUSVIEWER, <http://censusviewer.com/state/MA> (last visited Dec. 27, 2017) (listing the White population of Massachusetts as 80.41%).

- 536 Caroline Simon, *There is a Stunning Gap between the Number of White and Black Inmates in America’s Prisons*, BUS. INSIDER (June 16, 2016, 12:13 PM), <http://www.businessinsider.fr/us/study-finds-huge-racial-disparity-in-americas-prisons-2016-6/>. For another presentation of statistics reflecting the overrepresentation of Blacks and Hispanics in Massachusetts’ incarcerated population, see generally MARC MAUER & RYAN S. KING, THE SENTENCING PROJECT, *UNEVEN JUSTICE: STATE RATES OF INCARCERATION BY RACE AND ETHNICITY* (2007), http://prisonterminal.org/documents/rd_stateratesofincbyraceandethnicity.pdf.
- 537 BENJAMIN FORMAN ET AL., *THE GEOGRAPHY OF INCARCERATION: A SPECIAL REPORT FROM THE BOSTON INDICATORS PROJECT IN PARTNERSHIP WITH MASSINC AND THE MASSACHUSETTS CRIMINAL JUSTICE REFORM COALITION* 7 (2016), <https://massinc.org/wpcontent/uploads/2016/11/The-Geography-of-Incarceration.pdf>.
- 538 *Id.* at 8-9.
- 539 Michael Jonas, *Mapping Incarceration in Boston*, COMMONWEALTH MAG. (Nov. 10, 2016), <http://csgjusticecenter.org/jr/massachusetts/media-clips/mapping-incarceration-in-boston/>.
- 540 Sawyer, *supra* note 9.
- 541 *Id.*

incarcerated.⁵⁴² An estimated 67,000 people in Massachusetts are on probation.⁵⁴³ As discussed, probationers pay monthly fees to cover the cost of their own supervision. For each probationer, those fees amount to approximately \$850-1,300 over the course of an average probation of 17-20 months.⁵⁴⁴ The Prison Policy Initiative report found that low-income people are disproportionately represented among probationers in Massachusetts.⁵⁴⁵ The difference in rates is not small: probation rates in the poorest areas are nearly twice the rates of the wealthiest areas.⁵⁴⁶ Communities in Massachusetts where per capita income of population served by the district court is between \$20,000 and \$29,999 had the highest rates of probation and communities where the per capita income was more than \$50,000 had the lowest rates of probation.⁵⁴⁷ The report concluded: "As incomes go up, probation rates go down, which means the people who can least afford additional fees are more likely to be on probation and expected to pay up every month."⁵⁴⁸ A 2017 report by the New England Public Policy Center similarly concluded that "[o]f the 20 towns in Massachusetts with the highest number of individuals on probation per capita, all but three are in the bottom quarter of per capita income in the state."⁵⁴⁹

In light of this research, questions arise regarding the extent to which Massachusetts' budget has become dependent on criminal court payments and whether there has developed a financial—rather than public safety based—incentive to impose monetary sanctions.⁵⁵⁰

542 *Id.*

543 *Id.*

544 *Id.*

545 *Id.*

546 *Id.*

547 *Id.*

548 *Id.*

549 ROBERT CLIFFORD & RILEY SULLIVAN, THE CRIMINAL POPULATION IN NEW ENGLAND: RECORDS, CONVICTIONS, AND BARRIERS TO EMPLOYMENT 13 (2017), <https://www.bostonfed.org/publications/new-england-public-policy-center-policy-report/2017/the-criminal-population-in-new-england-records-convictions-and-barriers-to-employment.aspx>.

550 For a discussion of the dilemmas that arise from state dependency on criminal court debt, see generally Criminal Justice Policy Program at Harvard Law Sch., Confronting Criminal Justice Debt: A Guide for Policy Reform 2 (2016), <http://cjpp.law.harvard.edu/assets/Confronting-Crim-Justice-Debt-Guide-to-Policy-Reform-FINAL.pdf> ("[T]he dependence of courts and other government actors on criminal justice debt is itself part of the problem. It can distort governmental decision-making in individual cases by creating conflicts of interests when judges, police officers, or other criminal justice actors make

Given the weight of these payments on the financial resources of particularly the state's low-income communities of color, each of us must consider what these statistics reveal about the participation of Massachusetts' ground-level criminal courts—a key government institution—in exasperating historically entrenched economic and racial inequalities.

3. The Potential Application of Henry Beyond Restitution

While nowhere near a wholesale solution to the complex problem of debt-based incarceration and criminal court related debt, *Henry* constitutes a tremendous addition to the case law.⁵⁵¹ According to some scholars, the fact that debt-based jail time still occurs demonstrates that the existing case law has not been fully effective in preventing the punishment of poverty.⁵⁵² Scholars have identified the holes in the existing case law and *Henry* address at least two of the most significant holes.⁵⁵³ First, scholars have contended that monetary

decisions driven by revenue-raising considerations. This can also create a vicious cycle, where courts, jails, probation agencies, and others whose budgets draw from these revenue streams worry about the consequences of reducing the flow of court-generated revenue. Faced with these pressures, legislatures may resist policy changes that remove a major funding mechanism.”).

- 551 *Henry* is already carrying much force in the restitution context. See, e.g., Commonwealth v. Amendola, No. 16-P-1073, 2017 WL 2261067, at *1–2 (Mass. App. Ct. May 23, 2017) (vacating the restitution order because “in accordance with the recent guidance provided in [*Henry*], the judge was also required to determine the defendant’s ability to pay”); Commonwealth v. McMahon, No. 16-P-191, 2016 WL 7426977, at *1 (Mass. App. Ct. Dec. 23, 2016) (“[I]t is clear that the restitution payments could only be made from funds specifically earmarked for meeting the defendant’s minimum basic human needs. In these circumstances, we conclude that the judge, who did not have the benefit of the *Henry* case, erred in extending the defendant’s probation in light of the evidence that demonstrated the defendant’s financial inability to make restitution payments without causing substantial financial hardship.”).
- 552 For example, a professor at the University of Washington has written: “The incarceration of indigent debtors is surprising given the legal prohibitions on this practice . . . it is commonly assumed that indigent debtors living in the United States are no longer incarcerated for failing to meet their financial obligations. However, emerging evidence indicates that non-payment of three types of financial obligations—consumer debt, child support decrees, and legal financial obligations—by indigent persons results in arrest and incarceration in locales across the country.” See Beckett & Murakawa, *supra* note 57, at 228–29.
- 553 For example, a 2015 Harvard Law Review article contended that certain court actions “are discriminatory [based on wealth] long before incarceration” and that “*Bearden* simply requires assurance that a person who failed to pay a fine

sanctions unjustly straddle defendants with insurmountable debt and in so doing undermine rehabilitation;⁵⁵⁴ *Henry's* proclamation that ability to pay must be determined at the outset of the restitution order and that restitution payments cannot exceed a defendant's personal ability to pay is responsive. Second, scholars have argued that courts have applied the principles animating cases like *Bearden* too narrowly to effectively protect against the myriad ways in which poor defendants are unfairly entangled in the criminal system.⁵⁵⁵ To that, *Henry* offers that poverty-based punishment can occur in forms other than incarceration—unduly prolonged entanglement with the criminal courts most prominently. For these reasons, *Henry's* framework is much needed at all junctures of the criminal system where payments are demanded of poor individuals. And there is good legal reason to conclude that *Henry's* proclamations should be read to extend to other court-imposed payments.

A broad reading of *Henry* is supported first by the constitutional language threaded throughout it. The court referenced concepts that are soaked with constitutional implications; the decision contains terms such as “fundamental fairness,” “basic fairness,”

did not do so because she could not afford the fine, at the moment she would be incarcerated.” That article offers as a solution that: “courts could root out discrimination by requiring hearings into indigence at different points in the process” and “prohibit[ing] fees (or punishments for failing to pay fees) that extend a sentence.” *Developments in the Law – Policing*, 128 HARV. L. REV. 1723, 1741-42 (2015).

554 See, e.g., Katherine Beckett & Alexes Harris, *On Cash and Conviction: Monetary Sanctions as Misguided Policy*, 10 CRIMINOLOGY & PUB. POL'Y 505 (2011), <http://criminology.fsu.edu/wp-content/uploads/volume-10-issue-31.pdf> (“In the United States, fees and fine amounts are determined by statute and are not tethered to defendants’ earnings. Recent studies suggest that the assessment of these penalties often generate long-term debts that are sizeable relative to expected earnings and impede reintegration.”).

555 For instance, Colin Reingold, of the Orleans Public Defenders, offered this perspective: “Today, *Bearden* is invoked in courtrooms throughout America to protest when judges attempt to jail a defendant for reasons that directly or indirectly stem from poverty . . . So, how is it that every day in courtrooms across America criminal defendants are sent to jail for being poor? . . . The answer, in part, is that the criminal justice system sets up the poor by pressuring them to plead guilty and then requiring repeated and unreasonably frequent contact with the criminal justice system through probation oversight, court dates for paying outstanding court debt, and frequent drug testing.” See Colin Reingold, *Pretextual Sanctions, Contempt, and the Practical Limits of Bearden-Based Debtors' Prison Litigation*, 21 MICH. J. RACE & L. 361, 362-63 (2016), <http://repository.law.umich.edu/mjrl/vol21/iss2/9>.

“fundamental principle,” and “equal justice.”⁵⁵⁶ The court recognized a “fundamental principle that a criminal defendant should not face additional punishment solely because of his or her poverty.”⁵⁵⁷ There was nothing in the court’s language that limited that fundamental principle to restitution;⁵⁵⁸ to the contrary, the very nature of a fundamental principle is its broad applicability. It is well established that constitutional provisions “commonly are to be interpreted as stating a broad and general principle of government, regulative of all conditions arising in the future and falling within their terms.”⁵⁵⁹ If the duty to affirmatively ensure that a defendant does not experience additional punishment based on her impoverishment arises from a fundamental principle, it is only logical that the duty extends to other monetary sanctions.

Second, the court’s reasoning also supports a broad reading of *Henry*. As noted in the previous section, the *Henry* court explained that it was “requir[ing] a judge to consider the defendant’s ability to pay when setting the restitution amount because . . . the collection of restitution is enforced by the threat or imposition of a criminal sanction for violation of a probation condition.”⁵⁶⁰ This justification—that restitution was enforced by the threat or imposition of a criminal sanction—was not tangential to the court’s conclusion; it was central to the court’s reasoning. It follows that any other payments that are enforced by the threat or imposition of a criminal sanction are arguably subject to at least some of the constraints that the court expressly imposed on restitution orders. This would include the affirmative duty to consider ability to pay upfront and the prohibition on imposing any payment that would require the defendant to forgo basic human needs.

Finally, the court’s reliance on sources from outside of the restitution context supports this broad reading of *Henry*. The court drew broadly from preexisting law regarding the interaction between criminal punishment and poverty. The citations in *Henry* are not limited to restitution cases.⁵⁶¹ Instead, the court relied on

556 See generally Commonwealth v. Henry, 55 N.E.3d 943 (Mass. 2016).

557 *Id.* at 950 (emphasis added).

558 *See id.*

559 Opinion of Justices to Senate, 436 N.E.2d 935, 948 (Mass. 1982).

560 *Henry*, 55 N.E.3d at 950.

561 See, e.g., *id.* (citing United States v. McGiffen, 267 F.3d 581, 588-89 (7th Cir.2001) (regarding whether a defendant is financially able to contribute to cost of appointed counsel); Museitef v. United States, 131 F.3d 714, 716 (8th Cir.1997) (regarding ability to pay for costs of appointed counsel);

cases concerning *inter alia* legal counsel fees and default fees.⁵⁶² In writing that “[t]he payment of restitution, like any court-imposed fee, should not cause a defendant substantial financial hardship,”⁵⁶³ the court was drawing from Supreme Judicial Court Rule 3:10, the rule setting out the standards for determining if a defendant qualifies for court-appointed counsel.⁵⁶⁴ The court’s reliance on language and principles applicable to other court-imposed payments supports the notion that the court was viewing these various costs as related and subject to similar considerations. Thus, the various costs imposed by criminal courts, even though not identical in nature, are arguably constrained by the guidelines, prohibitions, and standards for measuring ability to pay that are announced in *Henry*.

Read broadly, *Henry* supports the proposition that criminal court judges have an obligation at the time of issuing any payment to consider how that payment interacts with the defendant’s poverty and whether the byproduct of that interaction is constitutionally tolerable.⁵⁶⁵ If the defendant’s impoverishment would render the court order more severe, render further punishment inevitable, lead to the defendant being entangled in the criminal system for far longer than the underlying offense would otherwise merit or force the defendant to forgo basic human needs, *Henry* suggests that the punishment is constitutionally problematic.⁵⁶⁶ *Henry* instructs that where payments are part of the resolution of a case, a defendant’s ability to pay is not an afterthought, cannot be delayed and cannot be perfunctory.⁵⁶⁷ If the rules and analytical perspective of *Henry* were to guide all criminal court-imposed payments in Massachusetts, *Henry* would be a watershed case in constraining the unjust imposition and collection of criminal court debt. As usual, it is the critical work of practitioners to push forward and refine these arguments in bustling

Commonwealth v. Porter, 971 N.E.2d 291, 297 (Mass. 2012) (regarding whether the defendant qualified for court-appointed counsel); Commonwealth v. Canadyan, 944 N.E.2d 93, 96–97 (Mass. 2010) (regarding the permissibility of punishing a homeless defendant for failing to maintain his GPS device operable); Commonwealth v. Gomes, 552 N.E.2d 101 (Mass. 1990) (regarding imposition of default fees); People v. Jackson, 769 N.W.2d 630, 633 (Mich. 2009) (regarding ability to pay fees for counsel)).

562 *Henry*, 55 N.E.3d at 950-54.

563 *Id.* at 954.

564 S.J.C. Rule 3:10, as amended, 416 Mass. 1306 (1993).

565 *Henry*, 55 N.E.3d at 950 (citing *Bearden v. Georgia*, 461 U.S. 660, 668–669 (1983); *Canadyan*, 944 N.E.2d at 96-97; *Gomes*, 552 N.E.2d at 105-06).

566 See *Henry*, 55 N.E.3d at 950.

567 See *id.* at 952.

courtrooms throughout the Commonwealth.

V. Conclusion

This political moment is a promising opportunity to reexamine long-standing practices in the criminal system. Criminal court practices must constantly be assessed to ensure that those practices are consistent with the federal and state constitutions, responsive to society's ever-evolving standards of decency and practically effective. In the aftermath of *Henry*, even those monetary sanctions that appear facially neutral must be closely examined to ensure that they do not unduly burden the already marginalized and reinforce social inequality. Massachusetts courts have recognized that "a person in collision with the government ought not to be punished for his poverty."⁵⁶⁸ At an opportune time, *Henry* brings renewed energy to that principle.

* * *

568 Commonwealth v. Payne, 602 N.E.2d 594, 595 (Mass. 1992).

Comity in US Courts

Thomas Schultz & Niccolò Ridi***

* Professor of Law, King's College London; SNSF Research Professor of International Law, Graduate Institute of International and Development Studies, Geneva; Editor-in-Chief, *Journal of International Dispute Settlement*. Thomas.Schultz@kcl.ac.uk; Thomas.Schultz@graduateinstitute.ch

** LLB, University of Florence; LLM, University of Cambridge; Modern Law Review Scholar and PhD Student, Dickson Poon School of Law, King's College London; Research Fellow, Graduate Institute of International and Development Studies, Geneva. Niccolo.Ridi@kcl.ac.uk; Niccolo.Ridi@graduateinstitute.ch.

Table of Contents

I.	Introduction	5
II.	The Problem of Comity: Definitions and Methodology	8
	A. <i>Acknowledging Comity and Defining It.....</i>	8
	B. <i>The Judicial Understanding of Comity: Methodological Remarks</i>	11
III.	The History of Comity in American Legal Thinking	13
	A. <i>Comity Enters the United States.....</i>	13
	1. <i>The Sister Notions of Comity and Sovereignty</i>	13
	2. <i>Comity in the United States: Livermore, Kent and Story</i>	19
IV.	The Judicial Evolution of Comity	23
	A. <i>"Legislative" or "Prescriptive" Comity: Restraint and Recognition.....</i>	23
	1. <i>Using Comity to Grant Recognition to the Law and the Acts of Other States</i>	25
	2. <i>Using Comity to Limit the Reach of American Law</i>	28
	a. Antitrust	28
	b. Outside the Antitrust Context.....	37
	c. Analysis	40
	B. <i>Comity and the Recognition of Foreign Judicial Acts</i>	42
	C. <i>Adjudicatory Comity, or "the Comity of Courts"</i>	48
	1. <i>The Origins of Adjudicatory Comity in Admiralty Courts and the Use of Forum Non Conveniens</i>	50
	2. <i>Comity as a Coordination Device for Pending or Potential Parallel Proceedings</i>	54
	3. <i>Anti-Suit Injunctions.....</i>	61
	D. <i>"Executive" or "Sovereign-Party" Comity.....</i>	65
	1. <i>The Act of State Doctrine.....</i>	66
	2. <i>Sovereign Immunity</i>	70
	3. <i>The Privileges of Suit</i>	74
V.	United States Experience and the Global Dimension of Comity	77
	A. <i>Comity Doctrines and Comity Reasoning.....</i>	77
	B. <i>The Problem of Deference to the Executive</i>	79
	C. <i>Beyond the Domestic: the "Comity of Courts" as a Global Ordering Principle</i>	83
VI.	Conclusion.....	86

I. Introduction

Throughout much of the world, the concept of comity has played a fundamental role in the shaping of modern private international law.¹ Sometimes labelled as a “principle,” sometimes as a “doctrine,” it provided the foundation and informed the evolution of a number of rules of conflict.² Granted, in many jurisdictions, comity gradually saw its importance fade as conflict came to be absorbed into the preserve of the domestic legislator.³ Scholarly attention declined too, prompting private international lawyers to dismiss comity as the flickering shimmer of a time long past, in the name of which courts would sometimes fine-tune the reach of their national substantive law and jurisdictional rules, refrain from questioning the lawfulness of another sovereign state’s acts, and restrict themselves from issuing such judgments and orders when to do so would have amounted to an unjustifiable interference.⁴

But fading and disappearing, to paraphrase Charles I, are clean different things.⁵ Comity never really vanished, and has been invoked—as a principle, a doctrine, or unidentified institution—in a great number of cases, prompting scholars to observe that it was perhaps premature to pronounce its demise.⁶ Comity, as the English Lord Collins of Mapesbury put it, “may be a discredited concept in

1 Thomas Schultz & David Holloway, *Retour sur la comity, deuxième partie: La comity dans l’histoire du droit international privé*, 139 J. DROIT INT. 571, 593–94 (2012).

2 Elliott E. Cheatham, *American Theories of Conflict of Laws: Their Role and Utility*, 58 HARV. L. REV. 361 (1945); See generally Donald Earl III Childress, *Comity as Conflict: Resituating International Comity as Conflict of Laws*, 44 U.C. DAVIS L. REV. 11, 13–14 n.1 (2010); William S. Dodge, *International Comity in American Law*, 115 COLUM. L. REV. 2071, 2072 (2015) (discussing the various faces of comity in American law).

3 Schultz & Holloway, *supra* note 1, at 593–94.

4 Adrian Briggs, *The Principle of Comity in Private International Law*, in 354 RECUEIL DES COURS DE L’ACADEMIE DE DROIT INTERNATIONAL 65, 82, 85–86 (2012). See generally Thomas Schultz & Niccolò Ridi, *Comity and International Courts and Tribunals*, 50 CORNELL INT’L L.J. 3, 4 (2017).

5 Charles I, King of England, King Charles his speech made upon the scaffold at Whitehall Gate, immediately before his execution, on Tuesday: With a relation of the manner of his going to execution (Jan. 30, 1648).

6 For a discussion of the issue in the Australian context, see Thomas Schultz & Jason Mitchenson, *Navigating Sovereignty and Transnational Commercial Law: The Use of Comity by Australian Courts*, 12 J. OF PRIV. INT’L L. 344 (2016). See also Briggs, *supra* note 4, at 80 (stating that even if comity is rarely used in academia, it is still used in common law in many cases).

the eyes of the text-writers, but it thrives in the judicial decisions.”⁷ What is more, there is evidence that comity may have made a quiet comeback: References to the concept have been gradually increasing for some time in the decisions of certain courts, namely those of the United States of America.⁸

To some, this peculiarity will not come as a surprise: Indeed, comity has long been acknowledged as a foundational principle—if not *the* foundation—of American conflict of laws.⁹ A foundation whose importance cannot be fully understood without digging deep into a complex narrative, which unfolds in intertwined tales of misunderstandings, ethical battles, and nation-building efforts.¹⁰ Further, while unique historical circumstances offered fertile ground for the concept of comity to adhere and prosper, changes in these conditions prompted the notion to evolve, perhaps transforming, at least to some degree, into “an expression of unexplained authority, imprecise meaning and uncertain application.”¹¹ It is certainly true that the term has come to refer to a variety of practices, many of which may rightly be considered conceptually distinct from its understanding, though they share with it “certain methods, values, and justificatory rhetoric.”¹²

Be that as it may, it is undeniable that no legal system has given as much weight to the notion of comity as that of the United States of America: Its tumultuous history and peculiar development, in and by themselves, well warrants an in-depth study of the life of the concept in the USA. While the topic is discussed in a number of studies, most of them fail to consider the phenomenon in a broader dimension and thus do not properly appraise the peculiarity of the

7 Laurence Collins, *Comity in Modern Private International Law*, in REFORM AND DEVELOPMENT OF PRIVATE INTERNATIONAL LAW: ESSAYS IN HONOUR OF SIR PETER NORTH 95 (James Fawcett ed., 2002).

8 As a general example, it can be added that a Westlaw search for the exact phrase “international comity” returned 1758 hits, of which 1295 concerned cases decided after 1995, and 782 decided after 2005. See also Childress, *supra* note 2, at 15 (discussing the growing emergence and prevalence of international comity in US courts).

9 Joel R. Paul, *Comity in International Law*, 32 Harv. Int'l L.J. 1, 78 (1991) [hereinafter Paul, *Comity in International Law*]; Joel R. Paul, *The Transformation of International Comity*, LAW CONTEMP. PROBL. 19, 19 (2008) [hereinafter Paul, *Transformation of International Comity*].

10 See *infra* Section III.

11 Michael D. Ramsey, *Escaping “International Comity,”* 83 IOWA L. REV. 893, 893 (1997).

12 Paul, *Comity in International Law*, *supra* note 9, at 21.

American understanding of comity.¹³

But there is one more, equally compelling reason that prompts us to carry out this investigation: Legal ideas circulate today much in the same way they did in the early nineteenth century, when Justice Joseph Story relied on the authority of the Dutch jurist Ulrich Huber to write his influential *Commentaries*.¹⁴ Inevitably, the American understanding of comity has proved persuasive. It has affected the development of legal doctrines elsewhere, and stimulated further reflection on the role of the concept, especially when employed by prominent American scholars.¹⁵ This has in turn prompted the revitalization of comity as a tool capable of alleviating problems of a global nature.¹⁶

The purpose of this study is thus, first, to contribute to the elucidation of the notion of comity as it is understood by American scholars and practitioners, and to illustrate the distinctiveness of this American understanding. Our research builds upon the existing literature and an extensive survey of Supreme Court and Federal Appellate decisions, covering more than two centuries' worth of judicial reasoning. Through our analysis, we aim to provide an account of how comity and its localized understanding evolved to meet unique historical and political circumstances. Accordingly, we will explore the significance of the notion—or, more correctly, several notions—of comity in shaping the American understanding of conflict, foreign relations law, and the concept of public international law jurisdiction.¹⁷ We will then proceed to assess the degree to which

13 Indeed, comity has been the object of three major studies, the findings of which the present article builds upon. See, e.g., Paul, *Transformation of International Comity*, *supra* note 9; Childress, *supra* note 2; Dodge, *supra* note 2.

14 For Story's reliance on Huber, see JOSEPH STORY, *COMMENTARIES ON THE CONFLICT OF LAWS, FOREIGN AND DOMESTIC: IN REGARD TO CONTRACTS, RIGHTS, AND REMEDIES, AND ESPECIALLY IN REGARD TO MARRIAGES, DIVORCES, WILLS, SUCCESSIONS, AND JUDGMENTS* § 29. (1834).

15 As a notable exception see ANNE-MARIE SLAUGHTER, *A NEW WORLD ORDER* 67–87 (2004) (“[d]iscussing the potential of the principle of comity in a global context”).

16 See, e.g., YUVAL SHANY, *THE COMPETING JURISDICTIONS OF INTERNATIONAL COURTS AND TRIBUNALS* (Phillipe Sands et al. eds., 2003); See generally Elisa D'Alterio, *From Judicial Comity to Legal Comity: A Judicial Solution to Global Disorder?*, 9 INT. J. CONST. L. 394 (2011) (discussing the implementation of different judicial techniques to express a court's “regulating function” of governing relations between different legal systems).

17 On the relationship between private international law and the concept of jurisdiction in public international law, see ALEX MILLS, *THE CONFLUENCE*

the American understanding of comity has, by reason of its isolated development, influenced related doctrines elsewhere.

II. The Problem of Comity: Definitions and Methodology

A. Acknowledging Comity and Defining It

Let us start from the beginning: Even by the most conservative assessment, the notion of comity is employed by nearly every common law court to an extent that is not negligible. Yet, there is comparatively little agreement as to what exactly comity is, what purposes it should serve and when and how it should operate. The problem is a significant one: On the one hand, certain commentators tend to be adamantly dismissive of the notion—which, it should be pointed out, is quite unhelpful, as the notion does not for this reason cease to exist and be employed.¹⁸ But even among those who do express a degree of interest and devote some scholarly attention to comity, views tend to diverge significantly.¹⁹ What is more, definition attempts seem to be doomed at the outset, to the point that one of the leading British conflicts of law scholars dodged the question by first explaining what comity is *not*, and then attempting to provide a rather lengthy list of the concept's distinctive characteristics.²⁰

A casual reader could be forgiven for playing down the importance of defining comity as yet another purely academic quest. All the same, such a reader would be mistaken. What is at stake here is the answer to significant questions, such as: which law should apply to a given dispute; whether a court should be able to decline to exercise its jurisdiction; and what consequences should be attached to the sovereign status (or sovereign capacity) of a party to a dispute. It is disingenuous to think that the solution to these problems would be straightforward if the answer must be based—at least to some extent—on a notion so vaguely defined that

OF PUBLIC AND PRIVATE INTERNATIONAL LAW: JUSTICE, PLURALISM AND SUBSIDIARITY IN THE INTERNATIONAL CONSTITUTIONAL ORDERING OF PRIVATE LAW 226–46, 303 (2009); CEDRIC RYNGAERT, JURISDICTION IN INTERNATIONAL LAW 156–57 (2008) (discussing that prior to the US asserting its jurisdiction over foreign sovereigns, it must first weigh several factors including foreign interests, policy and strength, and whether US interest and policy outweigh the foreign sovereign's).

18 CHESHIRE, NORTH & FAWCETT: PRIVATE INTERNATIONAL LAW 5 (James J. Fawcett et al. eds., 14th ed. 2008); *see generally* Louise Weinberg, *Against Comity*, 80 GEO. L.J. 53 (1991); Ramsey, *supra* note 11.

19 *See infra* Section IV.

20 Briggs, *supra* note 4, at 87, 180.

it “invites intuitive adjudication, and hence litigation-inspiring *ex ante* unpredictability.”²¹ And yet, it is arguable that the notion of comity is quite undeserving of the measure of criticisms that it has received. Indeed, if it is so indeterminate, the impudent question could—with good reason—arise of how generations of common law judges managed to live with concepts as abstract as those of “equity” and “due process.”²² Still, comity is treated with mistrust, and many have sought to look beyond the word’s “deceptively right ring, like good breeding and sweet disposition.”²³ Indeed, according to Andreas Lowenfeld, the wariness was such that the reporters of the Third Restatement on Foreign Relations Law avoided the word altogether, believing it too charged of “the idea of discretion or even political judgment,” preferring “reasonableness, which is conceived in terms of legal obligation.”²⁴

It is perhaps for this reason that most American scholars discussing the subject attempt to anchor the discussion to formally strong authority: the definition of comity offered in the early landmark case *Hilton v. Guyot*. In the words of Justice Gray, comity:

[I]n the legal sense, is neither a matter of absolute obligation, on the one hand, nor of mere courtesy and goodwill, upon the other. But it is the recognition which one nation allows within its territory to the legislative, executive, or judicial acts of another nation, having due regard both to international duty and convenience and to the rights of its own citizens or of other persons who are under the protection of its laws.²⁵

The influence of this definition, quite possibly the most cited in scholarly and judicial writing far beyond the American panorama,

21 Ramsey, *supra* note 11, at 894 (emphasis added).

22 See N. Jansen Calamita, *Rethinking Comity: Towards a Coherent Treatment of International Parallel Proceedings*, 27 U. PA. J. INT'L ECON. L. 601, 626 (2006).

23 Weinberg, *supra* note 18, at 59.

24 RESTAMENT (THIRD) OF FOREIGN RELATIONS LAW OF THE UNITED STATES § 403; Andreas F. Lowenfeld, *Conflict, Balancing of Interests, and the Exercise of Jurisdiction to Prescribe: Reflections on the Insurance Antitrust Case*, 89 AM. J. INT. L. 42–53, 52 n.50 (1995); Paul, *The Transformation of International Comity*, *supra* note 9, at 29–30.

25 *Hilton v. Guyot*, 159 U.S. 113, 163–64 (1895).

is immeasurable.²⁶ It is thus unsurprising that American scholars would accept it—although with varying degrees of criticism and different adjustments. For example, Harold Koh's monumental work on transnational litigation pays homage to this description, but clarifies that comity “flows from the respect that one sovereign is obliged to give to the sovereign acts of a coequal nation-state.”²⁷ It adds that the notion has been increasingly interpreted by American courts “as a reason why they should refrain from independent determination of cases under the law of nations.”²⁸ This statement, while not particularly helpful, has the merit of framing the issue in the broader perspective of a discourse on sovereignty and the allocation of regulatory authority. Indeed, as we will further discuss in Part III, comity is very much a predicate of the modern model of sovereignty and — more specifically — the idea of sovereign equality, as well as a parallel to the evolution of the international law on jurisdiction.²⁹

This connection introduces one further significant problem, that is, what relevance should be attributed to the historical origin of the principle or to its connection with other notions (such as that of state jurisdiction). In other words, this is the question of whether a better understanding of how comity came to develop or a contextual awareness of a number of related notions may contribute to shedding light on its current use.³⁰ Depending on the answer to this question, the discussion on comity generally focuses on its value as a general, overarching principle of private international law,³¹ or — abandoning “the language of the comity of sovereigns” and looking at the lower level of “what judges do and why they do it” — as a judicial tool.³² In fact, whether the problem of comity can be resolved by rigidly sticking to either end of the spectrum is doubtful at best: The notion tenaciously resists traditional definition efforts,

26 Calamita, *supra* note 22, at 626.

27 HAROLD HONGJU KOH, TRANSNATIONAL LITIGATION IN UNITED STATES COURTS 19 (2008).

28 *Id.*

29 As a leading scholar of public international law put it, “[c]omity arises from the horizontal arrangement of state jurisdictions”: JAMES CRAWFORD, BROWNLEE'S PRINCIPLES OF PUBLIC INTERNATIONAL LAW 485 (8th ed. 2012); Dodge, *supra* note 2, at 2073–74.

30 On this point, see Calamita, *supra* note 22, at 606. As the author argues, “the historic, common law principles of comity — in particular adjudicatory comity — are the very same principles that underlie current case law.”

31 See, *inter alia*, Childress, *supra* note 2.

32 Briggs, *supra* note 4, at 89.

based on gradual abstraction aimed at establishing its essence, as it seems to have countless meanings. It is precisely for this reason that discussion on the doctrine ends up focusing on what comity *does* — and has done — rather than what it is.

B. The Judicial Understanding of Comity: Methodological Remarks

According to Adrian Briggs, “legal thinking in the United States, from the earliest days to modern times, has found the principle of comity to be of assistance in getting the judge to the point where a case is decided.”³³ While this remark might perhaps overlook other roles of comity, it seems to accurately describe the use of the principle by American courts: A cursory search on any major database reveals a large number of mentions of comity in judicial decisions, which is hard to reconcile with the principle’s alleged demise. The table below shows the number of hits for different search operators. Our dataset included 104 Supreme Court cases and many Federal Appellate decisions. For Supreme Court cases, we employed Westlaw and LexisNexis to build a preliminary sample of all the decisions in which the term “comity” appeared; we then proceeded to discard the ones in which the word was employed in a purely domestic context.³⁴ For Federal Court decisions, the same methodology could not be employed, as the results yielded by database searches were in a whole different order of magnitude – tens of thousands. Accordingly, we have elected to rely on a combination of Westlaw search operator, West key numbers, and frequency of citation in other decisions and scholarly pieces to identify a number of significant and influential rulings.

33 *Id.* at 78.

34 The search for “comity” and “international” was devised with the goal of avoiding searching for “international comity” only. While the search has yielded a number of false positives, it has revealed a number of cases in which instances of “international comity” had been labelled with the simple mention of “comity”.

	“Comity” (unrestricted)	“Comity” (1995-)	“Comity” (2005-)	“Inter- national Comity”	“Inter- national Comity” (1995-)	Inter- national Comity (2005-)
SCOTUS	639	115	50	44	9	5
CTA	7,159	3046	1425	472	300	163
District Court	10000+	10000+	10000+	1093	911	573
All Federal	10000+	10000+	10000+	1758	1295	782

	"Comity" & "International" (unrestricted)	"Comity" & "International" (1995-)	"Comity" & "International" (2005-)
SCOTUS	212	48	24
CTA	1995	1068	525
District Courts	7465	6010	4322
All Federal	10000+	7590	5147

True, these numbers, while telling, do not necessarily expose the whole picture. Indeed, in a very significant number of cases courts employ the term to refer to “other” comity doctrines, which are wholly domestic in nature and arise from the complex relationship between state and federal institutions that inform the law of the United States.³⁵ These doctrines are beyond the scope of the present article, which focuses instead on the uses of comity in cases where an

35 For a discussion of the extensive reliance on comity by American Courts see generally, Michael Wells, *The Role of Comity in the Law of Federal Courts*, 60 N.C. L. REV. 59 (1981) (discussing the Supreme Court's failure to distinguish between comity-mandated restraint and other scenarios, as well as the idea that the Court may use comity to shield its arbitrary assignment of some cases to federal courts and some to state courts without deciding which interest is stronger); Gil Seinfeld, *Reflections on Comity in the Law of American Federalism*, 90 NOTRE DAME L. REV. 1309 (2014) (discussing US courts' routine description of comity and the lack of reliance on the concept to explain duties that run from the states to the federal government).

“international” element is present – in which “cross-border elements are in need of careful treatment.”³⁶ It follows that this study will focus on federal decisions, much more likely to concern disputes in which the “cross-border elements” are truly international in nature. Such inevitable constraints, however, must not be perceived as a limitation, all the more so in a study concerned with the elucidation of a legal concept in a system where federal judges have often been identified as the main interpretive community.³⁷

III. The History of Comity in American Legal Thinking

A. *Comity Enters the United States*

1. *The Sister Notions of Comity and Sovereignty*

Comity, it has been observed, “is a concept with almost as many meanings as sovereignty.”³⁸ This comparison is not accidental: Comity developed as a corollary of the invention of the concept of sovereignty, and the history of the former is very much an account of the latter’s creation, only told from a different perspective.³⁹

The two narratives share the same starting point: the development of the Westphalian system. The consecration of the principles of territorial sovereignty and freedom from interference made personal statuses irrelevant in the face of the territorial law

36 Briggs, *supra* note 4, at 89. It could be objected that a domestic context should not automatically exclude the relevance of comity. This objection has some merit, but only insofar as a historical perspective is adopted: after all, much of the reflection on the topic was prompted by conflicts between the policies and legislations of American federate states. Story himself thought as much when he wrote that the issue of the application of foreign law was of paramount importance in the United States, “since the union of a national government with that of twenty-four distinct, and in some respects independent states, necessarily creates very complicated relations and rights between the citizens of those states, which call for the constant administration of extra-municipal principles.” STORY, *supra* note 14, § 9 at 9. While the current degree of integration is much greater today than it was in Story’s day, early reflection on the role of comity in this context is taken into account to properly explain the history of the principle.

37 But see Eben Moglen & Richard J. Pierce, Jr., *Sunstein’s New Canons: Choosing the Fictions of Statutory Interpretation*, 57 U. CHI. L. REV. 1203, 1220 (1990).

38 Anne-Marie Slaughter, *Court to Court*, AM. J. INT’L L. 708, 708 (1998).

39 KURT LIPSTEIN, *PRINCIPLES OF THE CONFLICT OF LAWS: NATIONAL AND INTERNATIONAL* 8 (1981); For a discussion of the historical evolution and the relationship between the concepts of comity and sovereignty, see generally Schultz & Holloway, *supra* note 1; Schultz & Ridi, *supra* note 4, at 10–11. See also Schultz & Holloway, *supra* note 1, at 571, 574.

of the state. Compared to the previous approach, epitomized by the likelihood of witnessing a meeting between five men with no one law in common, this no doubt represented a fundamental shift.⁴⁰ Sovereign independence and non-interference served as the building blocks of the new world order,⁴¹ but proved to be at variance with the transnational relations that formed the backbone of seventeenth-century European society and commerce.⁴² It comes as no surprise that the doctrine of comity, meant to mitigate the adverse effect of strict territoriality, would develop in the Netherlands, a country that had much to lose from this tension. Following their independence from Spanish rule and their unification, the Low Countries found themselves in need of coping with the effects of territorial sovereignty. On the one hand, a weak central power left the provinces' jealousy of their different traditions, laws, and original privileges relatively unbridled. On the other hand, the Netherlands had become a major maritime and commercial power, and were mindful of the need to accord a favorable treatment to foreigners.⁴³ This context offered the ideal combination of "difference and deference,"⁴⁴ thus providing fertile ground for the need of rules capable of governing the allocation of regulatory authority.⁴⁵

In the end, the question affected one fundamental problem, that of which law should govern a specific private legal relationship. It was in these circumstances that jurists such as Paulus Voet and, most notably, Ulrich Huber, developed new theories to provide an answer to the conundrum.⁴⁶ These scholars, however, were mindful of the general importance of the issue: As Rodolfo de Nova points

40 Hessel E. Yntema, *The Comity Doctrine*, 65 MICH. L. REV. 9, 10 (1966) (Recounting the complaint of Saint Agobard against the lex Gundobadi of Burgundy: "Nam plerumque contingit ut simul eant aut sedeant quinque homines et nullus eorum communem legem cum altero habeat... Indeed, it frequently happens that five men get together or meet with each other and none of them has a law in common with any other.").

41 See generally *id.* at 10.

42 See generally RODOLFO DE NOVA, HISTORICAL AND COMPARATIVE INTRODUCTION TO CONFLICT OF LAWS 441, 441–445 (1966); See generally Schultz & Ridi, *supra* note 4, at 10–11 (2017).

43 Yntema, *supra* note 40, at 19.

44 MILLS, *supra* note 17, at 45–46.

45 Yntema, *supra* note 40, at 17–19; Paul, *Comity in International Law*, *supra* note 9, at 15.

46 Arthur Nussbaum, *Rise and Decline of the Law-of-Nations Doctrine in the Conflict of Laws*, 42 COLUM. L. REV. 189, 190 (1942). See generally Yntema, *supra* note 40, at 22–27 (1966).

out, they did not seek to present their work as mere suggestions to Dutch courts, but rather as a new model with universal validity.⁴⁷

The doctrine of comity was introduced, in a fairly primitive form, by Paulus Voet in a work published in 1661.⁴⁸ The learned treatise, “couched in a concise and sometimes crabbed style in the form of a catechism of questions and answers,”⁴⁹ conceived comity as a technique for mitigating the adverse effects of the inherent territoriality of statutes. In other words, the doctrine allowed states to give effect and recognition — *ex comitate* — to transactions concluded outside the borders of the state concerned. Voet’s theory seems to suggest that such recognition was to be considered fully discretionary. What is more, comity, in this sense, did not “so much operate as a constraint on the (extraterritorial) application of a forum State’s law. . . but rather as a constraint on the exclusive territorial application of the forum State’s law.”⁵⁰ The point is worth mentioning because, almost forty years later, Paulus Voet’s son, Johannes, stressed the importance of the fact that the extension of statutes beyond the territorial domain of a state was not constrained by any particular rule.⁵¹ In his restatement of his father’s doctrine, comity served the purpose of preserving “the primacy of the statute real, subject to such concessions as might be made by one nation to another.”⁵² In other words, the doctrine proved essential for the reconciliation of two diverging paradigms, but its application remained — at least at first glance — fundamentally discretionary.⁵³

When the history of comity is so presented, one could be forgiven to read the work of the Dutch theorists as dismissing the old, universalist — or “international” — approach to questions of conflict of laws with a view to endorsing an attitude more consistent

47 DE NOVA, *supra* note 42, at 449.

48 PAULUS VOET, DE STATUTIS EORUMQUE CONCURSU LIBER SINGULARIS 156 (A.B. Edwards & D.M. Kriel eds., 1661).

49 Yntema, *supra* note 40, at 22.

50 RYNGAERT, *supra* note 17, at 139.

51 JOHANNES VOET, COMMENTARIUS AD PANDECTAS 30 (1778).

52 Yntema, *supra* note 40, at 24.

53 For a discussion of the tension of sovereignty and territoriality with that state of affairs, see Schultz & Holloway, *supra* note 1 (discussing that Jean Voet’s, following his father, “stresses the idea that the comity constitutes an exception to the principle of territorial sovereignty, the foreign countries being applied *ex comitatis* on the basis of considerations of common sense, without any legal rule requiring it [and comity] adopts a positivist approach to private and public international law, according to which it is for each State to determine how it intends to rescind conflict of laws.”).

with the model of territorial sovereignty. In fact, as early as 1966 De Nova denounced such a conclusion as injudicious.⁵⁴ The tenuousness of the discretion thesis appears even more evident by reading the *magnum opus* of the most influential writer of the time on the topic of comity, Ulrik Huber's *De conflictu legum*.⁵⁵ One of the leading jurists of his day, Huber too devoted significant attention to the topic of the application of foreign law. Most notably, Huber formulated an elegant solution by the use of three "axioms," the normativity of which remains a controversial issue in the modern debate on the doctrine of comity. He wrote:

- (1) The laws of each state have force within the limits of that government and bind all subject to it, but not beyond.
- (2) All persons within the limits of a government, whether they live there permanently or temporarily, are deemed to be subjects thereof.
- (3) Sovereigns will so act by way of comity that rights acquired within the limits of a government retain their force everywhere so far as they do not cause prejudice to the power or rights of such government or of its subjects.⁵⁶

The third axiom has long represented a controversial point:⁵⁷ What

54 DE NOVA, *supra* note 42, at 449–450.

55 The original text can be found in Ernest G. Lorenzen, *Huber's De Conflictu Legum*, in SELECTED ARTICLES ON THE CONFLICT OF LAWS 375, 403 (1947). It reads: "Rectores imperiorum id comiter agunt, ut jura cujusque populi intra terminos ejus exercita, teneant ubique suam vim, quatenus nihil potestati aut juri alterius imperantis ejusque civium praejudicetur." See also Ernest G. Lorenzen, *Story's Commentaries on the Conflict of Laws: One Hundred Years After*, 48 HARV. L. REV. 15, 16 (1934). On Huber's solution see also Paul, *The Transformation of International Comity*, *supra* note 9, at 22–23; Schultz & Holloway, *supra* note 1, at 580.

56 Lorenzen, *Huber's De Conflictu Legum*, *supra* note 55, at 403.

57 See the discussion in Llewelyn Davies, *The Influence of Huber's De Conflictu Legum on English Private International Law*, 18 BRIT. Y.B. INT'L. L. 49, 56–59 (1937) (stating Huber's third axiom was generally accepted in England and the US, but was not "generally accepted among continental jurist", and was misunderstood by Dutch jurist.); See also Lorenzen, *Huber's De Conflictu Legum*, *supra* note 55, at 394–99 (describing the extent to which U.S. courts will decline to apply foreign law to validate polygamous marriages because such recognition would "prejudice vital interests"); See also Yntema, *supra* note 40 at 30 (Showing how Huber's approach to the issue was informed by a Grotian understanding of the *ius gentium* and thus resulted in the doctrine of comity).

did Huber actually mean? What kind of discretion did he envisage for a sovereign and its courts? What was the nature and cogency of the obligation — provided that there was any — to apply foreign law? Huber's work does not, in and by itself, provide simple answers to these hoary questions: But were they really appropriate to begin with? Two main points suggest otherwise.

First of all, Huber never employed the word *comitas*, let alone the expression *comitas gentium* ('the comity of nations'). Huber did, however, restrict himself to choosing an adverbial—and possibly less charged—form, *comiter*, which—it is worth recalling—had been adopted by Paulus Voet.⁵⁸ The language adopted in the Dutch edition does not even mention any such term: Acutely aware of the importance of ensuring the smoothness of business transactions, Huber supports the discussion with the allegorical image of governments extending a hand to each other.⁵⁹ While linguistic exegesis can only play a small role in legal scholarship, this small detail may contribute to the clarification of the original doctrine. In particular, it allows the concept of comity to be revealed as the cornerstone of the building, rather than the building itself and as a whole. The point was lucidly highlighted by Campbell McLachlan:

Huber, Voet and those the jurists that followed them . . . used the concept of comity as a springboard from which they proceeded to develop a highly organized and sophisticated set of choice of law rules. In this sense, "comity" did not remain a vague desideratum – an invitation to replace law with its antithesis in mere courtesy and discretion. On the contrary, it supplied the basis for the elaboration of a detailed set

having the features of an international usage); *See also ALAN WATSON, JOSEPH STORY AND THE COMITY OF ERRORS: A CASE STUDY IN CONFLICT OF LAWS* 28 (1992).

⁵⁸ See, e.g., VOET, *supra* note 51, at 143, 168. While Huber's work was published before Johannes Voet's *Commentarius*, he was undoubtedly familiar with the work of Paulus. Voet and Huber are traditionally regarded as representatives of the *doctrine hollandaise*' and their works were published only a few years apart. As Huber was one of the leading Dutch jurists of his generations, it may be inferred that he must have been familiar with the work of Paulus Voet. *See* Yntema, *supra* note 40, at 24–25, 29.

⁵⁹ Schultz & Holloway, *supra* note 1, at 580 ("Huber borrowing in his writings in Dutch the metaphor of states extending the hand reciprocally ... rests according to Huber on the principle of the comity").

of positive rules, grounded in practical reality.⁶⁰

Indeed, the system at issue could not be faulted as primitive: The Dutch jurists *did* develop fairly advanced rules, which we might, with a modern mind-set, define as rules of private international law. As is well known, due to the rise of positivism, the whole field of conflict was being drawn into the preserve of the national legislator.⁶¹ And yet, the understanding of comity outlined above allows the reconciliation of its value with the idea of territoriality and the doctrine of supremacy of sovereign command.⁶² Comity, in other words, can be considered the conceptual basis of the rules, rather than their formal source. The modern idea of the doctrine as necessitating exercises of discretion by a court does not follow directly from the essence of the principle.

The second point follows from the observation that Huber's conception of international law was fundamentally a Grotian one.⁶³ His third axiom spells out an international usage—if not an international custom—whereby “the effects of competent foreign laws are everywhere admitted, except when prejudicial to the forum State or its citizens, through the reciprocal indulgence of the sovereign authorities in each State.”⁶⁴ This is no small detail: on the one hand, reliance on the *jus gentium* allowed Huber to universalize his maxims; on the other hand, it lends credibility to the view that Huber's axioms were, at least to some extent, descriptive of then-current practices, and anticipative of their consolidation—which would have, in turn, yielded normative consequences—rather than simply prescriptive.

In this respect, remarks such as Dicey's well-known discussion on the application of foreign law as having little to do with courtesy between sovereigns fail to make the grade as a subtle critique of Huber.⁶⁵ Indeed, convenience and logic, rather than

60 CAMPBELL McLACHLAN, LIS PENDENS IN INTERNATIONAL LITIGATION 32 (2009) (emphasis added).

61 Schultz & Holloway, *supra* note 1, at 593.

62 The equation of law with a sovereign command found its finest jurisprudential expression in the work of Austin. See generally JOHN AUSTIN, THE PROVINCE OF JURISPRUDENCE DETERMINED (1832).

63 There is no question that Huber was familiar with Grotius, whom he cites in *De Conflictu Legum* too. See Lorenzen, *Story's Commentaries on the Conflict of Laws*, *supra* note 55, at 165. See also Schultz & Holloway, *supra* note 1, at 578.

64 Yntema, *supra* note 40, at 27, 30.

65 ALBERT VENN DICEY, A DIGEST OF THE LAW OF ENGLAND WITH

mere courtesy, were the goals with which the Dutch theorist was concerned—very much the same objectives that Dicey saw as best pursued without relying on his characteristic theoretical construct.⁶⁶ Needless to say, this was not the way Huber's ideas were received in the common law world, where his writings eventually made an impact, due to unique circumstances, on the minds of students and practitioners of the law.⁶⁷ The scholar's name became a shorthand to make a point on the supremacy of the forum's law, in a corruption of the doctrine that was to a great extent a corruption of his legacy.⁶⁸

2. Comity in the United States: Livermore, Kent and Story

When the concept is conceived as the basis of certain rules, Samuel Livermore's famous description of the expression "comity" as something "grating to the ear when it proceeds from a court of law" seems quite unfair and one-sided. "Comity between nations," Livermore wrote, "is to be exercised by those who administer the supreme power. The duty of judges is to administer justice according to law, and decide between parties' litigant according to their rights." And yet, Livermore and his passionate discussion of a point of law he regarded as fundamental played no small part in prompting a reconsideration of the concept of comity and the Dutch theories. But how did this happen?

According to Alan Watson, much can be traced back to the anger that Livermore, an influential attorney of civil-law Louisiana, felt towards the Anglo-American reliance on the Dutch jurists, and to his personal resentment against the judge who authored an influential opinion on conflict in a case where he happened to find himself on the losing side.⁶⁹ It has been argued that Livermore rejected the idea of comity, which he saw as allowing excessive discretion—a discretion, in other words, that international law did not allow.⁷⁰ Quite interestingly, he appeared rather ignorant and

REFERENCE TO THE CONFLICT OF LAWS 10 (London, Stevens & Sons, Ltd. and Sweet & Maxwell, Ltd. 1896).

66 See *id.* at 10, 11.

67 Davies, *supra* note 57, at 53–54; Schultz & Holloway, *supra* note 1, at 591.

68 See, e.g., WATSON, *supra* note 57 (suggesting that Story's understanding of the concept of comity resulted from a misunderstanding of Huber's writings).

69 WATSON, *supra* note 57, at 28–33; Saul v. His Creditors, 5 Mart. 569 (La. 1827).

70 Paul, *The Transformation of International Comity*, *supra* note 9, at 24. The expression in question is "something like an obligation upon sovereigns." SAMUEL LIVERMORE, DISSERTATIONS ON THE QUESTIONS WHICH

disdainful of Huber, whose authority may well have supported his thesis.⁷¹ Whether Livermore's view really was based on international law properly-so-called is at best doubtful: In his oft-cited *Dissertations*, he motivated the necessity of applying foreign law by reference to "a sense of mutual utility," "arisen from a sort of necessity" and "the inconveniences, which would result from a contrary doctrine."⁷² There is no question, however, that Livermore believed that the application of foreign law was conducive to maintaining peace, friendly intercourse, and even to "the general good," in that it would have pursued in the private sphere the same objective that the law of nations pursued in the public one.⁷³

This treatment of comity and of the Dutch jurists' theories was perhaps unsophisticated, but proved influential. True, it was only with Joseph Story's work that comity eventually became an important element of the interface between public international law and American conflict of laws. However, Livermore's treatise, along with the case that prompted its creation, was quoted in the influential work of James Kent,⁷⁴ to whom Joseph Story felt intellectually indebted.⁷⁵ And indeed, not unlike Kent, the learned jurist felt that the issue was one of central importance for the management of frictions resulting from radically different state policies.⁷⁶

Story insisted that no national law could have, in principle,

ARISE FROM THE CONTRARIETY OF THE POSITIVE LAWS OF DIFFERENT STATES AND NATIONS 30 (New Orleans, Benjamin Levy 1828).

71 WATSON, *supra* note 57, at 32–33.

72 LIVERMORE, *supra* note 70, at 28. Compare with the language adopted by STORY, *supra* note 14, at § 35.

73 LIVERMORE, *supra* note 70, at 30. In this regard, the theory is not too far away from Minor's famous reference to comity as the basis of international law. RALEIGH C. MINOR, *CONFLICT OF LAWS, OR, PRIVATE INTERNATIONAL LAW* 5 (1901).

74 2 JAMES KENT, *Of the Contract of Sale, in COMMENTARIES ON AMERICAN LAW* 363 (NEW YORK, O. HALSTEAD 1827). In reality, Kent attributed the work to Livermore with the title "Dissertations on Personal and Real Statutes," a corruption of the title of an earlier work, by the British scholar Henry, titled "Dissertation on Personal, Real and Mixed Statutes." Story, however, cites both authorities correctly. STORY, *supra* note 14, at 33.

75 John B. Cassoday, *James Kent and Joseph Story*, 12 YALE L.J. 146, 146 (1903).

76 See, e.g., STORY, *supra* note 14, at § 9 ("To no part of the world is it of more interest and importance than to the United States, since the union of a national government with that of twenty-four distinct, and in some respects independent states, necessarily creates very complicated relations and rights between the citizens of those states, which call for the constant administration of extra-municipal principles.").

extraterritorial effect: In his view, this conclusion was the natural consequence of public law principles, and was also supported by Vattel's writings on the sovereign equality of nations in the field of public international law.⁷⁷ "This branch of public law," he wrote in his *Commentaries*, "may be fitly denominated private international law."⁷⁸ But Story's adherence to this view should not be mistaken for a belief that the application of foreign law descended from some perfect international obligation—in fact, his reliance on Huber's theories, which may well have supported very different conclusions, has been described as the result of a misunderstanding.⁷⁹ Indeed, it is often argued that, to the Dutch theorist, the application of foreign law had a wholly different degree of bindingness.⁸⁰ To the American jurist, the state retained the fundamental ability to decide whether to give effect to foreign law: There was no duty to do so, but more of an "imperfect obligation, like that of beneficence, humanity, and charity."⁸¹ "Every nation," he continued, "must be the final judge for itself, not only of the nature and extent of the duty, but

77 STORY, *supra* note 14, at § 8. Consider the influence of this dictum in *Hilton v. Guyot*: "International law, in its widest and most comprehensive sense—including not only questions of right between nations, governed by what has been appropriately called the law of nations; but also questions arising under what is usually called private international law, or the conflict of laws, and concerning the rights of persons within the territory and dominion of one nation, by reason of acts, private or public, done within the dominions of another nation—is part of our law, and must be ascertained and administered by the courts of justice, as often as such questions are presented in litigation between man and man, duly submitted to their determination." 159 U.S. 113, 163 (1895).

78 STORY, *supra* note 14, at § 9. A citation of Vattel and Huber in the same decision might appear surprising to the modern reader. It was, however, hardly surprising at the time, nor was it unprecedented. For example, the authority of both Huber and Vattel was cited in *Desebats v. Berquier*, an 1808 Pennsylvania conflicts case relating to personal property under a will. Indeed, this constitutes another indication of the perceived contiguity—or, more precisely, of the absence of any real separation—of public and private international law. See WATSON, *supra* note 57, at 51. For an analysis of the issues concerning the questions concerning the unity of private and public international law see MILLS, *supra* note 17.

79 For an in-depth analysis of the issue, see WATSON, *supra* note 57.

80 See Yntema, *supra* note 40, at 29 (arguing that Huber saw the extraterritorial extension of statutes as mandated by the *jus gentium*); Davies, *supra* note 57, at 57. For a powerful rejection of the idea whereby comity should be the source of a perfect obligation to apply foreign law, see Calamita, *supra* note 22, at 617.

81 STORY, *supra* note 14, at § 33.

of the occasions, on which its exercise may be justly demanded.”⁸² An imperfect obligation indeed, which derived, in the words of Samuel Livermore, “from a sense of mutual utility . . . a sense of the inconveniences which would result from a contrary doctrine.”⁸³

The centrality of state discretion in Story’s treatment of comity is the crucial issue—indeed, one that makes Watson’s claim that he did, to some extent, misinterpret Huber seem not too harsh an assessment.⁸⁴ The point is not the claim that the predominance of Story’s authority over Huber’s altered the outcome of cases such as the infamous *Dred Scott v. Sandford*⁸⁵ due to the possibility—consistent with the former’s theory, but not with the latter’s—of exercising a choice as to which law to apply.⁸⁶ In fact, the point is the very centrality of the concept of discretion in Story’s theory. As Joel Paul observed, Story’s solution was an exceptionally elegant one, as it managed to universalize conflict and, at the same time, reflect his policy concerns by affirming the primacy of the forum’s law.⁸⁷ What is more, this thesis was in line with the rise of positivism—a characteristic that helped it survive and further develop in the American setting.⁸⁸

Given the popularity of his writings and the impact that they had on subsequent cases, a discussion of Story’s engagement with the idea of comity has traditionally represented the ideal *dénouement* of historical analysis of the concept—virtually all the studies cited in this article prove as much. There are good reasons to respect the tradition: upon publication of the *Commentaries*, Story remained an

82 STORY, *supra* note 14, at § 33.

83 LIVERMORE, *supra* note 70, at 28. See also STORY, *supra* note 14, at § 35.

84 On the reasons why a misunderstanding seems a more viable hypothesis than a deliberate, if subtle, rejection, see WATSON, *supra* note 57, at 71–72.

85 *Scott v. Sandford*, 60 U.S. (19 How.) 393 (1857).

86 This conclusion is grounded on Section 3 of Huber’s *De Conflictu Legum*, where it is said that “all transactions and acts, in court as well as out, whether mortis causa or inter vivos, rightly done according to the law of any particular place, are valid even where a different law prevails, and where, had they been so done, they would not have been valid.” Lorenzen, *Huber’s De Conflictu Legum*, *supra* note 55, at 404. Watson also observes that in case concerning slavery, the law of a slave-owning state could not be rejected on the basis of its being revolting, as Huber’s third axiom would be intimately linked to the law of nations, which ostensibly did not yet contain a prohibition to that effect. WATSON, *supra* note 57, at 64. It is arguable that this inference amounts to reading too much into Section 8 of *De Conflictu Legum*.

87 Paul, *The Transformation of International Comity*, *supra* note 9, at 25 (The word “elegantly” is used to describe Story’s theory).

88 Paul, *The Transformation of International Comity*, *supra* note 9, at 25.

important authority on the issue of comity, and his legacy was—at least formally—very much alive one century and a half later.⁸⁹ At the same time, it must be stressed that it is unthinkable to resolve cases involving comity considerations with the simple application of Story’s maxims. Comity—not unlike its parent concept, sovereignty—is inherently context-dependent. Its meaning and implications, as we seek to elucidate in the next sections, transformed as a reflection of the changes in the status quo. Such a shift is investigated in the following sections, under the lens provided by almost two centuries of judicial decision-making.

IV. The Judicial Evolution of Comity

A. “Legislative” or “Prescriptive” Comity: Restraint and Recognition

The history of comity, as we have seen above, clearly shows that the principle was borne out of the need to make sense of a new model of allocation of regulatory authority. With a central role played by the principle of territoriality, coupled with the need to cater to the needs of commerce and friendly relations, comity proved successful as a flexible mediating principle.

The idea of “flexibility” has long been attached to comity, but the term is somewhat misleading: Indeed, as Adrian Briggs has argued, it suggests an understanding of comity more akin to “the way in which the common law reaches a conclusion which its rules do not otherwise allow it to find,” which is regrettable.⁹⁰ Rather, comity, is “flexible” because it takes different shapes depending on the goals states need to accomplish. Consider the case of what is generally termed “legislative comity”: On the one hand, the principle mandates—or, at the very least—justifies the recognition of foreign law, in deference to the regulatory power of a foreign sovereign. On the other hand, and at the same time, it actively limits the reach of the law of the state acting according to comity, in an effort to avoid unreasonable interferences with the counterpart’s regulatory power. According to William Dodge, these are but two sides of the same coin:⁹¹ in other words, recognition and restraint are two faces of “prescriptive comity.”⁹² To be sure, this conclusion

89 See the citation of the *Commentaries* in Justice Scalia’s dissent in *Hartford Fire Ins. Co. v. California*, 509 U.S. 764, 817 (1993).

90 Briggs, *supra* note 4, at 87.

91 Dodge, *supra* note 2, at 2079.

92 The expression first appeared in Justice Scalia’s dissent in *Hartford Fire Ins. Co.*

appears sound. What is more, it appears to be a logical corollary of the most authoritative statement of the doctrine, Justice Gray's majority opinion in *Hilton v. Guyot*, where it was famously affirmed that “[n]o law has any effect, of its own force, beyond the limits of the sovereignty from which its authority is derived.”⁹³

It is possible to subscribe to this conclusion, but it must be recognized that it is premised on the idea that recognizing foreign law and limiting the state's own law are subject to different outer limits: In particular, the international law of jurisdiction plays a major role in the latter, but only a very minor one in the former.⁹⁴ What is more, the interplay between comity and international law—at best, a reasonably complex matter—is rendered particularly puzzling by the adoption of the adjective “prescriptive,” which, while appropriate to describe their role in this area, tends to add to

v. California, 509 U.S. 764, 817 (1993) (distinguishing “perspective comity” from “comity of the courts”, as a recognition of another sovereign’s laws and the legislative restraint to purposely limit the reach of a country’s own laws). The adjective “prescriptive” is generally used to refer to the concept of “jurisdiction to prescribe.” See ANDREAS F. LOWENFELD, PUBLIC LAW IN THE INTERNATIONAL AREA: CONFLICT OF LAWS, INTERNATIONAL LAW AND SOME SUGGESTIONS FOR THEIR INTERACTION 400 (1980).

93 Hilton v. Guyot 159 U.S. 113, 163 (1895). For a similar phrasing in a prior decision, see *Crapo v. Kelly*, 83 U.S. 610, 642–43 (1872) (“According to my view, whilst the disposition of his movable property by the owner is respected by the laws of all States everywhere, the laws of any particular State and transfers by operation of law, have no extra-territorial force which other States will concede, except by comity.”) (Bradley, J. dissenting). On the status of *Hilton* as the “classic statement” of the doctrine, see also Paul, *The Transformation of International Comity*, *supra* note 9, at 27 (describing the *Hilton* decision as the “classic statement” of the doctrine of comity).

94 We are restating the current understanding of the operation of the rules of public international law in this context, not claiming that such rules are inherently incapable of governing the recognition of foreign law. This is clearly not the case: Not only do a number of international instruments deal specifically with these matters, but it was observed as early as the 1930s that there may be cases in which the private international law requiring the recognition of foreign law would be overridden by the application of rules of public international law. See Alexander P. Fachiri, *Recognition of Foreign Laws by Municipal Courts*, 12 BRIT YB INTL L 95, 103 (1931). Further, and while a discussion of this issue is beyond the scope of the present study, certain rules of private international law may be—and, in fact, might have been—accorded the status of “general principles of law” in the sense of Article 38(1)(c) of the Status of the International Court of Justice and its predecessor, the PCIJ Statute. See *Payment of Various Serbian Loans Issued in France (Fr. v. Serb. & Croat. & Slovn.)*, Judgment, 1929 P.C.I.J. (ser. A) No. 14, at 44 (July 12). See also MILLS, *supra* note 17, at 214, 230.

the confusion. Finally, not unlike in the case of adjudicatory comity, the Restatement (Third) of Foreign Relations Law provides guidance in cases concerning the extraterritorial application of United States law, but sheds little light on the matter of recognition of foreign acts. Reasons of coherence and homage to tradition suggest that the two issues be treated under the same heading, but due consideration should be paid to the dissimilarities discussed above.

1. Using Comity to Grant Recognition to the Law and the Acts of Other States

While *Hilton v. Guyot* remains the most well-known decision on recognition in general, Statements of the Supreme Court invoking comity to allow the application of foreign law and acts predate it by almost one full century. As early as 1797 the Supreme Court had stated in *Emory v. Grenough* that:

[T]he laws of one commonwealth, cannot affect the integral parts, the territory of another commonwealth . . . it cannot be done by the immediate force and operation of a foreign law, but with the concurring consent of the supreme power in the other government, which gives an effect to foreign laws exercised upon property within its own jurisdiction, without any prejudice being received to its sovereignty or the rights of its citizens, regarding the mutual convenience of the two nations or governments, which is the foundation of all these rules . . . that it is not so much by force of law, as by the consent of the parties reciprocally communicating their rights to each other, by which means a change, or modification of property may arise, not less from matrimony than any other contract.⁹⁵

The decision, which was not immediately followed by other Supreme Court pronouncements invoking comity in the context of the recognition of foreign law, contains a statement of the doctrine almost as clear as the one contained in *Hilton*—indeed, all the elements of a traditional comity analysis are present.

95 *Emory v. Grenough* 3 U.S. (3 Dall.) 369, 374 (1797). The case contains an extract of Huber's *De Conflictu Legum* by Alexander Dallas. See Dodge, *supra* note 2, at 2086–87.

It was Justice Taney's 1839 opinion in the domestic case *Bank of Augusta v. Earle*, however, that brought the discussion of comity even closer to Story's theories. The decision which has been called "the original fountain head of the law of foreign corporations in America"⁹⁶ and had its drafter privately applauded by Story himself,⁹⁷ referred to the *Commentaries* to strengthen the proposition that foreign companies could, lacking an express prohibition, make business in another state.⁹⁸ Most importantly, it linked the notions of comity and sovereignty in the most principled manner to date:

The comity thus extended to other nations is no impeachment of sovereignty. It is the voluntary act of the nation by which it is offered; and is inadmissible when contrary to its policy, or prejudicial to its interests. But it contributes so largely to promote justice between individuals and to produce a friendly intercourse between the sovereignties to which they belong; that Courts of justice have continually acted upon it, as a part of the voluntary law of nations.⁹⁹

While *Earle* was a domestic case, its implications were clearly visible in *Canada Southern Railway. Co. v. Gebhard*.¹⁰⁰ This bankruptcy case concerned a Canadian railway company that was reorganized through a plan agreed upon by the majority creditors and the Canadian Parliament, who then passed a statute to bind the minority creditors. In a suit brought by an American resident, and faced with the question of giving effect to the Canadian statute, the Supreme Court reasoned, predictably, that "the laws of a country have no extra-territorial force is an axiom of international jurisprudence, but

96 GERARD CARL HENDERSON, THE POSITION OF FOREIGN CORPORATIONS IN AMERICAN CONSTITUTIONAL LAW 42 (1918).

97 Charles Grove Haines & Foster H. Sherwood, The Role of the Supreme Court in American Government and Politics 1835–1864, at 75 n.68 (1957).

98 "The Court can perceive no sufficient reason for excluding from the protection of the law the contracts of foreign corporations; when they are not contrary to the known policy of the state, or injurious to its interests... It is but the usual comity of recognising the law of another state. The states of the Union are sovereign states; and the history of the past and the events which are daily occurring, furnish the strongest evidence that they have adopted towards each other the laws of comity in their fullest extent." *Bank of Augusta v. Earle*, 38 U.S. (13 Pet.) 519, 520 (1839).

99 *Id.* at 589 *passim*.

100 109 U.S. 527 (1883).

things done in one country under the authority of law may be of binding effect in another country.”¹⁰¹ It followed that “every person who deals with a foreign corporation impliedly subjects himself to such laws of the foreign government[.]”¹⁰² Justice Harlan, however, dissented, arguing for a more nuanced analysis. In his view, comity was not sufficient reason to enable a foreign corporation to “benefit, in our courts—to the prejudice of our own people and in violation of their contract and property rights—of a foreign statute which could not be sustained had it been enacted by Congress or by any one of the United States[.]”¹⁰³ But was it really the case? Indeed, the proposition is particularly interesting insofar it also echoes the domestic case *Ogden v. Saunders*,¹⁰⁴ in which Chief Justice Marshall authored a powerful dissenting opinion—joined by none other than Joseph Story—arguing that the “single question for consideration” was whether the law of the act of discharge was “consistent with or repugnant” to the constitution of the United States.¹⁰⁵

Contrary to the enforcement of foreign judgments, it does not appear that reciprocity has ever had a significant role to play with regard to the recognition of foreign law in American courts. It must be pointed out, however, that such matters were, to a large extent, removed from this sphere by the Court’s decision in *Erie R.R. v. Tompkins*¹⁰⁶ and *Klaxon Co. v. Stentor Electric Mfg. Co.*,¹⁰⁷ which put back such private international law questions firmly within the domain of state law.¹⁰⁸

101 *Id.* at 536.

102 *Id.* at 537.

103 *Id.* at 539.

104 *Ogden v. Saunders*, 25 U.S. (12 Wheat.) 213 (1827).

105 *Id.* at 332. For an earlier comment, see also *Recognition of Foreign Bankruptcy Discharges*, 47 YALE L.J. 1020 (1938).

106 *Erie R.R. v. Tompkins*, 304 U.S. 64, 78 (1938) (“Except in matters governed by the Federal Constitution or by acts of Congress, the law to be applied in any case is the law of the State. And whether the law of the State shall be declared by its Legislature in a statute or by its highest court in a decision is not a matter of federal concern. There is no federal general common law.”).

107 313 U.S. 487, 496 (1941) (stating that “the prohibition declared in *Erie Railroad v. Tompkins* against such independent determinations by the federal courts extends to the field of conflict of laws”).

108 On the implications of *Erie* on private and public international law *see, inter alia*, Philip C. Jessup, *The Doctrine of Erie Railroad v. Tompkins Applied to International Law*, 33 AM. J. INT’L L. 740, 740–43 (1939); *see also* Donald Earl Childress III, *When Erie Goes International*, 105 NORTHWEST. UNIV. L. REV. 1156 (2011), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1691799 (last visited Apr 3, 2016), linking the Court’s approach in *Klaxon* “to the fact that Justice

2. Using Comity to Limit the Reach of American Law

a. Antitrust

At the turn of the millennium, Spencer Weber Waller penned a powerful analytical article in which he declared that comity had entered its twilight hours.¹⁰⁹ The principle, he said, was no longer important because “its advocates” had, at the same time, achieved a success and suffered a loss. On the one hand, the demands of comity were now more often satisfied by the increasingly cautious attitude of the United States in dealing with foreign interests; on the other hand, engaging in a comity analysis still fell severely short of being a requirement for courts tasked with cross-border cases.¹¹⁰ Most importantly, the principle had ceased to matter because the United States was no longer “the world’s antitrust policeman”: to Waller, comity had been a valuable tool, but the time was ripe to “say goodnight to an old friend.”¹¹¹

Waller has recognized a number of crucial characteristics of comity; for instance, that it is the result of historical contingency and that it is not a prerogative of courts only. While his article admittedly focused on antitrust issues, it does not seem apparent from even a cursory review of American case law that international comity did in fact come upon its demise.

To understand the point, it is helpful to start from the beginning. For a long time, comity has played a crucial role in limiting the reach of United States law. It has done so by acting as an upper limit to the exercise of jurisdiction or an interpretive canon capable of making sense of ambiguous statutes and treaties,¹¹² counselling

Story’s comity doctrine had been replaced by the time of Erie with Joseph Beale’s “vested rights” approach, which was influential during the first half of the twentieth century.”

109 Spencer Weber Waller, *The Twilight of Comity*, 38 Colum. J. Transnat'l. Law 563 (1999).

110 *Id.* at 565.

111 *Id.* at 566. For a partially diverging assessment, see KOH, *supra* note 27, at 59.

112 An early example is provided by the *Wildenhus's Case* 120 U.S. 1, 12 (1887), a case dealing with a homicide committed aboard a ship flying the Belgian flag. A treaty between the United States and Belgium stated that the jurisdiction of local courts should have been excluded for such offences—“except when a disorder arises of such a nature as to disturb tranquillity and public order on shore, or in the port, or when a person of the country or not belonging to the crew shall be concerned therein.” Convention between the United States and Belgium, concerning the rights, privileges and immunities of consular

restraint or a degree of intrusiveness depending on the interests at issue and the context.

While previous cases might be read to the same effect, the first instance of the use of comity to control the extraterritorial effect of American law is traditionally found in *American Banana Co. v. United Fruit Co.*¹¹³ An antitrust case, the dispute was noteworthy in that it concerned acts committed in Costa Rica and Panama in violation of American antitrust law. Delivering the opinion of the Court, Justice Holmes observed that the case was built on several “rather startling propositions,” chief among them “[the fact] the acts causing the damage were done, so far as appears, outside the jurisdiction of the United States, and within that of other states”; it was thus “surprising to hear it argued that they were governed by the act of Congress.”¹¹⁴ Citing an English landmark case,¹¹⁵ Holmes concluded that “[a]ll legislation [was] *prima facie* territorial” and with good reason too, for concluding otherwise—here, disregarding the role of local law in governing the legality of the acts—would have been unjust and “an interference with the authority of another sovereign, contrary to the comity of nations, which the other state concerned justly might resent.”¹¹⁶

officers. Belg.-U.S., art. XI, proclaimed Mar. 1, 1881, 21 Stat. 776, 785. Faced with the question of whether this was the case, the Supreme Court discussed the rationale of this provision: “And so by comity it came to be generally understood among civilized nations that all matters of discipline, and all things done on board, which affected only the vessel, or those belonging to her, and did not involve the peace or dignity of the country, or the tranquillity of the port, should be left by the local government to be dealt with by the authorities of the nation to which the vessel belonged as the laws of that nation, or the interests of its commerce should require. But, if crimes are committed on board of a character to disturb the peace and tranquillity of the country to which the vessel has been brought, the offenders have never, by comity or usage, been entitled to any exemption from the operation of the local laws for their punishment, if the local tribunals see fit to assert their authority. Such being the general public law on this subject, treaties and conventions have been entered into by nations having commercial intercourse, the purpose of which was to settle and define the rights and duties of the contracting parties with respect to each other in these particulars, and thus prevent the inconvenience that might arise from attempts to exercise conflicting jurisdictions.” *Id.* at 12.

¹¹³ *Am. Banana Co. v. United Fruit Co.*, 213 U.S. 347 (1909).

¹¹⁴ *Id.* at 355.

¹¹⁵ *Ex Parte Blain* (1879) 12 Ch D 522 at 528.

¹¹⁶ 213 U.S. at 356 (citing *Phillips v. Eyre* (1870) LRQB 225 at 239. It must be observed that this English case introduced a two-limbed test of “double actionability” (or “criminality”), which was applied until a different rule was adopted in *Boys v. Chaplin* [1971] AC 356. According to Koh, *American Banana*

American Banana relies heavily on comity—the idea is pervasive, if the invocation scarce. Its lesson is that legislation is presumed to be territorial, though there may be cases in which it is not, and other ones yet where no obstacle to regulation exists because there is simply no sovereign to offend.¹¹⁷ Its legacy, on the other hand, was relatively short-lived.¹¹⁸ The changed circumstances of transnational commerce—and perhaps the awareness of a different attitude towards jurisdiction in international law in the wake of the *S.S. Lotus* decision by the Permanent Court of International Justice,¹¹⁹ which embodies the belief that states have a wide measure of discretion in determining the limits of their jurisdiction—brought the Court of Appeals of the Second Circuit (the court of last resort as indicated by the Supreme Court) to partially reverse this approach in *United States v. Aluminum Corporation of America* (“*Alcoa*”).¹²⁰ While other decisions had progressively moved away from *American Banana*,¹²¹ the opinion was a true turning point. Writing for the

is as much a result of considerations of comity as it is of the application of rules of conflict, as the acts in question were not prohibited in the states where they had been committed and did not therefore satisfy the English test. See KOH, *supra* note 27, at 59.

- 117 213 U.S. at 356. See also Stephen D. Piraino, *A Prescription for Excess: Using Prescriptive Comity to Limit the Extraterritorial Reach of the Sherman Act*, 40 HOFSTRA L. REV. 1099, 1103 (2012).
- 118 The Supreme Court held in *Kirkpatrick* that *American Banana* had been “substantially overruled” by *Continental Ore Co. v. Union Carbide & Carbon Corp.*, 370 U.S. 690, 704 (1962).
- 119 S.S. *Lotus* (Fr. v. Turk.), Judgment, 1927 P.C.I.J. (ser. A) No. 10, at 253 (Sept. 7). See also KOH, *supra* note 27, at 59–60.
- 120 *United States v. Aluminum Co. of Am.*, 148 F.2d 416, 443–44 (2d Cir. 1945) (stating that “[o]n the other hand, it is settled law [...] that any state may impose liabilities, even upon persons not within its allegiance, for conduct outside its borders that has consequences within its borders which the state reprobates; and these liabilities other states will ordinarily recognize.”).
- 121 On the extraterritorial application of US antitrust law see *United States v. Pac. & Arctic Ry. & Navigation*, 228 U.S. 87, 98–99 (1913) (holding that foreign parties’ actions violating US antitrust law and effecting commerce within the US were in fact subject to US antitrust law despite their foreign location, and stating that “[i]f we may not control foreign citizens or corporations operating in foreign territory, we certainly may control such citizens and corporations operating in our territory, as we undoubtedly may control our own citizens and our own corporations”); *United States v. Sisal Sales Corp.*, 274 U.S. 268, 275 (1927) (distinguishing *American Banana* on the basis that the parties in *Sisal* entered into a contract and conducted their acts within the US, whereas the parties in *American Banana* and the violative actions occurred outside of the US).

Court, Judge Learned Hand affirmed that “any state may impose liabilities, even upon persons not within its allegiance, for conduct outside its borders that has consequences within its borders which the state reprobates; and these liabilities other states will ordinarily recognize.”¹²² The decision does not mention comity in the least—though its echoes can perhaps be perceived in the discussion of the issue of extraterritoriality¹²³—and relied upon a test allowing the application of American antitrust law to be triggered by acts committed abroad, provided that they would have been unlawful if committed in the United States and were intended to and actually did cause—an anticompetitive effect.¹²⁴ Remarkably, the decision managed to pay lip service to *American Banana*. Half a century on, a number of things can be traced back to this seminal case: the ideas of effects doctrine and protective principle as discussed in the Restatement; the “nationalist jurisprudence” that culminated in *Hartford Fire*; and above all, the growing willingness of American courts to interfere. And yet, the significance of comity goes hand in hand with the potential for interference warranted by grounds of jurisdiction that can be understood as “virtually unbounded in scope.”¹²⁵

While early on in certain areas the presumption against the extraterritorial effect of American law continued to operate in the traditional fashion,¹²⁶ antitrust cases remained fundamentally different. The new approach was codified in the Ninth Circuit case *Timberlane Lumber Co. v. Bank of America*.¹²⁷ The case generated a great deal of controversy and academic commentary, but for our purposes it is sufficient to mention that it introduced a tripartite analysis

122 *Aluminum Co. of Am.*, 148 F.2d at 443.

123 *Id.* (“Nevertheless, it is quite true that we are not to read general words, such as those in this Act, without regard to the limitations customarily observed by nations upon the exercise of their powers.”). The statement was quoted to this effect in *Timberlane Lumber Co. v. Bank of America*, N.T. & S.A., 549 F.2d 597, 613 (9th Cir. 1976). See also Piraino, *supra* note 118, at 1105–06.

124 *Aluminum Co. of Am.*, 148 F.2d at 444 (stating that company agreements that were made abroad and were in violation of US antitrust law “would clearly have been unlawful, had they been made within the United States; and it follows from what we have just said that both were unlawful, though made abroad, if they were intended to affect imports and did affect them”).

125 Jeffrey A. Meyer, *Dual Illegality and Geoambiguous Law: A New Rule for Extraterritorial Application of U.S. Law*, 95 MINN. L. REV. 110, 151 (2010).

126 See, e.g., *Lauritzen v. Larsen*, 345 U.S. 571, 589–91 (1953); see also Int’l Longshoremen’s Ass’n v. Allied Int’l, Inc., 456 U.S. 212 (1982).

127 549 F.2d 597 at 613.

requiring “some effect actual or intended on American foreign commerce”; an effect “sufficiently large to present a cognizable injury to the plaintiffs and, therefore, a civil violation of the antitrust laws.”¹²⁸ Most radically, it introduced the requirement of an interest of the United States strong enough—“including the magnitude of the effect on American foreign commerce”—“to justify an assertion of extraterritorial authority,” as the substantiality test could not be considered, in and by itself, “a sufficient basis on which to determine whether American authority should be asserted in a given case as a matter of international comity and fairness.”¹²⁹ The need to establish the existence of such an interest would have been consistent with a “jurisdictional rule of reason,” and required an appraisal of a number of elements such as:

[T]he degree of conflict with foreign law or policy, the nationality or allegiance of the parties and the locations or principal places of businesses or corporations, the extent to which enforcement by either state can be expected to achieve compliance, the relative significance of effects on the United States as compared with those elsewhere, the extent to which there is explicit purpose to harm or affect American commerce, the foreseeability of such effect, and the relative importance to the violations charged of conduct within the United States as compared with conduct abroad.¹³⁰

In short, *Timberlane* found the solution set out in *Alcoa* wanting—or “incomplete, because it fails to consider the interests of other nations in the application or nonapplication of United States law”¹³¹—and introduced interest-balancing to the picture.¹³² What is more, it did so based on comity.¹³³ It bears noting that this approach was more reflective of the *Zeitgeist*, as regulation was now to be assumed

128 *Id.*

129 *Id.*

130 *Id.* at 613–14.

131 Lowenfeld, *supra* note 24, at 44.

132 549 F.2d at 613–14.

133 Katherine J. Florey, *State Law, US Power, Foreign Disputes: Understanding the Extraterritorial Effect of State Law in the Wake of Morrison v. National Australia Bank*, 92 B. U. L. REV. 543 (2012).

to have some extraterritorial effects,¹³⁴ and its eventual fortune was no accident: The *Timberlane* interest-balancing methodology was incorporated in the Restatement (Third) of Foreign Relations law. Yet, the Restatement does not mention the term “comity,” perhaps “because the reporters believed that comity carries too much of the idea of discretion or even political judgment, as contrasted with the principle of reasonableness, which is conceived of in terms of legal obligation,”¹³⁵ and its commentary seems to suggest a qualification as an international legal obligation.¹³⁶ To some extent, this approach may have been due to the fact that the *Timberlane* approach had come under attack in the influential case *Laker Airways Ltd. v. Sabena, Belgian World Airlines*,¹³⁷ where it was memorably held that “[i]f promotion of international comity is measured by the number of times United States jurisdiction has been declined under the ‘reasonableness’ interest balancing approach, then it has been a failure.”¹³⁸ In the majority opinion, Judge Wilkey found that the notion of comity had been stretched beyond its bounds, relying on *Hilton v. Guyot* to state that “comity ‘never obligates a national forum to ignore the rights of its own citizens or of other persons who are under the protection of its laws,’”¹³⁹ and remained critical of the difficulty of reconciling interest-balancing with the function of courts, which he urged were

134 See e.g., the *European Woodpulp* cases, where comity arguments were raised by a number of Canadian applicants contending that by regulating their conduct—relating to activities performed outside of, but having effects within Europe—the Commission had ‘infringed Canada’s sovereignty and thus breached the principle of international comity. The Court swiftly dismissed the argument, stating that it amounted to questioning the Community’s jurisdiction to apply its competition rules. See Case C-89/85, *Ahlström v Commission (Woodpulp II)*, 1993 ECR 1307. See generally Schultz & Ridi, *supra* note 4 at 4 (Discussing how international courts have approached the issue of comity when dealing with extraterritorial regulation).

135 Lowenfeld, *supra* note 24, at 52. The author does, however, note that “[i]f agreement can be reached or approached on content, it may not be worthwhile continuing to debate the terminology.” *Id.*

136 RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW OF THE UNITED STATES § 403 (AM. LAW INST. 1987).

137 *Laker Airways Ltd. v. Sabena, Belgian World Airlines*, 731 F.2d 909 (D.C. Cir. 1984).

138 *Id.* at 950.

139 *Id.* at 943. This judgment may have been influenced by a seminal and skeptical—article by Harold G. Maier, *Interest Balancing and Extraterritorial Jurisdiction*, 31 AM. J. COMP. L. 579–597, 590–91 (1983) and *passim*. See Karl M. Meessen, *Antitrust Jurisdiction under Customary International Law*, 78 AM. J. INT’L. L. 783, 788 (1984).

not “organs of political compromise.”¹⁴⁰

It was, in any event, with *Hartford Fire Ins. Co. v. California* that the Supreme Court finally seized an opportunity to discuss comity in this context again.¹⁴¹ Often hailed as a “death blow” to the principle,¹⁴² the case concerned the conduct taken by American and British reinsurance and co-insurance companies, which had conspired to limit their offering in the United States, with anticompetitive consequences stretching to the United States. The London reinsurers argued that the Sherman Act did not apply to them, and justified the claim on the basis of comity: The activities they had carried out, they contended, were lawful in the United Kingdom, which had “established a comprehensive regulatory scheme governing the London reinsurance markets, and clearly has a heavy “interest in regulating the activity.”¹⁴³ The majority, however, was relatively unmoved by the claim. In the end, the Court held—though by a small margin—that it was not enough for the conduct of the reinsurers to be compliant with foreign law when compliance with foreign *and* American law was possible. The Court held that only a “true conflict” would have imported considerations of comity in resolution of the dispute and, since the situation in the case at issue did not amount to one, comity was no ground to decline the court’s jurisdiction.¹⁴⁴ In the powerful dissent he penned for the minority, destined to be—with all its wisdom and its deficiencies—more deserving of doctrinal attention than the majority opinion, Justice Scalia spotted the main shortcomings of the decision. As he observed, the question should not have been one of jurisdiction of the courts, but rather one “of substantive law turning on whether, in enacting the Sherman Act, Congress asserted regulatory power over the challenged conduct.”¹⁴⁵ Citing Story and his *Commentaries*, Scalia lectured on how the precedents cited by the majority did not refer to:

[T]he comity of courts, whereby judges decline to exercise jurisdiction over matters more appropriately adjudged elsewhere, but rather what might be

140 *Laker*, 731 F.2d at 953.

141 *Hartford Fire Ins. Co. v. California*, 509 U.S. 764 (1993).

142 Waller, *supra* note 110, at 564.

143 *Hartford Fire*, 509 U.S. at 819 (Scalia, J.) (citing RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW).

144 In this case, the court concerned was the District Court for the Northern District of California. *Hartford Fire*, 509 U.S. at 778–79, 820.

145 *Hartford Fire*, 509 U.S. at 813 (Scalia, J.).

termed “prescriptive comity”: the respect sovereign nations afford each other by limiting the reach of their laws . . . [C]omity in this sense includes the choice-of-law principles that, “in the absence of contrary congressional direction,” are assumed to be incorporated into our substantive laws having extraterritorial reach. Considering comity in this way is just part of determining whether the Sherman Act prohibits the conduct at issue.¹⁴⁶

On this point, Scalia was absolutely right, and it is possibly a matter of regret that he would reinforce an argument built on choice of law¹⁴⁷ with a confusing allusion to the public international law limits on jurisdiction to prescribe.¹⁴⁸ As we explained, comity *presupposes* the existence of any such jurisdiction — be it to prescribe, adjudicate, or enforce. It is unsurprising that it was Scalia’s version that prevailed in the subsequent Supreme Court opinion *F. Hoffman-La Roche Ltd. v. Empagran S.A.*, where it was “principles of prescriptive comity counsel[led] against the Court of Appeals’ interpretation of” United States antitrust law: “[I]f America’s antitrust policies could not win their own way in the international marketplace for such ideas, Congress, we must assume, would not have tried to impose them in an act of legal imperialism.”¹⁴⁹

But was this a return to comity? *Hartford Fire* decision did strike a blow to the principle that is hard to underestimate. Consider, for instance, *United States v. Nippon Paper Indus. Co.*, an antitrust case four

146 *Id.* at 817–18. Dodge observes—correctly—that this is not what Story meant, because in his time Courts did not have the authority to decline jurisdiction. William S. Dodge, *International Comity in American Law*, 115 COLUM. L. REV. 2071, 2106–07 (2015). This is true, but it might be added that, before *International Shoe*, courts did not have extraterritorial jurisdiction anyway (except, of course, in admiralty cases). *International Shoe, Co. v. Washington*, 326 U.S. 310 (1945). If this detail is taken into account, Scalia’s “rhetorical flourish” does not appear to be inconsistent with Story’s thinking. See Dodge, *supra* note 2, at 2106–07.

147 *Hartford Fire*, 509 U.S. at 818 (Scalia J.) (citing, among others, *Lauritzen*, 345 U.S. 571 (1953)).

148 *Id.* at 815–16 (Scalia, J.) (citing the canon of interpretation set out by Chief Justice Marshall in *Murray v. Schooner Charming Betsy*, 2 Cranch 64, 118 (1804) (“[A]n act of congress ought never to be construed to violate the law of nations if any other possible construction remains.”)).

149 *F. Hoffman-La Roche Ltd. v. Empagran S.A.*, 542 U.S. 155, 169 (2004).

years after *Hartford Fire*.¹⁵⁰ The First Circuit affirmed that comity was “more an aspiration than a fixed rule, more a matter of grace than a matter of obligation. In all events, its growth in the antitrust sphere has been stunted by *Hartford Fire*.”¹⁵¹ The conclusion in this case was that wholly foreign conduct having an intended and substantial effect in the United States could form the basis for *criminal* prosecution under American antitrust law, a finding hard to reconcile with any understanding of comity. *Empagran* too is not too far a stretch from *Hartford Fire*, as comity was grounds for the Court to dismiss the claims relating to alleged foreign damages, but not those concerning domestic harmful effects.¹⁵² In a concise article on the topic, Joel Paul made the argument that this conclusion is problematic in that the Court appeared to make a shift towards deference to interests of a global nature, rather than to the specific interest of a given sovereign counterpart.¹⁵³ Whether this was the case is questionable, and the idea that “[d]eference to the Market has nothing to do with respect for foreign law or private parties” might be criticized in light of the historical roots of comity.¹⁵⁴ In these respects, the court in *Empagran* did attempt to strike a balance, but while it claimed to construe “ambiguous statutes to avoid unreasonable interference with other nations’ sovereign authority,” it arguably fell short of the mark by leaving untouched an approach allowing for intrusive regulatory interference — rather than helping, as it professed, “the potentially conflicting laws of different nations work together in harmony.”¹⁵⁵

On the whole, the “breathtakingly broad”¹⁵⁶ holding of *Hartford Fire* has made it difficult for courts to engage in a comity analysis where a “true conflict” cannot be identified, and it is possibly responsible for much terminological and conceptual confusion, so that while courts in subsequent cases are more critical of this approach, it is sometimes difficult to tell where disagreement ends and misunderstanding begins.¹⁵⁷

150 United States v. Nippon Paper Indus. Co., 109 F.3d 1, 9 (1st Cir. 1997).

151 *Id.* at 8.

152 F. Hoffman-La Roche Ltd v Empagran S.A., 542 U.S. 155, 165–66 (2004).

153 Paul, *The Transformation of International Comity*, *supra* note 9, at 36–38 (discussing the court’s recognition that market and commercial interests were at least one of several more wide spread interests than simply deferring to the interests of one sovereign).

154 *Id.* at 38.

155 542 U.S. at 156.

156 509 U.S. at 820 (Scalia, J.).

157 See, e.g., *Mujica v. AirScan Inc.*, 771 F.3d 580, 599–604 (9th Cir. 2014)

b. Outside the Antitrust Context

For a long time, comity remained relatively absent from the debate on extraterritoriality in cases not concerning antitrust, though similar notions were often cited.¹⁵⁸ It may be argued that the absence of comity might be imputed to something more than mere terminological variety. For example, in the 1991 case *EEOC v. Arabian American Oil Co. (Aramco)*, the principle was only mentioned in the dissenting opinion of Justice Marshall.¹⁵⁹ This case concerned discrimination against female American workers by their employer, a Delaware corporation, operating in Saudi Arabia. In deciding whether the Civil Rights Act of 1964 was applicable to the case, the Court concluded in the negative, relying on the presumption against extraterritoriality as a guarantee “against unintended clashes between our laws and those of other nations which could result in international discord.”¹⁶⁰

More recently, the Supreme Court has employed comity to curb the territorial reach of United States law in a few interesting cases dealing with securities regulation and, most prominently, human rights.¹⁶¹ In *Morrison v. Nat'l Austl. Bank Ltd.*,¹⁶² a number of foreign investors initiated a class action against an Australian bank alleging securities fraud. In delivering the Court's opinion,

(affirming that the “true conflicts” approach is restricted to prescriptive comity and says nothing about adjudicatory comity (as in the case at issue), only to apply the comity considerations contained in § 403 of the Restatement, which deal with prescriptive comity too). See also Roger Alford, *The Ninth Circuit's Muddled Comity Analysis in Mujica*, OPINIO JURIS (Nov. 21, 2014, 1:40 PM), <http://opiniojuris.org/2014/11/21/ninth-circuits-muddled-comity-analysis-mujica/>; Symeon C. Symeonides, *Choice of Law in the American Courts in 2014: Twenty-Eighth Annual Survey*, 63 AM. J. COMP. LAW 299, 312 (2015).

158 See, e.g., *Foley Bros., Inc. v. Filardo*, 336 U.S. 281 (1949).

159 *EEOC v. Arabian Am. Oil Co.*, 499 U.S. 244, 265, 274 (1991).

160 *Id.* at 248 (citing *McCulloch v. Sociedad Nacional de Marineros de Honduras*, 372 U.S. 10 (1963)).

161 But see *Spector v. Norwegian Cruise Line Ltd.*, 545 U.S. 119 (2005), where it was held — citing the *Wildenhus's Case*, 120 U.S. 1, 12 (1887) — that in dealing with foreign vessels comity only requires that “general statutes are presumed not to impose requirements that would interfere with the internal affairs of foreign-flag vessels.” *Id.* at 142. However, the Court stated unambiguously “that general statutes are presumed to apply to conduct that takes place aboard a foreign-flag vessel in United States territory if the interests of the United States or its citizens, rather than interests internal to the ship, are at stake.” *Id.* at 130–31.

162 *Morrison v. Nat'l Austl. Bank Ltd.*, 561 U.S. 247 (2010).

Justice Scalia focused on the extraterritorial application of the Securities Exchange Act and concluded that the presumption against extraterritoriality applied “regardless of whether there is a risk of conflict between the American statute and a foreign law. When a statute gives no clear indication of an extraterritorial application, it has none.”¹⁶³ Interestingly, discussing a statement of the Solicitor General claiming that the “significant and material conduct” test was “in accord with prevailing notions of international comity,” Scalia appeared to draw the dubious inference that such an assertion of prescriptive jurisdiction, if there had been any, would not have amounted to a violation of international law.¹⁶⁴

This remark has been criticized in American scholarship on the grounds that it blurs the line between international law and international comity, but one wonders if the assessment is not too harsh.¹⁶⁵ Indeed, the Solicitor General raised the point of conformity to comity to suggest that an intent of extraterritorial application would not have been illogical. Scalia was surely right in remarking that this was no proof that Congress had made such an assertion.¹⁶⁶ What is more, he was also formally correct in stating that such a conduct would have been in conformity to customary international law: to breach it, one must ostensibly breach comity first, never the other way around. Ultimately, while we will never know whether this truism was the actual purpose of Scalia’s discussion of the issue, the resulting terminological confusion cannot be commended.¹⁶⁷

Other cases concerned claims for violation of the “law of nations” under the Alien Tort Statute (ATS). The leading authority in this context is *Kiobel v. Royal Dutch Petroleum Company*.¹⁶⁸ Until a few years before the facts of the case, the ATS had been relegated to little more than a historical oddity. Passed as part of the Judiciary Act of 1789, it granted the District Courts “original jurisdiction of any civil action by an alien for a tort only, committed in violation

163 *Id.* at 255 (citing *Sale v. Haitian Centers Council, Inc.*, 509 U.S. 155, 173–174).

164 *Id.* at 272.

165 Dodge, *supra* note 2, at 2120.

166 Briggs, *supra* note 4, at 104.

167 On the questions of extraterritorial application of state law raised by *Morrison*, see Florey, *supra* note 134.

168 *Kiobel v. Royal Dutch Petroleum Co.*, 133 S. Ct. 1659 (2013). On the history of the ATS, see Anne-Marie Burley, *The Alien Tort Statute and The Judiciary Act of 1789: A Badge of Honor*, 83 AM. J. INT’L L. 461 (1989); William S. Dodge, *The Historical Origins of the Alien Tort Statute: A Response to the Originalists*, 19 HASTINGS INT’L COMP. REV. 221 (1995).

of the law of nations or a treaty of the United States.”¹⁶⁹ The Act remained forgotten until the early eighties, when the Second Circuit rediscovered it in the landmark case *Filártiga v. Peña-Irala*, holding that torture amounted to a breach of the law of nations and that the ATS provided federal jurisdiction.¹⁷⁰

In *Kiobel*, Nigerian petitioners filed a suit against a number of corporations alleging that they had aided and abetted the Nigerian government in committing violations of the law of nations.¹⁷¹ The Court of Appeals had rejected the claims, reasoning that the law of nations did not recognize corporate liability, and certiorari was eventually granted by the Supreme Court.¹⁷² The case, however, was resolved on other grounds, namely that the facts of the case had taken place outside of the United States and that nothing in the ATS rebutted the presumption against extraterritoriality.¹⁷³ In order to do so, the Court traced the evolution of the ATS back to its historical roots¹⁷⁴ and considered the foreign policy implications of extraterritorial ATS jurisdiction in terms of interference with both other sovereign states and the executive as the sole responsibility of the United States foreign policy.¹⁷⁵ In particular, the Court used the historical argument to claim that the prosecution of admittedly extraterritorial conduct such as piracy did not interfere with the preserve of other sovereigns, and the foreign policy argument to reason that haling foreign defendants in United States courts for breaches of international law occurring abroad might have provoked the undesired result of the opposite also applying.¹⁷⁶

These conclusions seem consistent with Justice Breyer’s findings in his separate opinion in *Sosa v. Alvarez-Machain*.¹⁷⁷ However, in that case the Court had limited the applicability of the ATS to violations of “specific, universal, and obligatory” norms.¹⁷⁸ In his

169 28 U.S.C. § 1330 (2012). On the increased reliance on custom for suits under the ATS, see JAMES CRAWFORD, CHANCE, ORDER, CHANGE: THE COURSE OF INTERNATIONAL LAW, GENERAL COURSE ON PUBLIC INTERNATIONAL LAW 164, 234 (2014).

170 *Filartiga v. Pena-Irala*, 630 F.2d 876, 878 (2nd Cir. 1980).

171 *Kiobel*, 569 U.S. 108, 113–14.

172 *Id.*

173 *Id.* at 1669.

174 *Id.* at 1665–66.

175 *Id.* at 1663–65.

176 *Id.* at 1669.

177 *Sosa v. Alvarez-Machain*, 542 U.S. 692 (2004) (Breyers, S., concurring).

178 *Id.* at 748 (quoting *Hilao v. Marcos*, 25 F.3d 1467, 1475 (9th Cir. 1994)).

concurrence, Breyer endorsed this choice on the basis that “universal criminal jurisdiction necessarily contemplates a significant degree of civil tort recovery,” and the fact that universal tort jurisdiction could threaten “the practical harmony that comity principles seek to protect” militated in favor of a limited number of actionable norms.¹⁷⁹ Accordingly, Breyer disagreed again with the Court’s reasoning in *Kiobel*: In his view, it was wrong to decide the case on the basis of the presumption against extraterritoriality. Rather, he stated that the ATS provided jurisdiction where:

- (1) the alleged tort occurs on American soil, (2) the defendant is an American national, or (3) the defendant’s conduct substantially and adversely affects an important American national interest, and that includes a distinct interest in preventing the United States from becoming a safe harbor (free of civil as well as criminal liability) for a torturer or other common enemy of mankind.¹⁸⁰

To Breyer and those who joined the minority opinion, to interpret the ATS this way would be consistent with the international law of jurisdiction and with the comity concerns of minimizing international friction.¹⁸¹

c. Analysis

There can no longer be any doubt that comity plays a role in limiting the territorial reach of American law, though it is nevertheless distinct from the presumption against extraterritoriality.¹⁸² It is, however, clear from the above that the use of the principle to this end is not necessarily consistent.

Let us consider, for example, the *Aramco* case. As Posner and Sunstein observed, “does Saudi Arabia really care about sex

179 *Id.* at 763.

180 *Kiobel*, 569 U.S. 108, 132–33.

181 *Id.*

182 See *In re Maxwell Commc'n Corp. plc* by Homan, 93 F.3d 1036, 1047 (2d Cir. 1996) (“Moreover, international comity is a separate notion from the ‘presumption against extraterritoriality,’ which requires a clear expression from Congress for a statute to reach non-domestic conduct Because the principle of comity does not limit the legislature’s power and is, in the final analysis, simply a rule of construction, it has no application where Congress has indicated otherwise.”).

discrimination by American businesses practiced against American employees? Even if it does, does it care enough that the discriminatory practice should be tolerated?"¹⁸³ The principle of comity, as it has been understood by American courts, does allow this type of analysis—therein lies its importance, and the reason why it was precisely comity that Justice Marshall mentioned in his dissent in *Aramco*.¹⁸⁴ And yet, it is interesting to observe that the Supreme Court has been much bolder in affirming that comity was no bar to the exercise of extraterritorial prescriptive jurisdiction in the antitrust sphere than it has been in other significant areas, such as worker discrimination or tortious claims under the ATS.¹⁸⁵ True, such choices lie within the conceptual boundaries of the principle, but it is arguably difficult to justify a difference in treatment of anticompetitive conduct and discrimination against American nationals when both cause—to play with Huber's words—"prejudice to the power or rights of such governments or of its subjects."¹⁸⁶ As Louise Weinberg lucidly put it "[t]his is a vivid example of how 'comity' can mean accommodation to values repugnant to this country."¹⁸⁷

Finally, the attitude of the American judiciary suggests that we spend a few words on the relationship between comity and territoriality. The two concepts have always gone hand in hand, the former being a predicate of the latter, but always in the sense that one state should allow foreign law to have extraterritorial effect in its territory unless their fundamental domestic interests militated against it.¹⁸⁸ This model coherently assumes—rather than presumes—

183 Eric A. Posner & Cass R. Sunstein, *Chevronizing Foreign Relations Law*, 116 YALE L.J. 1170, 1173 (2007).

184 EEOC v. Arabian Am. Oil Co., 499 U.S. 244, 265 (1991). *But see* Briggs, *supra* note 4, at 96 (maintaining that the Court applied principles descending from comity).

185 On ATS litigation post-*Kiobel* landscape, see Donald Earl Childress III, *Is an International Arbitral Tribunal the Answer to the Challenges of Litigating Transnational Human Rights Cases in a Post-Kiobel World?*, 19 UCLA J. INT'L FOREIGN AFF. 31, 34 (2015) (discussing the SCOTUS's decision in *Kiobel* casting doubt over a plaintiff's ability to bring ATS claims against foreign nationals for human rights violations before a US court.).

186 Lorenzen, *Huber's De Conflictu Legum*, *supra* note 55, at 403.

187 Weinberg, *supra* note 18, at 74 (emphasis added). Speaking of the *Aramco* case, Weinberg notes that some of her readers may conclude that *Aramco*'s result might have been "a practical necessity, particularly at a time when hundreds of American companies are competing for contracts to rebuild Kuwait." Weinberg, *supra* note 18, at 74.

188 For Huber's original statement of the doctrine, see *supra* note 5 and accompanying text.

legislation to be territorial: to state that national interests justify a greater reach of the forum's law just because they allow limitation of foreign law is stretching the theoretical boundaries of the paradigm. To clarify, we are not claiming that presumptively limiting the reach of the forum's law to the forum's territory is inconsistent with comity. Quite the contrary, we observe that extending it because of the importance of domestic interests does not descend naturally from the principle. If anything, the point contributes to highlight the significance of the distinction between matters of international law of jurisdiction and comity, in that state interests are the justification of exorbitant jurisdiction and comity an incentive to its restriction.

B. Comity and the Recognition of Foreign Judicial Acts

The doctrine of comity has for a long time provided the basis for the recognition of foreign judgments in the United States.¹⁸⁹ In this regard, Justice Gray's opinion in *Hilton v. Guyot* has undeniably represented "the lodestar for all transnational enforcement doctrines in the U.S."¹⁹⁰ The Court famously held that comity was:

[N]either a matter of absolute obligation, on the one hand, nor of mere courtesy and good will, upon the other. But it is the recognition which one nation allows within its territory to the legislative, executive or judicial acts of another nation, having due regard both to international duty and convenience, and to the rights of its own citizens, or of other persons who are under the protection of its laws.¹⁹¹

The case concerned the recognition of a decision rendered in a French court. The comity analysis in the case was partly consistent with the theories of Story and the Dutch writers, predicated as it

189 A number of authorities have classified the recognition of foreign judgments under the heading of adjudicative comity. While we agree that there is merit to that view, we prefer to privilege the aspects concerning comity as a principle of recognition and treat it separately. For different views, see *Ungaro-Benages v. Dresdner Bank AG*, 379 F.3d 1227, 1238 (11th Cir. 2004) ("The doctrine of international comity can be applied retrospectively or prospectively. When applied retrospectively, domestic courts consider whether to respect the judgment of a foreign tribunal or to defer to parallel foreign proceedings."). See also *Dodge*, *supra* note 2, at 2083; *Childress*, *supra* note 2, at 16.

190 *Koh*, *supra* note 27, at 206.

191 *Hilton v. Guyot*, 159 U.S. 113, 163–64 (1895).

was on the idea that any such recognition should not have caused harmful effects to the United States, its citizens, or their rights, and resulted in the rule whereby if:

the foreign judgment appears to have been rendered by a competent court, having jurisdiction of the cause and of the parties, and upon due allegations and proofs, and opportunity to defend against them, and its proceedings are according to the course of a civilized jurisprudence, and are stated in a clear and formal record, the judgment is *prima facie* evidence, at least, of the truth of the matter adjudged; and it should be held conclusive upon the merits tried in the foreign court, unless some special ground is shown for impeaching the judgment, as by showing that it was affected by fraud or prejudice, or that by the principles of international law, and by the comity of our own country, it should not be given full credit and effect.¹⁹²

The Court thus provided a principled framework for the recognition of foreign judgments. However, there is at least one more reason to hail *Hilton* as a “watershed moment in the history of international comity,”¹⁹³ and that of abidance to the rule of reciprocity, which the Court—perhaps reading too much in Story’s *Commentaries*—placed firmly within “the structure of international jurisprudence.”¹⁹⁴ The Court reasoned that “if the judgment had been rendered in this country, or in any other outside of the jurisdiction of France, the French courts would not have executed or enforced it, except after examining into its merits.” Accordingly, arguing on the basis that “international law is founded upon mutuality and reciprocity,” “the principles of international law recognized in most civilized nations, and by the comity of our own country,” the Court held that “the judgment [was] not entitled to be considered conclusive.”¹⁹⁵

192 *Id.* at 205–06.

193 Childress, *supra* note 2, at 31.

194 *Hilton v. Guyot*, 159 U.S. 113, 227 (1895) (citing STORY, *supra* note 14, at § 618) (Story only observes that the rule of reciprocity is “a very reasonable” one.).

195 *Hilton*, 159 U.S. at 228.

Hilton's significance for the recognition of judgments also faded to some degree after the Supreme Court's decisions in *Erie* and *Klaxon*. Indeed, Hilton has been considered in contrast to the more efficient rules of state law,¹⁹⁶ and the reciprocity rule denounced by courts more or less expressly.¹⁹⁷ In any event, even before *Erie* many American legal minds were aware of the limitations of Hilton's reciprocity. As early as 1925, Judge Learned Hand affirmed that

[w]hatever may be thought of that decision, the court certainly did not mean to hold that an American court was to recognize no obligations or duties arising elsewhere until it appeared that the sovereign of the locus reciprocally recognized similar obligations existing here. That doctrine I am happy to say is not a part of American jurisprudence.¹⁹⁸

The criticism of Hilton's model of reciprocity was in fact not too uncommon.¹⁹⁹ Perhaps the most compelling exposition of the flaws inherent in a reciprocity analysis may be found in Judge Cuthbert Pound's opinion in *Johnston v. Compagnie Generale Transatlantique*²⁰⁰:

It is argued with some force that questions of

196 See, e.g., *Banque Libanaise Pour Le Commerce v. Khreich*, 915 F.2d 1000, 1004 (5th Cir. 1990) ("Although comity is not a rule of law, it is more than mere courtesy and accommodation . . . Under the Texas Recognition Act the rules relating to the recognition of foreign country money-judgments are statutory and therefore more predictable.") (citation omitted).

197 See, e.g., *De la Mata v. Am. Life Ins. Co.*, 771 F. Supp. 1375, 1383 (D. Del. 1991) ("[T]he court predicts that the Delaware Supreme Court would no longer regard reciprocity as a precondition for the recognition of a foreign judgment."). See also *Johnston v. Compagnie Generale Transatlantique*, 242 N.Y. 381, 387 (1926) ("Comity is not a rule of law, but it is a rule of 'practice, convenience and expediency. It is something more than mere courtesy, which implies only deference to the opinion of others, since it has a substantial value in securing uniformity of decision, and discouraging repeated litigation of the same question.'").

198 *Direction der Disconto-Gesellschaft v. U.S. Steel Corp.*, 300 F. 741, 747 (S.D.N.Y. 1924).

199 *Disconto Gesellschaft v. Umbreit*, 208 U.S. 570, 580 (1908) ("There being, then, no provision of positive law requiring the recognition of the right of the plaintiff in error to appropriate property in the state of Wisconsin and subject it to distribution for the benefit of foreign creditors as against the demands of local creditors, how far the public policy of the state permitted such recognition was a matter for the state to determine for itself").

200 *Johnston v. Compagnie Generale Transatlantique*, 242 N.Y. 381 (1926).

international relations and the comity of nations are to be determined by the Supreme Court of the United States But the question is one of private rather than public international law, of private right rather than public relations, and our courts will recognize private rights acquired under foreign laws and the sufficiency of the evidence establishing such rights. . . . Comity is not a rule of law, but it is a rule of ‘practice, convenience and expediency. It is something more than mere courtesy, which implies only deference to the opinion of others, since it has a substantial value in securing uniformity of decision, and discouraging repeated litigation of the same question.’ It therefore rests, not on the basis of reciprocity, but rather upon the persuasiveness of the foreign judgment. When the whole of the facts appear to have been inquired into by the French courts, judicially, honestly, and with full jurisdiction and with the intention to arrive at the right conclusion, and when they have heard the facts and come to a conclusion, it should no longer be open to the party invoking the foreign court against a resident of France to ask the American court to sit as a court of appeal from that which gave the judgment. I reach the conclusion that this court is not bound to follow the Hilton Case and reverse its previous rulings.²⁰¹

Be that as it may, comity has largely lost its role as ground upon which foreign judgments are recognized, at least as a common law doctrine, as the matter is now covered by Uniform Acts,²⁰² which provide more straightforward solutions. It must be observed, however, that the doctrine has not been sidestepped completely. First of all, comity played a role for the period of time in which the Acts had yet to receive a homogeneous adoption by states. In a Tenth Circuit case, it was affirmed that comity governed the

201 *Id.* at 386–87 (citations omitted).

202 UNIFORM FOREIGN MONEY-JUDGMENTS RECOGNITION ACT (NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS 1962); UNIFORM FOREIGN-COUNTRY MONEY JUDGMENTS RECOGNITION ACT (NATIONAL CONFERENCE OF COMMISSIONERS ON UNIFORM STATE LAWS 2005).

recognition of foreign judgments in Utah, a state that had not yet adopted the Uniform Acts.²⁰³ Second, even a cursory reading of the 2005 Uniform Foreign-Country Money Judgments Recognition Act shows as much, as its “saving clause” states that it “does not prevent the recognition under principles of comity or otherwise of a foreign-country judgment not within [its] scope[.]”²⁰⁴ Third, the Acts’ focus on concepts such as “repugnancy” and their understanding by the courts is certainly revealing of their foundation in comity.²⁰⁵ What is more, the courts have used comity as a canon of statutory construction for the Uniform Acts: In a 2013 case, the Ninth Circuit held that the act’s public policy exception did not apply “unless a foreign-country judgment or the law on which it is based is ‘so antagonistic to California [or federal] public policy interests as to preclude the extension of comity.’”²⁰⁶ It might be added that the Uniform Acts do not, for example, cover injunctions. Accordingly, in *Yahoo! Inc. v. La Ligue Contre Le Racisme Et L’Antisemitisme*, the Ninth Circuit stated that “[b]ecause the Uniform Act does not cover injunctions, we look to general principles of comity followed by the California courts.”²⁰⁷ Finally, it has been argued in at least one case that the demands of comity might be met by requiring that non-enforceable judgments be granted recognition and their enforcement being left untouched.²⁰⁸ This might, of course, be a meaningful result on its own—one may

203 12 Soc'y of Lloyd's v. Reinhart, 402 F.3d 982, 999 (10th Cir. 2005).

204 *Id.* at § 11.

205 See Wolff v. Wolff, 40 Md. App. 168, 175 (1979) (“Thus the Uniform Foreign Money-Judgments Recognition Act was intended to promote principles of international comity by assuring foreign nations that their judgments would, under certain well-defined circumstances, be given recognition by courts in states which have adopted the Uniform Act. . . . The Act, therefore delineates a Minimum of foreign judgments which Must be recognized in jurisdictions which have adopted the Act, and in no way constitutes a Maximum limitation upon foreign judgments which May be given recognition apart from the Act.”). See also Marc P. Epstein, *Comity Concerns Are No Joke: Recognition of Foreign Judgments Under Dormant Foreign Affairs Preemption*, 82 FORDHAM REV 2317, 2321 (2013).

206 Naoko Ohno v. Yuko Yasuma, 723 F.3d 984, 1002 (9th Cir. 2013) alteration in original (citing *Crockford's Club Ltd. v. Si-Ahmed*, 203 Cal. App. 3d 1402, 1406 (Ct. App. 1988)).

207 Yahoo Inc. v. La Ligue Contre Le Racisme Et L’Antisemitisme, 433 F.3d 1199, 1213 (9th Cir. 2006).

208 Guinness PLC v. Ward, 955 F.2d 875, 889 n.9 (4th Cir. 1992) (stating “[W]e nonetheless believe that such goal as well as the principles of comity are still sufficiently served by the fact that judgments which are not enforceable might still be entitled, if consistent with the Act's criteria, to recognition.”).

think of the *res judicata* effect of such a judgment—and exemplifies a conclusion that could not be reached by the reading of the Acts alone.

In conclusion, comity reasoning appears to be deeply entrenched in matters of recognition of foreign judgments and, more generally, judicial acts. In this regard, it can be observed that *Hilton*'s legacy remains uncontested authority for the view that comity constitutes the basis for the recognition of foreign judgments, filling the gaps left by and fine-tuning the more detailed regulation of the matter by state sources. Conversely, reciprocity, while sometimes invoked as a “relevant factor” for the granting of comity, has failed to become an essential element of a comity analysis in this context.²⁰⁹

Indeed, the main problem of *Hilton*'s legacy is not connected to its definition of the doctrine,²¹⁰ or to its linking the concept of reciprocity to the recognition of judgments, but rather to the way its rationale has been extended to other manifestations of comity—namely, those concerning applicable law and management of parallel proceedings.²¹¹ Some see *Hilton*'s continued citation where comity is invoked as a matter of regret. Childress calls *Hilton*'s treatment of the matter “woefully inadequate” because of its relinquishment of a sovereign interests rationale for reciprocity and an unsatisfactory foundation that “does not provide courts with concrete direction in applying the doctrine.”²¹² As Louise Weinberg put it, reciprocity “is simply not as safe an item as ‘motherhood’ or ‘apple pie’.”²¹³

It is, however, not entirely true that the recognition of foreign judgments is entirely devoid of foreign relations consequences: the recent *Chevron* cases²¹⁴ are testament to this proposition. As it is well

209 *Cunard S.S. Co. v. Salen Reefer Servs.* AB, 773 F.2d 452, 460 (2nd Cir. 1985).

210 As Calamita put it, *Hilton*'s definition does convey the idea that comity “stands as a short-hand term to denote a collection of identifiable interests and values that must be balanced in given cases and from which rules of law develop.” Calamita, *supra* note 22, at 626.

211 *In re Maxwell Commc'n Corp.*, 93 F.3d 1036, 1046 (2d Cir. 1996) (“Although Hilton addressed the degree to which a foreign judgment is conclusive in a court of the United States, the principle expressed is one of broad application.”).

212 Childress, *supra* note 2, at 33–34. Partially against this view, see Briggs, *supra* note 4, at 145.

213 Weinberg, *supra* note 18, at 70.

214 On the delicate political implications of the *Cevron* dispute, see generally *Chevron Corp. v. Donziger*, 768 F. Supp. 2d 581 (S.D.N.Y. 2011) (granting an injunction to prevent the Ecuadorian judgment from being enforced), *rev'd and remanded sub nom.*, *Chevron Corp. v. Naranjo*, 667 F.3d 232, 244 (2d Cir. 2012) (holding that the judgment of the District Court is vacated until Ecuador

known, the District Court had, acting under the recognition law of the forum state, granted an injunction enjoining and restraining the defendants “from directly or indirectly funding, commencing, prosecuting, advancing in any way, or receiving benefit from any action or proceeding, outside the Republic of Ecuador, for recognition or enforcement of the judgment previously rendered . . .”²¹⁵ The rationale for the decision was that the judgment was “rendered under a system which does not provide impartial tribunals or procedures compatible with the requirements of due process of law . . .”²¹⁶ Vacating the injunction, the Second Circuit held that it was a “particularly weighty matter for a court in one country to declare that another country’s legal system is so corrupt or unfair that its judgments are entitled to no respect from the courts of other nations.”²¹⁷ This, if comity is to be respected, would have been a correct conclusion even if the matter had been limited to the recognition of foreign judgment. The Court, however, was also concerned with the implications of an injunction with such wide-ranging, global effects, as it found one such injunction to be very close to tacitly holding the courts of other countries “insufficiently trustworthy to recognize what is asserted to be the extreme incapacity of the legal system from which the judgment emanates.”²¹⁸ As Briggs points out, solutions to these issues are not readily available: If these questions are to be resolved on the basis of comity, possibly infringed by any such judgment of non-recognition, the answer, lacking a clear, sovereign—or international law—command overriding comity considerations, “depends, inevitably, on what comity is.”²¹⁹

C. *Adjudicatory Comity, or “the Comity of Courts”*

Comity has also been invoked to justify certain approaches towards adjudication carried out in foreign countries and, to some extent, international courts and tribunals. This is an area in which the notion of “comity” falls close to the notion of “deference,” as these doctrines are clearly understood to arise not from a lack of jurisdictional power, but rather from the discretionary decision not

attempts to enforce it in a New York court).

215 *Donziger*, 768 F. Supp. 2d at 660, 631, 633.

216 *Id.* at 636.

217 *Naranjo*, 667 F.3d at 244, 242, 247 (dictum).

218 *Id.* at 244. The Court based its reasoning on New York state law. For our purposes this distinction is immaterial. *See Epstein, supra* note 206, at 2352.

219 Briggs, *supra* note 4, at 141.

to exercise one that exists—indeed, if it did not, no conflict would occur—for a number of reasons.²²⁰ Within the debate on the modern idea of comity, the notion that the principle may provide the basis for the coordination of the exercise of adjudicatory power, or is at least implicated in it, is usually assumed as a starting point.²²¹ This role of comity, however, can hardly be taken for granted. As Campbell McLachlan observed in his 2009 Hague lectures on *Lis Pendens in International Litigation*:

[T]he early Dutch fathers of the Conflict of Laws did not [...], proceed to consider the application of *lis pendens* to the problem of the conflict of litigation internationally.... cases of this kind did not begin to tax the English courts until the nineteenth century. In any event, the application of the doctrine of *lis pendens* in the form in which Huber and Voet recognized it would have required a much stronger form of comity than they had otherwise accepted for the effects of foreign legal systems. It was not merely a matter of courts giving effect to foreign law and foreign judgments. Rather, it would have required the court to cede the power of adjudication to a foreign court.²²²

This latter problem has traditionally represented the most significant hurdle: Indeed, it is one thing to accept that *acts* issued by a foreign court must be entitled to recognition insofar as they accord with the fundamental policies of the forum state; to concede that the same respect should be granted to mere *proceedings* might sound to some as crossing the Rubicon of judicial abdication. To be sure, adjudicative comity encompasses a wide range of hypotheses and not all of them call for the same reactions. Dismissing local proceedings when none are pending before a foreign court is not the same as doing so when they have been instituted and a clash is inevitable. Different yet is the case of anti-suit injunctions issued for the purposes of frustrating the effects of potential or actual foreign

220 This, of course, does not exclude the possibility that in some cases “the exercise of apparently discretionary rules could mask an underlying objective of compliance with international limitations on judicial authority.” MILLS, *supra* note 17, at 6.

221 Calamita, *supra* note 22, at 614.

222 McLACHLAN, *supra* note 60, at 32–33 (emphasis added).

parallel proceedings. As Briggs observed, these cases do not have the same implications for comity.²²³ Yet, comity does have a role to play in all of them. According to William Dodge, it does so by informing doctrines of abstention that “defer to foreign courts by restraining the exercise of U.S. courts’ jurisdiction.”²²⁴

1. *The Origins of Adjudicatory Comity in Admiralty Courts and the Use of Forum Non Conveniens*

Before the Supreme Court’s decision in *International Shoe Co. v. Washington*, the jurisdiction of American Courts was understood to be strictly territorial.²²⁵ On the other hand, the jurisdiction of admiralty courts was without limitations. This was by design as the interests of justice so demanded in many cases, most importantly in those brought by seamen for their wages or injuries suffered during employment.²²⁶ And yet, taking jurisdiction in such cases often meant holding a ship in an American port, an action liable to be perceived as an unjustified interference by the flag state.²²⁷ In this regard, a choice to decline jurisdiction echoed comity concerns. As one early case put it, “upon the one hand the courts are not without ample power to hear and determine such suits, when the circumstances of the case before them seem to render it fit that they should do so; while, upon the other hand, they are not bound to do this, but will, in general, from motives of international comity, of delicacy, and of convenience, decline the suit.”²²⁸

One way the courts addressed this problem was through the doctrine of *forum non conveniens*.²²⁹ *Forum non conveniens* is a common law doctrine which results in the “discretionary power of a court to decline to exercise a possessed jurisdiction whenever it appears that the cause before it may be more appropriately tried elsewhere.”²³⁰

223 Briggs, *supra* note 4, at 116.

224 Dodge, *supra* note 2, at 2116.

225 See *Int’l Shoe Co. v. Washington*, Office of Unemployment Comp. & Placement, 326 U.S. 310, 316 (1945) (clarifying that the jurisdiction of a court should be conditioned to the defendant having at least “minimum contacts” with the forum’s territory).

226 Alexander M. Bickel, *Doctrine of Forum Non Conveniens As Applied in the Federal Courts in Matters of Admiralty*, 35 CORNELL L. REV. 12, 19–20 (1949).

227 *Id.* at 20–21.

228 *Davis v. Leslie*, 7 F. Cas. 134, 137 (S.D.N.Y. 1848).

229 *Am. Dredging Co. v. Miller*, 510 U.S. 443, 464 (1994) (Kennedy, J., dissenting). (“From the beginning, American admiralty courts have confronted this problem through the *forum non conveniens* doctrine.”).

230 Paxton Blair, *The Doctrine of Forum Non Conveniens in Anglo-American Law*, 29

The “plea” of *forum non conveniens* probably originated in Scotland²³¹ and was further developed in English courts.²³² The doctrine, however, is largely unknown to civil law systems, though it might—after having suffered a significant blow²³³—be in the process of being rediscovered under the Recast Brussels Regulation.²³⁴

While the relationship between comity and *forum non conveniens* is generally complex throughout the common law tradition,²³⁵ it is particularly so in the American context: As early as 1929, Paxton Blair wrote that while it was “a principle of exceptions to the doctrine of comity,”²³⁶ employing *forum non conveniens* in a particular case “does not involve ignoring the requirements of comity; on the contrary in some cases only by doing so can the requirements of comity be met.”²³⁷ The argument has also been made that the doctrine does not

COLUM. L. REV. 1, 1 (1929).

- 231 Atlantic Star v. Bona Spes (*In re The Atlantic Star*) [1974] A.C. 436. (“The appellants’ counsel first referred to the law of Scotland, where for a very long time the plea of *forum non conveniens* has been recognised as valid. No doubt it is a desirable objective to diminish remaining differences between the laws of the sister countries.”). See Alexander Reus, *Judicial Discretion: A Comparative View of the Doctrine of Forum Non Conveniens in the United States, the United Kingdom, and Germany*, 16 LOY. OF LOS ANGELES INT'L & COMP. L. J. 455, 459 (1993).
- 232 See, in particular, *The Abidin Daver* [1984] 1 AC 398; Spiliada Maritime Corp v Cansulex Ltd [1987] AC 460.
- 233 Case C-281/02 Owusu v. Jackson, 2005 E.C.R. I-01383 (holding that “[t]he Brussels Convention precludes a court of a Contracting State – whose jurisdiction is established on the basis of Article 2 of that Convention – from exercising a discretionary power to decline to exercise jurisdiction on the ground that a court of a non-Contracting State would be more appropriate for the trial of the action, where the latter court has not been designated by any agreement conferring jurisdiction, has not previously been seised of any claim liable to give rise to *lis pendens* or related actions and the links connecting the dispute with that non-Contracting State are of a kind other than those referred to in Article 16 of the Brussels Convention”). See also Barry J. Rodger, *Forum Non Conveniens Post-Owusu*, 2 J. PRIV. INT'L L. 71, (2006) (discussing the Owusu Court decision to exclude a plea of *forum non conveniens* “where the court’s jurisdiction, under the Brussels Convention, is based on the defender’s domicile and the alternative forum is a non-Contracting State”).
- 234 See generally 2012 O.J. (L 351) 20.12.2012, art. 33–34 (on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters).
- 235 For an analysis of the relationship between comity and *forum non conveniens* in English law, see Thomas Schultz & Jason Mitchenson, *Navigating Sovereignty and Transnational Commercial Law: The Use of Comity by Australian Courts*, 12 J. PRIV. INT'L L. 344, 366 (2016) (discussing the relationship between comity and *forum non conveniens*).
- 236 Blair, *supra* note 231, at 33.
- 237 *Id.* at 6.

fully accord with comity: on the one hand, a decision to dismiss local proceedings in favor of a different forum entails an evaluation of the comparative suitability of a foreign court;²³⁸ on the other hand, if the proceedings before the foreign court have not already been initiated, a decision by the local court not to exercise jurisdiction may amount to “dumping” cases—that is, telling another court what to do.²³⁹

Indeed, while after *International Shoe* American courts swiftly understood the significance of the doctrine, the Supreme Court did not link *forum non conveniens* and comity in its first decisions on the matter, though it must be noted that the Court’s main concern was to overcome the constitutional hurdles to its application. In *Gulf Oil Corp. v. Gilbert*,²⁴⁰ the Court did not play down the constitutional duty to exercise jurisdiction; rather, it held that “[t]he principle of *forum non conveniens* is simply that a court may resist imposition upon its jurisdiction even when jurisdiction is authorized by the letter of a general venue statute.”²⁴¹ By the same token, comity remained, with few exceptions, notably absent from the Supreme Court’s discussion of the doctrine for a long time, with very few timid mentions as exceptions.²⁴²

If lower courts have more often considered the two concepts together, they have largely done so for the purpose of distinguishing one from the other in the most ambiguous cases. In *Laker*, the D.C.

238 This aspect has implications of a certain magnitude in international adjudication, where it might be considered in tension with the principle of *Kompetenz-Kompetenz* — that is, the principle that a tribunal is capable of adjudicating on its own jurisdiction. See Schultz & Ridi, *supra* note 4 (discussing the tension between *forum non conveniens* and the principle of *Kompetenz-Kompetenz* — that is, the principle that a tribunal is capable of adjudicating on its own jurisdiction.); see also McLACHLAN, *supra* note 60, at 454.

239 Briggs, *supra* note 4, at 119.

240 *Gulf Oil Corp. v. Gilbert*, 330 U.S. 501 (1947).

241 *Id.* at 507. According to Calamita, this conclusion followed from the contention that congressional grants of jurisdiction are to be read on the background of the common law, which allow for such discretion. See Calamita, *supra* note 22, at 638. See also *Zivotofsky ex rel. Zivotofsky v. Clinton*, 566 U.S. 189, 206 (2012) (“[A]bstention accommodates considerations inherent in the separation of powers and the limitations envisioned by Article III, which conferred authority to federal courts against a common-law backdrop that recognized the propriety of abstention in exceptional cases.”).

242 See e.g. *Am. Dredging Co. v. Miller*, 510 U.S. 443, 467 (1994) (Kennedy, J., dissenting); *Sinochem Int’l Co. v. Malay. Int’l Shipping Corp.*, 549 U.S. 422 (2007); *Daimler AG v. Bauman*, 134 S. Ct. 746, 771 (2014) (Sotomayor, J., concurring).

Circuit observed that the interest supporting the grant of anti-suit injunctions enjoining the parties from commencing proceedings abroad were to be balanced with comity principles requiring respect for the judicial function of the foreign court.²⁴³ Accordingly, the avoidance of “hardship to parties” and the promotion of “economies of consolidated litigation” would have been better pursued by a motion for dismissal on grounds of *forum non conveniens*. In other cases, comity and *forum non conveniens* have been discussed together, but stressing that they were considered as alternative grounds for dismissal of the same suit.²⁴⁴ However, connection between the two is drawn insofar as the procedural necessity of showing that an alternative and suitable forum might exist is imposed by the circumstance that a court is precluded by comity from passing judgment on a foreign court.²⁴⁵

Yet, *forum non conveniens* is regarded as much narrower a ground for dismissal than comity. As the Second Circuit put it in *Norex Petroleum Ltd. v. Access Indus., Inc.*:

[i]t may well be that a plaintiff that is precluded from litigating a matter in a foreign jurisdiction because of an adverse earlier judgment by its courts will not be able to pursue the claim further in the United States, but the reason for dismissal in such circumstances is our recognition of the foreign judgment in the interest of international comity, not *forum non conveniens*.²⁴⁶

Most importantly—and this truly is the core of the problem—*forum non conveniens* may be considered much narrower than comity because its focus on sovereign interests is less significant. True, it can be argued that comity is the basis of the recognition of the “local interest in having localized controversies decided at

243 *Laker Airways Ltd. v. Sabena, Belgian World Airlines*, 731 F.2d 909, 928 (D.C. Cir. 1984).

244 *Republic of Panama v. BCCI Holdings (Luxembourg) S.A.*, 119 F.3d 935, 951 (11th Cir. 1997).

245 *Abdullahi v. Pfizer, Inc.*, 562 F.3d 163, 189 (2d Cir. 2009) (“The defendant bears the burden of establishing that a presently available and adequate alternative forum exists, and that the balance of private and public interest factors tilts heavily in favor of the alternative forum . . . Absent a showing of inadequacy by a plaintiff, ‘considerations of comity preclude a court from adversely judging the quality of a foreign justice system.’”).

246 *Norex Petroleum Ltd. v. Access Indus., Inc.*, 416 F.3d 146, 159 (2d Cir. 2005).

home.”²⁴⁷ Further, one fails to see how any such distinction could be understood from the phrasing of Scalia’s famous dissent in *Hartford Fire*, which defined the “comity of courts” as declining “to exercise jurisdiction over matters *more appropriately adjudged elsewhere.*”²⁴⁸ It must, however, be observed that comity analysis, while sometimes recognized as “ultimately intertwined with the *forum non conveniens* calculus,”²⁴⁹ is generally regarded by the courts as separate.²⁵⁰

2. Comity as a Coordination Device for Pending or Potential Parallel Proceedings

While “[c]oncurrent jurisdiction does not necessarily entail conflicting jurisdiction,”²⁵¹ differences and tensions do arise from parallel adjudication of the same or similar disputes in different countries. Comity has served as a powerful tool to resolve this problem by providing exceptions to the “virtually unflagging obligation” of American Courts to exercise their jurisdiction.²⁵²

This is particularly common in bankruptcy cases. As the Second Circuit put it in *Cunard S.S. Co. v. Salen Reefer Servs. AB*, this promoted efficiency insofar as:

[t]he granting of comity to a foreign bankruptcy proceeding enables the assets of a debtor to be dispersed in an equitable, orderly, and systematic manner, rather than in a haphazard, erratic or piecemeal fashion. Consequently, American courts

247 *Gulf Oil Corp. v. Gilbert Storage & Transfer Co.*, 330 U.S. 501, 509 (1947). See also Anne-Marie Slaughter, *Judicial Globalization*, 40 VA. J. INT'L L. 1103, 1113 (1999).

248 *Hartford Fire Ins. Co. v. California*, 509 U.S. 764, 817 (1993) (emphasis added).

249 *Ford v. Brown*, 319 F.3d 1302, 1304 n. 3 (11th Cir. 2003) (emphasis added).

250 But see *Dodge*, *supra* note 2, at 2109–10 (“Because the doctrine of *forum non conveniens* allows U.S. courts to restrain their exercise of jurisdiction in deference to foreign courts, it is properly considered a doctrine of international comity.”).

251 *Laker Airways Ltd. v. Sabena, Belgian World Airlines*, 731 F.2d 909, 926 (D.C. Cir. 1984).

252 On the Courts’ obligation to exercise their jurisdiction, see *Colorado River Water Conser. Dist. v. United States*, 424 U.S. 800, 800–01, 817 (1976) (holding that Colorado Federal Court had original jurisdiction to adjudicate all civil matters between US and Indian tribe unless “otherwise provided by [an] Act of Congress,” but dismissal was warranted because dismissal furthered policy of Act against concurrent proceedings.).

have consistently recognized the interest of foreign courts in liquidating or winding up the affairs of their own domestic business entities.²⁵³

The grant of comity, however, was not automatic, but subject to certain conditions. In *Allstate Life Ins. Co. v. Linter Grp. Ltd.*, the same Court set out a number of factors as “indicia of procedural fairness,” including:

- (1) whether creditors of the same class are treated equally in the distribution of assets; (2) whether the liquidators are considered fiduciaries and are held accountable to the court; (3) whether creditors have the right to submit claims which, if denied, can be submitted to a bankruptcy court for adjudication; (4) whether the liquidators are required to give notice to the debtors’ potential claimants; (5) whether there are provisions for creditors’ meetings; (6) whether a foreign country’s insolvency laws favor its own citizens; (7) whether all assets are marshalled before one body for centralized distribution; and (8) whether there are provisions for an automatic stay and for the lifting of such stays to facilitate the centralization of claims.²⁵⁴

As the same Court held years later, comity is thus afforded “to foreign bankruptcies only if those proceedings do not violate the laws or public policy of the United States.”²⁵⁵

These conclusions reflect an understanding of comity which is largely derived from *Hilton*. It follows that the same logic governing the recognition of judgments applies, to a significant extent, to the question of deference to foreign proceedings. Outside of the bankruptcy context, the Eleventh Circuit openly admitted as much in *Turner Entm’t Co. v. Degeto Film GmbH*, eventually concluding that comity concerns included:

253 *Cunard S.S. Co. Ltd. v. Salen Reefer Servs. AB*, 773 F.2d 452, 458 (2d Cir. 1985).

254 *Allstate Life Ins. Co. v. Linter Grp. Ltd*, 994 F.2d 996, 999 (2d Cir. 1993).

255 *Finanz AG Zurich v. Banco Economico S.A.*, 192 F.3d 240, 246 (2d Cir. 1999) (citing *Cunard*, 773 F.2d 452 at 460).

(1) whether the judgment was rendered via fraud . . . (2) whether the judgment was rendered by a competent court utilizing proceedings consistent with civilized jurisprudence, see *id.*; and (3) whether the foreign judgment is prejudicial, in the sense of violating American public policy because it is repugnant to fundamental principles of what is decent and just.²⁵⁶

Interestingly, however, the Court went further, adding that “the relative strengths” of the two countries’ interests were also relevant to the comity analysis: In the case at issue, the Court relied on this point to conclude that Germany had a greater interest in the determination of the dispute.²⁵⁷ Complications may, of course, arise. For example, in *Jota v. Texaco, Inc.*²⁵⁸ the Second Circuit dealt with a shift in the position of the other sovereign party involved, Ecuador, which first opposed and then advocated for the jurisdiction of United States courts.²⁵⁹

The existence of governmental interests is considered vital for the purposes of comity-based abstention. In *Royal & Sun All. Ins. Co. of Canada v. Century Int'l Arms, Inc.* it was held that:

[C]ircumstances that routinely exist in connection with parallel litigation cannot reasonably be considered exceptional circumstances, and therefore the mere existence of an adequate parallel action, by itself, does not justify the dismissal of a case on grounds of international comity abstention. Rather, additional circumstances must be present—such as a foreign nation’s interest in uniform bankruptcy proceedings—that outweigh the district court’s

256 *Turner Ent Co. v. Degeto Film GmbH*, 25 F.3d 1512, 1519 (11th Cir. 1994) (internal citations omitted).

257 *Id.* at 1521. (“There appears to be no clear federal interest in trying this case. Certainly much is at stake in this litigation for both parties. However, the public interest in the litigation is more conspicuous in Germany”). Comity may also explain the reasoning in *Bi v. Union Carbide Chems. & Plastics Co.*, 984 F.2d 582, 585–86 (2d Cir. 1993) (recognizing the Indian Government’s interest, reflected in its domestic legislation, to be the only entity with standing to sue).

258 *Jota v. Texaco, Inc.*, 157 F.3d 153 (2d Cir. 1998).

259 *Id.* at 160.

general obligation to exercise its jurisdiction.²⁶⁰

The Court reached its conclusion by reference to the Supreme Court's dictum that “[a]bstention from the exercise of federal jurisdiction is the exception, not the rule.”²⁶¹ Yet, even the circuits that do not follow this approach adopt a similar methodology when confronted with parallel proceedings pending abroad. For example, in *Ungaro-Benages v. Dresdner Bank AG*,²⁶² the Eleventh Circuit drew a line between the situations where the application of the comity doctrine has a “retrospective” character and those in which it is “prospective”:

When applied retrospectively, domestic courts consider whether to respect the judgment of a foreign tribunal or to defer to parallel foreign proceedings... When applied prospectively, domestic courts consider whether to dismiss or stay a domestic action based on the interests of our government, the foreign government and the international community in resolving the dispute in a foreign forum . . . Applied prospectively, federal courts evaluate several factors, including the strength of the United States' interest in using a foreign forum, the strength of the foreign governments' interests, and the adequacy of the alternative forum.²⁶³

Accordingly, a “retrospective” approach is warranted when proceedings are already pending before another forum. It is only before proceedings are initiated that a “prospective” approach may apply. *Ungaro-Benages* concerned a suit by a descendant of the heir to a German company against German banks that, the plaintiff claimed, had stolen the stock belonging to Jewish heirs in aryanization processes. In 2000, President Clinton entered into an agreement with Germany to establish a foundation to hear claims from victims

260 Royal & Sun All. Ins. Co. of Can. v. Century Int'l Arms, Inc., 466 F.3d 88, 95 (2nd Cir. 2006).

261 Colorado River Water Conservation Dist. v. United States, 424 U.S. 800, 813 (1976) (holding that the District Court was correct in finding that the conditions for abstaining were not present in the instant case). See also Dodge, *supra* note 2, at 2112–13.

262 *Ungaro-Benages v. Dresdner Bank AG*, 379 F.3d 1227 (11th Cir. 2004).

263 *Id.* at 1238.

of the Nazi regime.²⁶⁴ The Eleventh Circuit found that the United States' support for the Foundation as the exclusive forum for such claims, along with the German government's interests "in having the Foundation be the exclusive forum for these claims in its efforts to achieve lasting legal peace with the international community," and the adequacy of the Foundation as a forum all supported dismissal.²⁶⁵

In some cases, the very act of adjudication of a dispute might be regarded, in and by itself, as an extremely unfriendly act. Indeed, "extreme cases might be imagined where a foreign sovereign's interests were so legitimately affronted by the conduct of litigation in a United States forum that dismissal is warranted without regard to the defendant's amenability to suit in an adequate foreign forum."²⁶⁶

These might be cases in which a court is concerned about the far-reaching effects of entertaining a suit with foreign relations implications. All the same, such implications might have to be balanced with the policies of the forum state. *Bigio v. Coca-Cola Co.*, a Second Circuit case, concerned Canadian citizens who alleged that the unlawful expropriation of their Egyptian corporation by the Egyptian government was carried out because of their Jewish lineage.²⁶⁷ The corporation was later acquired by a company, which the plaintiffs proceeded to sue on the grounds that it had done so in full knowledge of the unlawfulness of the expropriation.²⁶⁸ The Court responded to a plea of dismissal on grounds of comity that "the only issue of international comity properly raised here is whether adjudication of this case by a United States court would

264 *Ungaro-Benages*, 379 F.3d 1227 at 1231.

265 *Id.* at 1239.). But see *GDG Acquisitions, LLC v. Gov't of Belize*, 749 F.3d 1024, 1026, 1031 (11th Cir. 2014) (holding that dismissal on grounds of international comity was appropriate). A similar standard as *GDG Acquisitions* has been adopted, though perhaps not fully understood, in the Fifth Circuit decision *Perforaciones Exploracion Y Produccion v. Maritimas Mexicanas, S.A. de C.V.*, 356 F. App'x 675, 681 (5th Cir. 2009) ("Dismissal of a suit on international comity grounds may sometimes be appropriate when there is litigation pending in a foreign forum or, even absent such litigation, when allowing a case to proceed in the United States would intrude on the interests of a foreign government."). However, the Third Circuit has remained skeptical of "prospective comity" analyses. See e.g. *Gross v. German Found. Indus. Initiative*, 456 F.3d 363, 393 (3d Cir. 2006) ("Absent true conflicts, a judgment from a foreign court, or parallel proceedings in a foreign forum, rarely have United States courts abstained from deciding the merits of a case on international comity grounds.").

266 *Jota v. Texaco, Inc.*, 157 F.3d 153, 160 (2d Cir. 1998).

267 *Bigio v. Coca-Cola Co.*, 448 F.3d 176 (2d Cir. 2006).

268 *Id.* at 178.

offend ‘amicable working relationships’ with Egypt.”²⁶⁹ In the case at issue, the Court could hold that comity did not justify dismissing the local proceedings by reference to the fact that the Egyptian Government had “never raised the slightest objection to adjudication of the instant controversy by United States courts” and its prior determination that “resolution of this case by United States courts will “not likely impact on international relations.”²⁷⁰

However, a different situation was faced in *Khulumani v. Barclay National Bank Limited*,²⁷¹ a suit brought under the Alien Tort Statute by South African plaintiffs claiming that the defendants had “actively and willingly collaborated with the government of South Africa in maintaining a repressive, racially based system known as ‘apartheid,’ which restricted the majority black African population in all areas of life while providing benefits for the minority white population.”²⁷² The case was an extremely political one, with statements of interest by the government of both countries involved.²⁷³ Without delving into the complicated history of the case,²⁷⁴ the case is notable because the District Court eventually held on remand that “[i]nternational comity comes into play only when there is a true conflict between American law and that of a foreign jurisdiction.”²⁷⁵ The Court further held that even if there was a “true conflict,” the decision whether to dismiss on comity grounds depends on the degree of legitimate offense to the foreign sovereign, steps the foreign sovereign may have taken to address the issues in the litigation, and the extent of the United States’ interest in the underlying issues.²⁷⁶ The Court further observed that:

[t]he absence of conflict between this litigation and the TRC process is fatal to the argument that international comity requires dismissal. The TRC process was not exclusive—by its terms, only upon a grant of amnesty was the right of ‘victims and/or their families to institute criminal and/or civil proceedings...’ Plaintiffs have now come to this Court

269 *Id.*

270 *Id.*

271 *Khulumani v. Barclay Nat'l Bank Ltd.*, 504 F.3d 254 (2d Cir. 2007).

272 *Id.* at 258.

273 *Id.* at 259.

274 For further information on this topic, see Childress, *supra* note 2, at 53–54.

275 *In re S. African Apartheid Litig.*, 617 F. Supp. 2d 228, 283 (S.D.N.Y. 2009).

276 *Id.*

to exercise their rights.

Defendants do not argue—and South Africa has not claimed—that ‘an adequate forum exists in the objecting nation’ . . . Nor does this litigation conflict with the *goals* of the TRC process; thus it is not an ‘extreme case’ requiring dismissal even without the existence of an alternative forum for the plaintiffs’ claims.²⁷⁷

The problem with this reasoning is evident: As Childress has observed, it was not quite clear from *Hartford Fire* that a true conflict analysis should guide adjudicatory comity cases.²⁷⁸ Even if it did, it is arguable that it should focus on a conflict between sovereigns—something that a *Hartford Fire* type of reasoning does not seem to allow.

This point was precisely one of the most critical issues of the recent Ninth Circuit decision in *Mujica v. AirScan Inc.*,²⁷⁹ hailed by many as a missed opportunity that contributed to further confusion on the issue. This too was an action brought under the Alien Tort Statute and the Torture Victim Prevention Act, this time by Colombian nationals against an American corporation and its private security firm, for the defendants’ alleged “complicity in the bombings” of their village.²⁸⁰ No other proceedings were pending elsewhere and the Court endeavored to elucidate the concept of comity and ultimately dismissed the suit on that basis.²⁸¹ The Court concluded—by any measure, correctly—that “*Hartford Fire* does *not* require proof of a ‘true conflict’ as a prerequisite for invoking the doctrine of comity, at least in a case involving adjudicatory comity.”²⁸² It reached this conclusion by reading a number of precedents to this effect,²⁸³ and ultimately expanded the Eleventh Circuit’s test developed in *Ungaro-*

277 *Id.* at 285–86 (citing 6 *Final Report of the Truth and Reconciliation Commission* 100. Accord TRC Act § 20(7)).

278 Childress, *supra* note 2, at 55.

279 *Mujica v. AirScan Inc.*, 771 F.3d 580 (9th Cir. 2014).

280 *Id.* at 584.

281 *Id.* at 585–86, 596–97, 603–615 (analysing each factor weighing of and against comity and ultimately concluding that factors favoured application of adjudicatory comity).

282 *Id.* at 600.

283 *Id.* at 600–01.

Benages to formulate its own.²⁸⁴ In particular, where the latter test focused on “the United States’ interest in using a foreign forum, the strength of the foreign governments’ interests, and the adequacy of the alternative forum,”²⁸⁵ the Court in *Mujica* attempted to clarify the first element with the prescriptive comity factors the same Court had considered in *Timberlane*, and which made their way into § 403 of the Restatement.²⁸⁶ The Court held that they constituted “a general list of indicia to which we may look when weighing U.S. and foreign interests and the adequacy of the alternative forum.”²⁸⁷ This choice has been harshly criticized, and it has been observed that such an analysis seems to be the result of a misunderstanding.²⁸⁸ Yet, a second look reveals that the most important component of the Court’s comity analysis might have been the deference granted to the Executive’s Statement of Interest, thus seriously downplaying the significance of the other elements in the picture.²⁸⁹

3. Anti-Suit Injunctions

Finally, a few words must be spent on the topic of anti-suit injunctions. These may be considered the flipside of *forum non conveniens* and comity-based abstention doctrines insofar as their need arises when a court does *not* dismiss the local suit in favor of foreign proceedings. Contrary to the English tradition, where “[i]t is easy to take anti-suit injunctions for granted,”²⁹⁰ American courts have on the whole granted the remedy sparingly. That federal courts have the power to grant them is not at all controversial.²⁹¹ It has long

284 *Ungaro-Benages v. Dresdner Bank AG*, 379 F.3d 1227, 1238 (11th Cir. 2004).

285 *Id.*

286 *Mujica v. AirScan Inc.*, 771 F.3d 580, 603 (9th Cir. 2014).

287 *Id.* at 604.

288 See, e.g., William Dodge, *International Comity Running Amok*, JUST SECURITY (Feb. 3, 2015, 9:22 AM), <https://www.justsecurity.org/19640/international-comity-run-amok/> (describing the analysis in *Mujica* as “oddly engrafted” and “broad and unnecessary”); Roger Alford, *The Ninth Circuit’s Muddled Comity Analysis in Mujica*, OPINIO JURIS (Nov. 21, 2014, 1:40 PM), <http://opiniojuris.org/2014/11/21/ninth-circuits-muddled-comity-analysis-mujica/>.

289 *Mujica*, 771 F.3d at 610 (“Accordingly, we ‘give serious weight to the Executive Branch’s view of [this] case’s impact on foreign policy,’ and we conclude that the United States’ interest in having the case adjudicated exclusively in Colombia is strong.” (citing *Sosa v. Alvarez-Machain*, 542 U.S. 692, 733 n.21 (2004))).

290 Richard Fentiman, *Anti-Suit Injunctions – Comity Redux?*, 71 CAMB. L. J. 273, 273 (2012).

291 See, e.g., *Seattle Totems Hockey Club, Inc. v. Nat'l Hockey League*, 652 F.2d 852, 855 (9th Cir. 1981) (“A federal district court with jurisdiction over the

been recognized, however, that “the power to enjoin a party before it from pursuing litigation before a foreign tribunal . . . often effectively restricts the jurisdiction of the foreign tribunal and should therefore be used sparingly.”²⁹²

Indeed, the grant of an anti-suit injunction can be a delicate matter: While English courts and writers often justify their existence with the need to enforce an agreement not to bring proceedings before a foreign court, there is good evidence for the proposition that the remedy does interfere with the sovereign act of adjudication—though the extent to which it does, depending on whether a foreign court has or has not been seized at the time of the injunction, might be a matter of degrees.²⁹³ American courts, on the other hand, are generally more mindful of the implications of any such gesture. As early as 1849, the Supreme Court held in *Peck v. Jenness*²⁹⁴ that “[t]he fact . . . that an injunction issues only to the parties before the court, and not to the court, is no evasion of the difficulties that are the necessary result of an attempt to exercise that power over a party who is a litigant in another and independent forum.”²⁹⁵ In other words, courts granting such injunctions too liberally would raise serious comity concerns.²⁹⁶

This conclusion reflects the belief that “[c]oncurrent jurisdiction does not necessarily entail conflicting jurisdiction”—a statement that holds true with reference to both jurisdiction to prescribe and to adjudicate.²⁹⁷ Accordingly, concurrent proceedings represent the rule, and, since they “are ordinarily tolerable, the initiation before a foreign court of a suit concerning the same parties and issues as a suit already pending in a United States court does not, without more, justify enjoining a party from proceeding in the foreign forum.”²⁹⁸ In other words, parallel proceedings normally

parties has the power to enjoin them from proceeding with an action in the courts of a foreign country.”).

292 *United States v. Davis*, 767 F.2d 1025, 1038 (2d Cir. 1985).

293 *Briggs, supra* note 4, at 125–26.

294 *Peck v. Jenness*, 48 U.S. 612 (1849).

295 *Id.* at 625.

296 *China Trade & Dev. Corp. v. M.V. Choong Yong*, 837 F.2d 33, 35 (2d Cir. 1987) (“The fact that the injunction operates only against the parties, and not directly against the foreign court, does not eliminate the need for due regard to principles of international comity.”).

297 *Laker Airways Ltd. v. Sabena, Belgian World Airlines*, 731 F.2d 909, 926 (D.C. Cir. 1984).

298 *China Trade & Dev. Corp.*, 837 F.2d 33 at 36.

represent a necessary condition for the grant of the remedy, but seldom suffice alone, even when their concurrence might produce an “embarrassing race to judgment” and “potentially inconsistent adjudications.”²⁹⁹ Indeed, comity considerations militating against the grant of anti-suit injunctions are no minor hurdle: The Sixth Circuit referred to the issuance of such an order as “a drastic step.”³⁰⁰ The First Circuit, too, held as much in the often-cited decision in *Quaak v. Klynveld Peat Marwick Goerdeler Bedrijfsrevisoren*,³⁰¹ where it held that “considerations of international comity must be given substantial weight—and those considerations ordinarily establish a rebuttable presumption against the issuance of an order that has the effect of halting foreign judicial proceedings.”³⁰²

To some extent, the grant of anti-suit injunctions tends to follow the same logic of the recognition of judgments and, even more closely, comity-driven abstention doctrines.³⁰³ While conceding that no hard and fast rules were available, the D.C. Circuit observed that:

[t]he equitable circumstances surrounding each request for an injunction must be carefully examined to determine whether, in light of the principles outlined above, the injunction is required to prevent an irreparable miscarriage of justice. Injunctions are most often necessary to protect the jurisdiction of the enjoining court, or to prevent the litigant’s evasion of the important public policies of the forum.³⁰⁴

Over the years, the Ninth Circuit has elaborated a more principled—and seemingly less conservative³⁰⁵—approach to anti-suit injunctions, based on a three-pronged test, which the court

299 *Laker Airways Ltd.*, 731 F.2d at 928.

300 *Gau Shan Co. v. Bankers Tr. Co.*, 956 F.2d 1349, 1354 (6th Cir. 1992).

301 *Quaak v. Klynveld Peat Marwick Goerdeler Bedrijfsrevisoren*, 361 F.3d 11 (1st Cir. 2004).

302 *Id.* at 12.

303 *Laker Airways Ltd.*, 731 F.2d at 931 (“This principle is similar to the rule that a foreign judgment not entitled to full faith and credit under the Constitution will not be enforced within the United States when contrary to the crucial public policies of the forum.”).

304 *Id.* at 927.

305 *Goss Int'l Corp. v. Man Roland Druckmaschinen Aktiengesellschaft*, 491 F.3d 355, 360 (8th Cir. 2007).

described in full in the 2012 case *Microsoft Corp. v. Motorola, Inc.*³⁰⁶ First, the court must consider whether the parties and the issues are the same and “whether or not the first action is dispositive of the action to be enjoined.”³⁰⁷ Second, the court must determine whether:

[T]he foreign litigation would (1) frustrate a policy of the forum issuing the injunction; (2) be vexatious or oppressive; (3) threaten the issuing court’s in rem or quasi in rem jurisdiction; or (4) where the proceedings prejudice other equitable considerations.³⁰⁸

Finally, the court must “assess whether the injunction’s impact on comity is tolerable.”³⁰⁹ In this regard, the Court reasoned that there are significant differences between cases where the enforcement of a choice of forum agreement is sought—unlikely, in principle, to infringe comity—and those politically sensitive situations in which foreign relations implications are expected, and perhaps even an object of a governmental statement of interest.³¹⁰ For those cases situated at neither end of the spectrum, the comity analysis may take into account other elements, such as whether the plaintiff initiated the foreign suit at a later moment for seemingly abusive purposes.³¹¹

Irrespective of the differences between circuits, it is easy to understand how the doctrine of comity governs the grant of anti-suit injunctions. In broad terms, the two main grounds to grant them are the protection of a court’s own jurisdiction and the prevention of abusive forum selection designed to escape fundamental American policies. The first ground explains the more relaxed attitude in granting the remedy for *in rem* or *quasi-in rem* actions. As it was held in a Sixth Circuit case:

306 *Microsoft Corp. v. Motorola, Inc.*, 696 F.3d 872, 881 (9th Cir. 2012). *See also E. & J. Gallo Winery v. Andina Licores S.A.*, 446 F.3d 984, 991–94 (9th Cir. 2006).

307 *Microsoft Corp. v. Motorola, Inc.*, 696 F.3d 872, 881 (9th Cir. 2012) (citing *E. & J. Gallo Winery v. Andina Licores S.A.*, 446 F.3d 984, 991 (9th Cir. 2006)).

308 These are commonly referred to as the “Unterweser factors.” *See In re Unterweser Reederei GMBH*, 428 F.2d 888, 890 (5th Cir. 1970).

309 *Microsoft Corp.*, 696 F.3d at 881 (citing *E. & J. Gallo Winery v. Andina Licores S.A.*, 446 F.3d 984, 991 (9th Cir. 2006)).

310 *Id.* at 887.

311 *Id.*

Where jurisdiction is based on the presence of property within the court's jurisdictional boundaries, a concurrent proceeding in a foreign jurisdiction poses the danger that the foreign court will order the transfer of the property out of the jurisdictional boundaries of the first court, thus depriving it of jurisdiction over the matter.³¹²

The second is a public policy ground, and in that it does not differ much from the phenomenon of recognition of foreign judgments. The latter, as the D.C. Circuit put it in *Laker*, is a strict standard, but that governing anti-suit injunctions is even stricter in recognition of the much greater inference with the judicial process of another country.³¹³ An evasion of public policy warranting the remedy cannot thus be found in "the availability of slight advantages in the substantive or procedural law,"³¹⁴ such as, for example, the unavailability of a treble damages remedy.³¹⁵ Interestingly, courts instead grant anti-suit injunctions because they frustrate a United States policy favoring forum selection clauses³¹⁶ and "the liberal enforcement of arbitration clauses."³¹⁷

D. "Executive" or "Sovereign-Party" Comity

There is one more type of comity, which has traditionally been assigned the confusing label of "executive."³¹⁸ The adjective has

312 *Gau Shan Co. v. Bankers Trust Co.*, 956 F.2d 1349, 1356 (6th Cir. 1992).

313 *Laker Airways Ltd. v. Sabena, Belgian World Airlines*, 731 F.2d 909, 931–32 (D.C. Cir. 1984).

314 *Laker Airways Ltd.*, 731 F.2d at 931 n.73.

315 *Gau Shan Co.*, 956 F.2d at 1358.

316 *E. & J. Gallo Winery v. Andina Licores S.A.*, 446 F.3d 984, 993 (9th Cir. 2006) ("We hold that Andina's pursuit of litigation in Ecuador, in violation of the forum selection clause, frustrates a policy of the United States courts . . .").

317 *Paramedics Electromedicina Comercial, Ltda v. GE Med. Sys. Info. Techs., Inc.*, 369 F.3d 645, 654 (2d Cir. 2004). The opinion cites *Mitsubishi Motors Corp. v. Soler Chrysler-Plymouth, Inc.*, 473 U.S. 614, 629 (1985), holding that "[c]oncerns of international comity, respect for the capacities of foreign and transnational tribunals, and sensitivity to the need of the international commercial system for predictability in the resolution of disputes require that we enforce the parties' agreement, even assuming that a contrary result would be forthcoming in a domestic context."

318 See, e.g., Childress, *supra* note 2, at 47; see also Molly Warner Lien, *The Cooperative and Integrative Models of International Judicial Comity: Two Illustrations Using Transnational Discovery and Breard Scenarios*, 50 CATH. U. L. REV. 591, 595 n.9 (2000).

been employed to convey the idea that this type of comity is intended to command deference “with regard to the executive acts of foreign countries,”³¹⁹ and usually identifies the Act of State doctrine, foreign sovereign immunity and the privilege of foreign governments to bring suit in United States courts. Whether the expression achieves its objective is, however, doubtful at best: In fact, this terminology adds to the confusion stemming from the likely involvement of the Executive Branch in the cross-border affairs that normally justify the use of such types of comity. To avert the problem, William Dodge has proposed the alternative definition of “sovereign party comity,” meant to identify those cases where deference is given to foreign governments as litigants.³²⁰ This choice too is problematic, as it comes with the corollary of not covering the Act of State doctrine, which Dodge classifies as a principle of recognition under the heading of prescriptive comity.³²¹

For our purposes, the distinction is not of disproportionate consequence. Accordingly, we acknowledge the difficulty in capturing the essence of these doctrines and adopt, for reasons of intelligibility, the traditional approach of grouping the three of them together under the same heading.

1. The Act of State Doctrine

The “Act of State Doctrine” requires American courts not to question the validity of an act concluded by a foreign government in its territory. For greater precision, as the Court put it in *First National City Bank v. Banco Nacional de Cuba*, “the doctrine precludes any review whatever of the acts of the government of one sovereign State done within its own territory by the courts of another sovereign State.”³²² The classical statement of the doctrine is generally understood to be contained in the 1897 decision *Underhill v. Hernandez*, where the Supreme Court affirmed that:

Every sovereign state is bound to respect the independence of every other sovereign state, and the courts of one country will not sit in judgment on the acts of the government of another, done within its own territory. Redress of grievances by reason of

319 Ramsey, *supra* note 11, at 937.

320 Dodge, *supra* note 2, at 2099.

321 *Id.* at 2079.

322 *First Nat'l City Bank v. Banco Nacional de Cuba*, 406 U.S. 759, 763 (1972).

such acts must be obtained through the means open to be availed of by sovereign powers as between themselves.³²³

The doctrine has traditionally been considered to stem from considerations of international comity, and justifiably so. As the Supreme Court put it, “[t]o permit the validity of the acts of one sovereign state to be reexamined and perhaps condemned by the courts of another would very certainly ‘imperil the amicable’ relations between governments and vex the peace of nations.”³²⁴ But what does comity require in this context? The answer is not as straightforward as is sometimes argued.

It has been suggested that the interests to be protected by the doctrine have evolved, moving from mutual convenience, to encompass respect for sovereignty, and, finally, consideration for the foreign relations interests of the United States and prerogatives of the political branches.³²⁵ The foreign relations element of the doctrine is particularly evident in *Banco Nacional de Cuba v. Sabbatino*, where the Court considered a dispute arising from an expropriation by the Cuban government to the detriment of an American company, made all the more interesting by the circumstance that the suit was brought by the former.³²⁶ The defendant had argued that this element, along with the fact that the expropriation was to be considered unlawful under international law, would have rendered the doctrine inapplicable to the case.³²⁷ The Court, however, held otherwise, stating that:

The doctrine as formulated in past decisions expresses the strong sense of the Judicial Branch that its engagement in the task of passing on the validity of foreign acts of state may hinder rather than further this country’s pursuit of goals both for itself

323 *Underhill v. Hernandez*, 168 U.S. 250, 252 (1897).

324 *Oetjen v. Cent. Leather Co.*, 246 U.S. 297, 304 (1918).

325 Dodge, *supra* note 2, at 2096; Paul, *Comity in International Law*, *supra* note 9, at 30–31.

326 *Banco Nacional de Cuba v. Sabbatino*, 376 U.S. 398 (1964).

327 Indeed, this hypothesis is now covered by the Foreign Assistance Act of 1964 (amended 2000), 28 U.S.C. § 1605 (2012), which excludes the application of the act of state doctrine for declining jurisdiction over confiscations of property violating international law after January 1, 1959. *Banco Nacional de Cuba v. Sabbatino*, 376 U.S. at 423.

and for the community of nations as a whole in the international sphere.³²⁸

In this context, comity serves as the basis of the rule.³²⁹ This respect for other sovereign entities, though, is not just a textbook definition of comity itself, but rather a consequence of precise policy determinations. It is precisely in this way that courts begin to pay deference, rather than to the acts of a foreign state as such, to these policy concerns and to the branch of government best suited to pursue them: the Executive.³³⁰ The Court expanded on these conclusions twice. In *First National City Bank*, it confirmed that the act of state doctrine, much like sovereign immunity,³³¹ was “judicially created to effectuate general notions of comity among nations and among the respective branches of the Federal Government.”³³² The corollary of this statement was that the doctrine was not meant to be “inflexible,” but rather as malleable as the needs of the executive

328 376 U.S. at 423.

329 Though, according to Ramsay, the fact that “conflict with foreign territorial acts carries strong foreign policy implications so that in the U.S. system courts presume that such decisions lie with the so-called political branches. However, it adds little to import the label ‘comity’ into the discourse of the act of state doctrine, which can be described and explained without it.” Ramsey, *supra* note 11, at 916.

330 See *Sabbatino*, 376 U.S. 398 at 423 (“The act of state doctrine does, however, have ‘constitutional’ underpinnings. It arises out of the basic relationships between branches of government in a system of separation of powers. It concerns the competency of dissimilar institutions to make and implement particular kinds of decisions in the area of international relations.”) According to Paul, the Court’s desire to give the government the widest possible discretion in dealing with communist states prompted the ironic and contradictory result of replacing deference to party autonomy with deference to the Executive. *See also* Paul, *Transformation of International Comity*, *supra* note 9 at 31–32.

331 However, “[u]nlike a claim of sovereign immunity, which merely raises a jurisdictional defense, the act of state doctrine provides foreign states with a substantive defense on the merits. Under that doctrine, the courts of one state will not question the validity of public acts (acts *jure imperii*) performed by other sovereigns within their own borders, even when such courts have jurisdiction over a controversy in which one of the litigants has standing to challenge those acts.” *Republic of Austria v. Altmann*, 541 U.S. 677, 700 (2004). *See also* *Samantar v. Yousuf*, 560 U.S. 305, 322 (2010). Most importantly, as Justice Marshall pointed out in a dissent, the two “differ fundamentally in their focus and in their operation. Sovereign immunity accords a defendant exemption from suit by virtue of its status. By contrast, the act of state doctrine exempts no one from the process of the court.” *See also* *Alfred Dunhill of London, Inc. v. Republic of Cuba*, 425 U.S. 682, 725–26 (1976).

332 *First Nat'l City Bank v. Banco Nacional de Cuba*, 406 U.S. 759, 762 (1972).

required. In the case at issue, the Executive Branch had “expressly stated that an inflexible application of the act of state doctrine by this Court would not serve the interests of American foreign policy.”³³³ By the same token, the application of the rule out of deference would have been illogical if the interested party—the Executive—had counselled otherwise.³³⁴ According to Koh, by explicitly linking the Act of State Doctrine to separation of powers, *Sabbatino* implied that determinations regarding the legality of foreign state acts are quasi-political questions, whose decision is appropriately confided in the Executive or Congress, not the court.”³³⁵ Almost two decades later, *W.S. Kirkpatrick & Company, Inc., et al. v. Environmental Tectonics Co., International*³³⁶ offered an opportunity to discuss the shifting foundations of the doctrine. In its opinion, delivered by Justice Scalia,³³⁷ the Court was unambiguous in describing it as “a consequence of domestic separation of powers, reflecting ‘the strong sense of the Judicial Branch that its engagement in the task of passing on the validity of foreign acts of state may hinder’ the conduct of foreign affairs,” thus confirming the first part of the comment above.³³⁸ And yet, while no question was made expressly of the need to defer to the Executive, a reading of *Kirkpatrick* suggests a narrower interpretation of the Act of State doctrine: it may only be employed when there is one specific act of state, and that the act has bearing on the outcome of the dispute is less obvious a remark than might appear at first glance.³³⁹ In other words, the underlying policies giving rise to the doctrine cannot, by themselves, give rise to the application of the doctrine lacking an act of the foreign government the validity of which is at issue: The “embarrassment” of any government is either devoid of any relevance, or in any event fails, in and by itself, to constitute a sufficient reason for a court to decline its jurisdiction.³⁴⁰

For our purposes, the way the doctrine operates need not be clarified further: But what is, then, the function of comity in this

333 *Id.* at 768.

334 *Id.* at 769–70.

335 Harold Koh, *Transnational Public Law Litigation*, 100 YALE L. J. 2347, 2362 (1991).

336 *Kirkpatrick & Co. v. Envtl. Tectonics Corp.*, 493 U.S. 400 (1990).

337 Interestingly, Scalia was a vocal supporter of separation of powers doctrines. See Koh, *supra* note 336, at 2366.

338 *Kirkpatrick*, 493 U.S. at 404.

339 *Id.* at 406.

340 *Id.* at 410.

context? In a recent study, William Dodge has argued that the Act of State doctrine should not be classified as a form of “executive comity,” but must on the contrary be understood as constituting a principle of recognition, a *species* of prescriptive comity. The remark is not devoid of merit, but tends to describe the result of the doctrine rather than its basis, and happens to be no news at all. As the Court noted in *Sabbatino*, “[t]he act of state doctrine . . . although it shares with the immunity doctrine a respect for sovereign states, concerns the limits for determining the validity of an otherwise applicable rule of law.”³⁴¹ In other words, while the end product of the rule is the recognition of a foreign act, the fact that this recognition may extend to manifestly unlawful actions does render the element of respect for foreign states and relations preponderant.³⁴² Further, as the Court put it in *First National City Bank*, “the act of state doctrine justifies its existence primarily on the basis that juridical review of acts of state of a foreign power could embarrass the conduct of foreign relations by the political branches.”³⁴³ True, Kirkpatrick has, at the very least, mitigated the significance of comity for the use of the doctrine—and indeed, as far as its application is concerned, Joel Paul is absolutely correct in observing that “[t]he risk of embarrassing the executive is a curious rationale for a conflicts principle . . .”³⁴⁴ The point is that, if “embarrassment” belongs within the semantic spectrum of comity, the contrary is not necessarily true. In fact, the comity that explains the origin of the act of state doctrine is very much the same comity that remains relevant for the modern-day life of the institution. Indeed, it was through comity that American courts created the doctrine in their quest to “accommodate respect for foreign sovereignty with growing American intercourse with other nations.”³⁴⁵ In other words, comity, intended as a tool to promote successful political and commercial relations, supports the idea of deference to the Executive, assumed as the branch capable of best pursuing these objectives.

2. *Sovereign Immunity*

341 Banco Nacional de Cuba v. Sabbatino, 376 U.S. 398, 438 (1964).

342 See *Id.* at 458–59 (White, J., dissenting) (“Where a clear violation of international law is not demonstrated, I would agree that principles of comity underlying the act of state doctrine warrant recognition and enforcement of the foreign act.”).

343 First Nat'l City Bank v. Banco Nacional de Cuba, 406 U.S. 759, 765 (1972).

344 Paul, *The Transformation of International Comity*, *supra* note 9 at 32.

345 Koh, *supra* note 336, at 2257.

Doctrines of sovereign immunity have long been recognized to be founded on comity. True, the early landmark case *The Schooner Exchange v. McFaddon*, often labelled as “[t]he classic case illustrating the relationship between territorial jurisdiction and sovereign immunity,”³⁴⁶ does not refer to comity as such.³⁴⁷ While affirming that “[t]he jurisdiction of the nation within its own territory” was “necessarily exclusive and absolute” and “susceptible of no limitation not imposed by itself,” Chief Justice Marshall observed that exceptions did exist, though they ultimately had to be traced to the consent of the nation, “either express or implied. In the latter case, it is less determinate, exposed more to the uncertainties of construction; but, if understood, not less obligatory.”³⁴⁸ Marshall’s skillful and balanced language in putting forward an exception to sovereignty while reaffirming its absoluteness,³⁴⁹ along with his oblique reference to the language of the law of nations,³⁵⁰ have somehow diminished the significance of comity elements in the case.³⁵¹ And yet, these elements were clearly there:

The world being composed of distinct sovereignties, possessing equal rights and equal independence, whose mutual benefit is promoted by intercourse with each other, and by an interchange of those good offices which humanity dictates and its wants require, all sovereigns have consented to a relaxation in practice, in cases under certain peculiar circumstances, of that absolute and complete jurisdiction within their respective territories which sovereignty confers.³⁵²

A number of authorities, American and foreign, support the proposition that the *The Exchange* was founded on comity. In the English case *The Parlement Belge*, for example, the Court of Appeal cited the American decision to support the proposition that as “a consequence of the absolute independence of every sovereign

³⁴⁶ MALCOLM N. SHAW, INTERNATIONAL LAW 699 (6th ed. 2008).

³⁴⁷ *Schooner Exchange v. McFaddon*, 11 U.S. 116 (1812).

³⁴⁸ *Id.* at 136.

³⁴⁹ PHILIP C. JESSUP, TRANSNATIONAL LAW 54 (Yale University Press 1956).

³⁵⁰ For a comparison with similar Grotian expressions see similar language in Austin *supra*, note 62, at 144.

³⁵¹ John G. Hervey, *The Immunity of Foreign States When Engaged in Commercial Enterprises: A Proposed Solution*, 27 MICH. L. REV. 751, 761 (1929).

³⁵² *Schooner*, 11 U.S. at 137.

authority and of the international comity” sovereign states should decline to exercise territorial jurisdiction.³⁵³ When a similar case reached the House of Lords in 1938, Lord Maugham relied on *The Exchange* to state that “the word ‘comity,’ whatever may be its defects in regard to other rules of private international law, has a very powerful significance” in the context of immunity.³⁵⁴

Again, the relation between comity and immunity was addressed in *Guaranty Trust Co. v. United States*, where it was held “that upon the principle of comity foreign sovereigns and their public property are held not to be amenable to suit in our courts without their consent.”³⁵⁵ The relation was further addressed in Justice Reed’s dissent in *City Bank of New York v. Republic of China*,³⁵⁶ where he asserted that any consent to a relaxation of jurisdictional rules could, “in some instances, be tested by common usage, and by common opinion, growing out of that usage.”³⁵⁷ To Reed, the word “comity” represented as good a shorthand as any to describe the whole phenomenon; further, it called to mind the idea that the sovereign could easily revoke any such consent.³⁵⁸ Indeed, there exist examples in which the severance of diplomatic relations was considered sufficient grounds to deny the invocation of immunity

353 The Parlement Belge (1880) 5 P.D. 197.

354 *Compania Naviera Vascongado v. Steamship “Cristina”* [1938] A.C. 485, 519. For a more detailed discussion on the case see SHAW, *supra* note 347, at 705.

355 *Guaranty Trust Co. v. United States*, 304 U.S. 126, 134–35 (1938).

356 *Nat'l City Bank v. Republic of China*, 348 U.S. 356, 366 (1955).

357 *Id.* at 367 (“citing *The Exchange v McFadden*”).

358 *Id.* at 376–68. Similar conclusions on the meaning of *The Exchange* had been reached in *La Santissima Trinidad*, where it was held that “[exemption] was not founded upon any notion that a foreign sovereign had an absolute right, in virtue of his sovereignty, to an exemption of his property from the local jurisdiction of another sovereign, when it came within his territory; for that would be to give him sovereign power beyond the limits of his own empire. But it stands upon principles of public comity and convenience, and arises from the presumed consent or license of nations, that foreign public ships coming into their ports, and demeaning themselves according to law, and in a friendly manner, shall be exempt from the local jurisdiction. But as such consent and license is implied only from the general usage of nations, it may be withdrawn upon notice at any time, without just offence, and if afterwards such public ships come into our ports, they are amenable to our laws in the same manner as other vessels.” *The Santissima Trinidad* 20 U.S. 283, 352–53 (1822) (finding that upon the principles of comity foreign public ships were exempt from local jurisdiction; however, the consent and license could be withdrawn at any time, without just offence).

by a foreign government.³⁵⁹ However, it was *Verlinden B.V. v. Central Bank of Nigeria*³⁶⁰ that included a definitive interpretive statement clarifying the basis of immunity: “As *The Schooner Exchange* made clear . . . foreign sovereign immunity is a matter of grace and comity on the part of the United States, and not a restriction imposed by the Constitution.”³⁶¹ For the Court, this circumstance justified the consistent deference to the political branches in determining whether it was appropriate to exercise jurisdiction against foreign sovereigns and their instrumentalities.³⁶² The point was further expanded upon in *Dole Food Co. v. Patrickson*, where the court held that the purpose of foreign sovereign immunity was to give “states and their instrumentalities some protection from the inconvenience of suit as a gesture of comity between the United States and other sovereigns.”³⁶³ Interestingly, comity was invoked in *First National City Bank v. Banco Para el Comercio Exterior de Cuba* to rule that duly established government instrumentalities were presumed to have independent status.³⁶⁴ Citing *Hilton v. Guyot*, the decision affirmed as much based on “[d]ue respect for the actions taken by foreign sovereigns and for principles of comity between nations.”³⁶⁵

More recent cases seem to rely on comity to a lesser extent, but its importance still seems to be significant. In *Republic of Austria v. Altmann*,³⁶⁶ a case dealing with an action under the FSIA, the court referenced comity and discussed its role in shaping the doctrine of immunity at length, but the concept does not appear to have played a major role in the solution of the case.³⁶⁷ *Republic of*

359 See, e.g., *The Gul Djemal*, 296 F. 563 (S.D.N.Y. 1921), aff’d on other grounds, 264 U.S. 90 (1924), cited in *Sabbatino* 376 U.S. 398.

360 *Verlinden B.V. v. Cent. Bank of Nig.*, 461 U.S. 480 (1983).

361 *Id.* at 486. See also *Samantar v. Yousuf*, 560 U.S. 305, 311 (2010) (seemingly suggesting that *The Exchange* had generally been interpreted too broadly); *Republic of Argentina v. NML Capital, Ltd.*, 134 S. Ct. 2250, 2255 (2014) (“Foreign sovereign immunity is, and always has been, ‘a matter of grace and comity on the part of the United States, and not a restriction imposed by the Constitution.’”).

362 *Verlinden B.V.*, 461 U.S. at 486.

363 *Dole Food Co. v. Patrickson*, 538 U.S. 468, 479 (2003) (citing *Verlinden B.V.*, 461 U.S. at 486).

364 *First Nat'l City Bank v. Banco Para El Comercio Exterior De Cuba*, 462 U.S. 611, 626–627 (1983).

365 *Id.*

366 *Republic of Aus. v. Altmann*, 541 U.S. 677 (2004).

367 *Id.* At 694–95.

Philippines v. Pimentel,³⁶⁸ though, paints a different picture: The case concerned a litigation in which the Republic of the Philippines and the Philippine Presidential Commission on Good Governance had invoked immunity, but the lower courts had allowed the litigation to proceed with regards to the other parties. In delivering the opinion of the Court, Justice Kennedy criticized the approach and held that “the District Court and the Court of Appeals failed to give full effect to sovereign immunity when they held the action could proceed without the Republic and the Commission. Giving full effect to sovereign immunity promotes the comity interests that have contributed to the development of the immunity doctrine.”³⁶⁹

The conclusion is particularly interesting, because it resulted in a different interpretation of Rule 19 of the Federal Rules of Civil Procedure. This interpretation was in turn mandated by comity, the operation of which required a broader conception of immunity.³⁷⁰

3. *The Privileges of Suit*

Finally, comity has traditionally served as the traditional justification for the privilege granted to foreign governments to bring suit in United States Courts. As early as 1870, Justice Bradley affirmed as much in delivering the opinion for a unanimous Court in *The Sapphire*³⁷¹:

The first question raised is as to the right of the French Emperor to sue in our courts. On this point not the slightest difficulty exists. A foreign sovereign, as well as any other foreign person, who has a demand of a civil nature against any person here, may prosecute it in our courts. To deny him this privilege would manifest a want of comity and friendly feeling. . . . The Constitution expressly extends the judicial power to controversies between a State, or citizens thereof, and *foreign States*, citizens, or subjects, without reference to the subject-matter of the controversy. Our own government has largely availed itself of the like privilege to bring suits in the English courts in

368 *Philippines v. Pimentel*, 553 U.S. 851 (2008).

369 *Id.* at 865.

370 See *Pimentel*, 553 U.S. at 868–69.

371 78 U.S. (11 Wall.) 164 (1870).

cases growing out of our late civil war.³⁷²

The decision relied heavily on comity, more so than the precedent that the Court cited, where the basis for allowing a sovereign's suit was grounded solely in the Constitution.³⁷³ One consequence of this approach to granting of the privilege, of course, was the ability to preserve "the discretion of the United States to deny it, at least to foreign states that are at war with the United States or not recognized by it."³⁷⁴ While this conclusion is not clear from the wording of the decision, this is certainly the conclusion that was reached in later cases.³⁷⁵ But the privilege of suit, in this regard, differs fundamentally from immunity. As the Court put it in a later case, "[b]y voluntarily appearing in the role of suitor it abandons its immunity from suit and subjects itself to the procedure and rules of decision governing the forum which it has sought."³⁷⁶ Interestingly, the issue of privilege of suit was later raised in a case which dealt with a question of statutory interpretation. In *Pzifer*, the Court clarified whether a sovereign state damaged by anticompetitive conduct could sue for treble damages in United States district courts.³⁷⁷ The question ultimately came down to whether sovereign states qualified as persons under the Clayton Act,³⁷⁸ which the Court answered in the affirmative, responding to a vigorous dissent of the Chief Justice by invoking the hoary rule of the privilege of suit and its basis in comity.³⁷⁹

It has been argued that comity, the roots of the privilege

372 *Id.* at 167–68.

373 *Id.* at 167. The decision refers to an earlier decision by the Pennsylvania Circuit Court, where it was affirmed that "the constitution (*sic*) of the United States gives jurisdiction to the courts of the United States, in cases where foreign states are parties; and the judicial act gives jurisdiction to the circuit courts, in all cases between aliens and citizens." *King of Spain v. Oliver*, 14 F. Cas. 577, 579 (C.C.D. Pa. 1810) (No. 7814).

374 *Dodge*, *supra* note 2, at 2091 (citing the authority of *Pfizer, Inc. v. Gov't of India*, 434 U.S. 308, 319–20 (1978)).

375 "Although comity is often associated with the existence of friendly relations between states... prior to some recent lower court cases which have questioned the right of instrumentalities of the Cuban Government to sue in our courts, the privilege of suit has been denied only to governments at war with the United States[.]" *Banco Nacional de Cuba v. Sabbatino*, 376 U.S. 398, 409 (1964).

376 *Guaranty Trust Co. v. United States*, 304 U.S. 126, 134–35 (1938).

377 434 U.S. 308, 311–12, 318.

378 15 U.S.C. § 12 (2012).

379 *Pfizer*, 434 U.S. at 318–19.

notwithstanding, has now a much smaller role to play, and that questions concerning suits brought by foreign governments can be resolved by clear-cut—if judge-made—rules, as opposed to standards.³⁸⁰ But as a less dated Second Circuit decision puts it, reliance on comity highlights “that foreign nations are external to the constitutional compact, and it preserves the flexibility and discretion of the political branches in conducting this country’s relations with other nations.”³⁸¹ The Act of State doctrine, sovereign immunity, and the privilege of foreign governments to bring suit in United States courts all have their basis in comity. But is the role of comity still significant? While other elements may play a role, what lies beneath the surface reveals a more complex reality.

First of all, it must be observed that the comity basis of these doctrines also provides a convincing explanation to the problem of deference to the political branches of the government: Indeed, such deference is intended to increase the odds of successful conduct of foreign relations, especially where sensible matters were implicated, and thus mutual convenience. It is worth recalling that comity—much like smooth talk—was never there for its own sake. Whether the doctrines “reflect not a concern about entanglement” or “a rough consequentialist judgment on the part of the federal courts”³⁸² really does not make that much of a difference. Be that as it may, the idea of deference to the Executive and the separation of powers rationale always reflected,³⁸³ and still reflects, serious policy concerns, also echoed in how the doctrine developed off of American soil.³⁸⁴

Be that as it may, it appears that American courts have less discretion at their disposal. *Kirkpatrick* has significantly restricted the operation of the Act of State doctrine; a comprehensive set of rules governs immunities after “Congress abated the bedlam in 1976, replacing the old executive-driven, factor-intensive, loosely common-

380 See Dodge, *supra* note 2, at 2126 (asserting that comity doctrines are more akin to rules than standards).

381 Price v. Socialist People's Libyan Arab Jamahiriya, 294 F.3d 82, 97 (D.C. Cir. 2002).

382 Posner & Sunstein, *supra* note 184, at 1227.

383 See the Executive’s intervention in *Schooner Exchange v. McFaddon*, 11 U.S. (7 Cranch) 116, 117–18 (1812).

384 For a leading case signaling the development of the doctrine elsewhere, see *Buttes Gas & Oil Co. v. Hammer* [1982] AC 888, 6–7 (UK). See generally Michael Singer, *The Act of State Doctrine of the United Kingdom: An Analysis, with Comparisons to United States Practice*, 75 AM. J. INT’L L. 283–323 (1981); CRAWFORD, *supra* note 29, at 85.

law-based immunity regime.”³⁸⁵ It is precisely with this in mind that Harold Koh writes about “*a general decline in the deference* shown by U.S. courts to foreign sovereignty.”³⁸⁶ Let us consider again the examples made above: it is clear that Kirkpatrick limits the operation of the Act of State doctrine, but it is not quite evident from the decision that, necessity of a tangible act of state aside, no instance of foreign embarrassment should be covered by the doctrine.³⁸⁷ Much in the same way, the FSIA has been found not to be as all-encompassing as some may have thought, thus leaving open the possibility of common law foreign sovereign immunity effectively barring a suit.³⁸⁸ To some extent, this description matches Dodge’s suggestive picture of “an international law ‘core’ and a comity ‘penumbra’,” though it is certainly true—as Dodge concedes—that many phenomena are regulated by comity alone.³⁸⁹

The better view appears to be that comity, though relegated to an ancillary role, remains a widely invoked and discussed concept in the case law of American courts dealing with the three doctrines discussed above and maintains its value as an interpretive principle.

V. United States Experience and the Global Dimension of Comity

In the foregoing sections we have examined the approach American courts and scholars have adopted when dealing with the concept of comity. It is now time to understand its peculiarities and overall fortune in the global discourse on the notion.

A. Comity Doctrines and Comity Reasoning

An analysis of the notion of comity in American law does

385 Republic of Argentina v. NML Capital, Ltd., 134 S. Ct. 2250 (2014).

386 KOH, *supra* note 27, at 88.

387 KOH, *supra* note 27, at 101; Kirkpatrick & Co. v. Envtl. Tectonics Corp., 493 U.S. 400, 409 (observing that, in any case, “the policies underlying the act of state doctrine should be considered in deciding whether, despite the doctrine’s technical availability, it should nonetheless not be invoked”). On the reduced relevance of embarrassment and political implications see also Hourani v. Mirtchev, 796 F.3d 1, 14–15 (D.C. Cir. 2015) (stating that “The function of the Act of State doctrine is to promote “international comity, respect for the sovereignty of foreign nations on their own territory, and the avoidance of embarrassment to the Executive Branch in its conduct of foreign relations.”).

388 Samantar v. Yousuf, 560 U.S. 305, 325–26 (2010) (noting that the Court was careful in “emphasiz[ing] … the narrowness of [its] holding” entrusting to the District Court the question of “[w]hether [an individual foreign official] may be entitled to immunity under the common law”).

389 Dodge, *supra* note 2, at 2084.

not simply highlight the variety of meanings that are commonly associated with the expression; rather, the most striking finding is perhaps the sheer number of doctrines and rules which in comity find their rationale. Through an examination of references to comity, we have found that the notion is taken into account in a variety of settings, such as the limitation of the reach of foreign law, the management of competing proceedings, the grant of foreign immunity or of the privilege to sue to foreign states or persons acting in official capacity. It is difficult to drive out the impression that, today as in Huber's time, comity has been to judges "a springboard from which they proceeded to develop a highly organized and sophisticated set of choice of law rules."³⁹⁰

True, as Donald Childress has convincingly observed, American courts have largely lost touch with the conflict of laws roots (and rationale) of the comity doctrine, which would provide "a more principled basis for applying the doctrine in transnational cases by bringing sovereign interests to light" and allowing for a more reasoned mediation between them.³⁹¹ And yet, other factors must be taken into account. First, reliance on a "conflict of laws" rationale might be useful and appropriate for the resolution of certain types of clashes between sovereign interests, but one wonders if the development of international law might have left any space for this approach in questions such as, for example, immunity. Secondly, the possibility should be entertained that courts have quite different things in mind when they think about such clashes, and possibly not the greatest clarity on the matter—the antitrust cases discussed above serve as a striking example. But what is, then, the significance of comity in the resolution of modern cross-border disputes? How is it possible to reconcile its nature and origins with modern decision-making?

While fundamentally accurate and evocative of the original roots of the doctrine, the metaphor of "comity as a springboard" does not fully describe its role in American law. We need not enter discussions on whether its demands are met by the application of rules or standards—indeed, while the distinction is undoubtedly relevant, it is of no consequence for our purposes.³⁹² What we submit is that comity often enters the picture in a more oblique manner than it is generally suggested, and that what we may label "comity

390 See Schultz & Holloway, *supra* note 1, and accompanying text.

391 Childress, *supra* note 2, at 63.

392 Dodge, *supra* note 2, at 2124.

reasoning” is as important as the reliance on the “principle” or “the doctrine” of comity.

One illustration is offered by the numerous references to comity in the context of decisions involving international law. To a number of commentators, American courts have consistently struggled to distinguish one from the other. There is certainly truth to this claim, but the American judiciary might have received disproportionately bad press. In fact, in many cases, it appears that courts have simply used the concept of comity as a lens through which they were required to ascertain what exactly international law permitted or required.³⁹³ As Jörn Kämmerer observed, the idea that the interpretation of norms of international law may draw inspiration from comity is not theoretically illogical, especially if comity’s theoretical vicinity with the principle of good faith is taken into account.³⁹⁴ When these aspects are considered, the cogency of comity appears in another light: statements to the effect that “[t]he king is wise and good”³⁹⁵ clearly work much better in the context of domestic statutory interpretation than they do with regards to international sources. In other words, there might be method in the American judiciary’s apparent confusion of comity and public international law, in that the former allows American courts to give proper effect to the latter.

B. The Problem of Deference to the Executive

The debate on comity in the United States has often turned into a discussion on separation of powers. It should be clear by now that comity—as a doctrine, as a theoretical construct—is largely agent-agnostic in terms of which among the powers exercises it. Not only does this observation find solid grounding in Huber’s own statement of the doctrine, where it was simply stated that the *rectores imperorum*—the sovereigns—were to act *comiter*, according the “rights acquired within the limits of a government retain their force everywhere . . .”³⁹⁶ Indeed, it also accords with the rationale

393 See Meyer, *supra* note 6, and accompanying text.

394 JÖRN KÄMMERER, *Comity*, 4–5 (2006), <http://opil.ouplaw.com/view/10.1093/law:epil/9780199231690/law-9780199231690-e691?prd=EPIL>.

395 EDMUND S. MORGAN, *INVENTING THE PEOPLE: THE RISE OF POPULAR SOVEREIGNTY IN ENGLAND AND AMERICA* 29 (1989).

396 Lorenzen, *Huber’s De Conflictu Legum*, *supra* note 55, at 376; Hartford Fire Ins. Co. v. California, 509 U.S. 764, 813 (1993) (Scalia, J.) (citing *RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW OF THE UNITED STATES* 235, 231 (AM. LAW INST. 1987)) (discussing Congress’s power to “legislative

of many canons of statutory interpretation, first and foremost the presumption against extraterritoriality, which assumes comity to have been exercised by the legislative branch, the role of the judiciary being limited to an *ex post* recognition of such exercise. Yet, by looking at the instances in which comity has been invoked and the issue of deference to the Executive raised, one would not conclude that the problem of separation of powers fits so neatly in the life of the doctrine. In fact, it is arguable that the most troubling links between comity and deference to the executive stem from sensitive matters arising from cross-border or international disputes.

This conclusion seems obvious, but it need not be: indeed, comparable occurrences before courts of other nations have not prompted courts to grant the political branches such a substantial degree of deference.³⁹⁷ More precisely, a more limited number of hypotheses call for this type of deference, which is, conversely, deeply rooted in American legal thinking and continues to encourage one of the most unbending “myths” of international comity.³⁹⁸ Indeed, this difficulty has less to do with questions relating to the United States’ form of government than with certain attitudes of the American judiciary, divided—when dealing with issues of foreign affairs—between the opposed approaches of a “customary” practice of judicial abdication³⁹⁹ and their constitutional “province and duty . . . to say what the law is.”⁴⁰⁰ Further, contrary to obsequiousness to the will of Congress, deference to the Executive could also be seen as undermining the principle of democratic accountability, and the advocacy of Executive power in this area is linked to the growing

jurisdiction” / “jurisdiction to prescribe” as “the authority of a state to make its law applicable to persons or activities” and wholly separate from the court’s “jurisdiction to adjudicate”).

397 For example, in the United Kingdom the question of deference to executive power normally arises in the context of political questions and issues of justiciability. For an example concerning the grant of diplomatic protection and the position of the British Court of Appeal that an unreasonable refusal would have been considered justiciable, see *R. (Abbas) v. Sec'y of State for Foreign and Commonwealth Affairs* [2002] EWCA (Civ.) 1598, [2003] UKHRR 76, [51], [85], [106] (C.A.). *See also Canada v. Khadr*, [2010] 1 S.C.R. 44 (Can.) (holding that a foreign party’s rights were, but the court left it up to the Crown to decide how to best respond given “its responsibilities to foreign affairs, and the Charter” in question.)

398 *Dodge*, *supra* note 2, at 2132.

399 *Jama v. Immigration & Customs Enf't*, 543 U.S. 335, 338 (2005).

400 *Marbury v. Madison*, 5 U.S. (1 Cranch) 137, 177 (1803). *See Curtis A. Bradley, Chevron deference and foreign affairs*, 86 VA. LAW REV. 649, 650 (2000).

fear of the “Soviet threat.”⁴⁰¹ As Curtis Bradley has persuasively observed, the main problem with the hoary problem of deference is that it has been traditionally perceived as a unitary phenomenon, when in fact it cannot be labelled as such.⁴⁰²

While there appears to be no reason why a court cannot be invested of an issue having cross-border significance and conduct its comity analysis, where needed, alone, it goes without saying that common sense and institutional courtesy limit this kind of behavior, and the argument that at least some deference should be given to the Executive appears quite compelling: But how to explain and conceptualize a framework for deference?

It has been suggested that borrowing the doctrine of *Chevron*⁴⁰³ deference from administrative law could be particularly fruitful.⁴⁰⁴ Broadly speaking, the *Chevron* doctrine requires courts to engage in a two-step analysis in the interpretation of statutes, first determining whether Congress has spoken clearly on the question at issue, and then giving deference to the reading put forward by the governmental agency tasked with administration of the statute itself, insofar as it is permissible.⁴⁰⁵ The rationale for this deference is that “[j]udges are not experts in the field, and are not part of either political branch of the Government,” while agencies are both better placed and more politically accountable—if only through the Chief Executive.⁴⁰⁶

With comity in mind, reliance on the *Chevron* rationale can be either hailed as a helpful contribution or criticized as an unnecessary and problematic complication: the truth probably stands somewhere inbetween and, most importantly, is highly context-dependent. Indeed, the Executive’s interpretation of a treaty is not the same thing as relying on its determination as to whether the Act of State doctrine should not be applied, or a foreign head of state granted immunity from jurisdiction.⁴⁰⁷ The concept of deference is clearly not unitary.⁴⁰⁸ And as far as comity is concerned, William Dodge has persuasively argued that deference to the executive is highly

401 Paul, *The Transformation of International Comity*, *supra* note 9, at 33.

402 Bradley, *supra* note 402, at 649.

403 *Chevron U.S.A., Inc. v. Natural Res. Def. Council, Inc.*, 467 U.S. 837 (1984).

404 Bradley, *supra* note 402, at 651; Posner & Sunstein, *supra* note 184, 1204–07, 1227–28.

405 *Chevron*, 467 U.S. at 842–43, 865.

406 *Id.* at 865–66.

407 See generally, *Samantar v. Yousuf*, 560 U.S. 305, 322 (2010).

408 Bradley, *supra* note 402, at 665–66, 659.

inappropriate wherever the doctrine comes down to “the core responsibility of the courts to manage their dockets and decide cases.”⁴⁰⁹ We understand Dodge’s argument as mostly based on the *domestic* dimension of separation of powers; we restrict ourselves to appending a minor observation on the *international* dimension of the phenomenon, noting that deference to the executive in circumstances such as, for example, the recognition of foreign judgments and acts, the whole area of adjudicatory comity, could—in itself—be perceived by foreign sovereigns as revealing of a want of comity.

Beyond this aspect, the involvement of the executive need not, in and by itself, be considered incompatible with the doctrine of comity on the international plane, though the extent to which it is defensible is a matter of degrees and balance. To bring this further, we may consider the role of deference within the broader framework of “foreign relations” law, a legal category that sounds rather unfamiliar to those writing from other shores of the Atlantic, but a well-established one in the United States.⁴¹⁰ As McLachlan, of one of the few non-American lawyers who has devoted considerable attention to the field, put it, foreign relations law performs an “allocative function.”⁴¹¹ As the author contends, this function is an aggregate of two different ones, in that it controls the jurisdiction and applicable law “in the external exercise and control of the public power of states” and contributes to the ordering of “the allocation of foreign affairs competence within the municipal constitution.”⁴¹² This allocative function obeys a conflict of laws logic: while conflict involves determinations on jurisdiction and applicable law that follow a “two-dimensional” approach considering two systems of municipal law, foreign relations law implicates a “three-dimensional” judgment on the allocation of institutional competence.⁴¹³ In the

409 Dodge, *supra* note 2, at 2132.

410 See, e.g., LOUIS HENKIN, FOREIGN AFFAIRS AND THE UNITED STATES CONSTITUTION 88 (2d ed. 1997).

411 Campbell McLachlan, *The Allocative Function of Foreign Relations Law*, 82 BRIT. Y.B. INT’L L. 349, 377 (2012).

412 *Id.*

413 *Id.* at 380. Quite remarkably, McLachlan refers to Lauterpacht’s own treatment of issues of foreign relations law in HERSCH LAUTERPACHT, THE FUNCTION OF LAW IN THE INTERNATIONAL COMMUNITY 397 (2011). As Martti Koskenniemi discusses in the foreword to the revised edition, Lauterpacht was wary of areas of non-justiciability, more so on the international plane. However, to the question of whether the “limitation of the competence of the courts in these matters [could] be construed as a limitation of the rule of law within the State,” as he continued, “care must be taken not to confuse the

United States, foreign relations law is a fundamentally internal matter and reflects the American Constitution's structural aspects. It follows that constitutional prerogatives of the Executives make it so that deference is, if not always necessary, justifiable. To this extent, whether this is compatible with the Courts' "judicial duty to know and to declare" the "comity of our own country" is probably a question.⁴¹⁴

On the one hand, the Executive may appeal because of its expertise and accountability;⁴¹⁵ on the other hand, there may be questions of legitimacy and fear of being led onto a short-term focused agenda with potentially harmful long-term consequences.⁴¹⁶ But above all, Dodge rightly observes that "[e]ach opportunity for deference invites pressure from foreign governments and creates the possibility of diplomatic backlash if the Executive decides not to support their positions."⁴¹⁷ In this regard, the pattern established in limiting the reliance on executive determinations in the areas of foreign sovereign immunity supports the proposition that the "government need not, and should not, speak in every case."⁴¹⁸ As then Legal Adviser to the Department of State Harold Koh put it, "[i]n domestic litigation, [the Department's] ultimate goal is, in fact, not more verbiage, but more silence."⁴¹⁹

C. Beyond the Domestic: the "Comity of Courts" as a Global Ordering Principle

In 1998, in the wake of *Breard v. Greene*⁴²⁰ and the provisional measures issued by the International Court of Justice (ICJ) in the *Case concerning the Vienna Convention on Consular Relations*,⁴²¹ Anne-Marie Slaughter penned a seminal essay on the *American Journal of International Law* in which she discussed the implications of the

limitation upon the unrestricted freedom of judicial decision with a limitation of the rule of law." *Id.*

414 *Hilton v. Guyot*, 159 U.S. 113, 228 (1895).

415 KOH, *supra* note 27, at 150.

416 Derek Jinks & Neal Kumar Katyal, *Disregarding Foreign Relations Law*, 116 YALE L.J. 1230, 1262 (2007).

417 Dodge, *supra* note 2, at 2140.

418 Harold Hongju Koh, *Foreign Official Immunity After Samantar: A United States Government Perspective*, 44 VAND. J. TRANSNAT'L L. 1141, 1161 (2011).

419 *Id.*

420 *Breard v. Greene*, 523 U.S. 371 (1998).

421 Vienna Convention on Consular Relations (Para. v. U.S.), Order, 1998 I.C.J. Rep. 248 (Apr. 1998).

Supreme Court's decision not to accord a stay of execution in compliance with the World Court's order.⁴²² Slaughter observed that, irrespective of whether the measures issued by the ICJ were binding,⁴²³ the Supreme Court should have nonetheless honored the request "as a matter of judicial comity."⁴²⁴ In this essay and her subsequent work, Slaughter observed that the United States judiciary was re-discovering the concept of "judicial comity," building the case on the basis of Scalia's dissent in *Hartford Fire*. It has been long argued that "the comity of courts" terminology left much to be desired, but it most certainly provide an opportunity for Slaughter to describe a number of approaches and attitudes as driven by comity, and comity as "the lubricant of transjudicial relations."⁴²⁵ Slaughter's understanding of "judicial comity" later evolved as one of the building blocks of the theoretical model she developed in later writings for the construction of a global legal system through the concerted work of domestic courts. In *A New World Order* she described it as providing:

[T]he framework and the ground rules for a global dialogue among judges in the context of specific cases. It has four distinct strands. First is a respect for foreign courts qua courts, rather than simply as the face of a foreign government, and hence for their ability to resolve disputes and interpret and apply the law honestly and competently. Second is the related recognition that courts in different nations are entitled to their fair share of disputes—both as co-equals in

422 Anne-Marie Slaughter, *Court to Court*, *supra* note 8.(stating, with regards to Breard, that "The Supreme Court should have honored this request as a matter of judicial comity, offering the IJC the same respect that U.S. courts are increasingly according their counterparts around the world.").

423 There can now be very little doubt as to the binding nature of the provisional measures issued by international courts and tribunals. There exists a large amount of literature on the topic. See, e.g., CHESTER BROWN, A COMMON LAW OF INTERNATIONAL ADJUDICATION 146–49 (2007) (finding commonalities in the procedural approaches of international courts and tribunals); SHABTAI ROSENNE, PROVISIONAL MEASURES IN INTERNATIONAL LAW: THE INTERNATIONAL COURT OF JUSTICE AND THE INTERNATIONAL TRIBUNAL FOR THE LAW OF THE SEA (2005); see also Hugh Thirlway, *The Law and Procedure of the International Court of Justice 1960–1989: Part Eleven*, 71 BRIT. Y.B. INT'L. L. 71 (2001).

424 Slaughter, *Judicial Globalization*, *supra* note 428, at 708.

425 *Id.*

the global task of judging and as the instruments of a strong “local interest in having localized controversies decided at home.” Third is a distinctive emphasis on individual rights and the judicial role in protecting them. Fourth, although seemingly paradoxically, is a greater willingness to clash with other courts when necessary, as an inherent part of engaging as equals in a common enterprise.⁴²⁶

To be sure, Slaughter’s claims have sometimes been portrayed as vaguely starry-eyed: most notably, while acknowledging that conflict between courts is inevitable, she argues that it is conducive to greater dialogue, and thus comity.⁴²⁷ As Alex Mills and Tim Stephens have observed, such a claim relies on the notion that “the ‘special’ character of courts” and “a capacity of a free market of legal ideas to avoid distortions caused by inequalities of power” will allow substantive conflict avoidance through agreed procedure.⁴²⁸ Yet, Slaughter’s theory has proved fascinating to many, especially to those scholars and practitioners who dealt with the most informal of all legal regimes, that of public international law. In the context of fragmentation of international law and its “institutional aspect,”⁴²⁹ the proliferation of international courts and tribunals, it has been often argued that comity might have the potential of mitigating the resulting problems.⁴³⁰ The practice of international adjudication and arbitration too seems to provide a number of indications that international courts and tribunals are conscious of the value of comity, and respect each other’s competence and decision-making capacity, displaying an awareness of their status as “co-equals in the global task of judging” and recognizing different specializations.⁴³¹

⁴²⁶ SLAUGHTER, *supra* note 15, at 87; Many of these aspects had been dealt with in Slaughter, *Judicial Globalization*, *supra* note 248; Anne-Marie Slaughter, *A Global Community of Courts*, 44 HARV. INT’L L.J. 191 (2003).

⁴²⁷ SLAUGHTER, *supra* note 15, at 89–90.

⁴²⁸ Alex Mills & Tim Stephens, *Challenging the Role of Judges in Slaughter’s Liberal Theory of International Law*, 18 LEIDEN J. INT’L L. 1, 30 (2005).

⁴²⁹ CRAWFORD, *supra* note 170, at 298, 302–303, 355.

⁴³⁰ See, e.g., *id.* (defining proliferation as institutional fragmentation). See generally Schultz & Ridi, *supra* note 4. See also SHANY, *supra* note 16, at 260; YUVAL SHANY, *REGULATING JURISDICTIONAL RELATIONS BETWEEN NATIONAL AND INTERNATIONAL COURTS* 166 *passim* (2007).

⁴³¹ Slaughter, *Judicial Globalization*, *supra* note 248, at 1113. See, e.g., MOX Plant Case (Ir. v. U.K.), Case No. 10, Order of Dec. 3, 2001, <https://www.itlos.org/en/cases/list-of-cases/case-no-10/#c667>, where the tribunal relied on

This attitude seems to transcend simple institutional dialogue,⁴³² rather amounting to “an emerging general principle of international procedural law.”⁴³³ Indeed, there are indications that comity has served as a valuable tool even when dealing with competing proceedings before courts of different orders.⁴³⁴

Reasons of space and context prevent us from examining these aspects in further detail. We restrict ourselves to observing that the roots of the global comity discourse are, in both its real-world and theoretical dimensions, unmistakably American.⁴³⁵

VI. Conclusion

Centuries after Westphalia, sovereignty and territoriality still hold their significance as fundamental notions in the global ordering of regulatory authority. Despite proclamations of their absolute value, they remain—as they have always been—relative notions. Their evolution, informed by change and tensions, has, time after time, resulted in troubled waters where the safest course has often been inspired by the comity doctrine. Indeed, the development of the doctrine in the United States has shadowed these changes, rendering comity a lens through which a world of competing sovereignties

“considerations of mutual respect and comity which should prevail between judicial institutions” to justify the suspension of its proceedings in the face of an almost certain involvement of Court of Justice of the European Union. For a discussion of the meaning and significance of comity in a variety of different international fora, see generally Schultz & Ridi, *supra* note 4.

432 See, e.g., Rosalyn Higgins, *A Babel of Judicial Voices? Ruminations from the Bench*, 55 INT'L COMP. L. Q. 79 (2006) (discussing international judges of courts and tribunals similarly appearing to be mindful of the need to engage in institutional dialogue and pay appropriate deference to each other’s judgments, thus appearing to embrace a pluralist structure of the international legal order); Cesare P.R. Romano, *Deciphering the Grammar of the International Jurisprudential Dialogue*, 41 N.Y.U. J. INT’L L. & POL. 755 (2008); Relying almost exclusively on the concept of comity, see generally, Elisa D’Alterio, *From Judicial Comity to Legal Comity: A Judicial Solution to Global Disorder?*, 9 INT’L J. CONST. L. 394 (2011) (discussing international court’s deference to each other’s respective interpretations of their own local law).

433 Bruno Simma, *Universality of International Law from the Perspective of a Practitioner*, 20 EUR. J. INT’L. L. 265, 287 (2009).

434 See, e.g., Société Générale de Surveillance S.A. v. Republic of the Phil., ICSID Case No. ARB/02/6, Objections to Jurisdiction, 34, (Jan. 29, 2004), 8 ICSID Rep. 518 (2005). See also, SHANY, *supra* note 432; Schultz & Ridi, *supra* note 4.

435 See, e.g., *Bank of Augusta v. Earle*, where Justice Taney saw comity as promoting “a friendly intercourse between the sovereignties.” 38 U.S. (13 Pet.) 519 (1839).

could be better understood, and a source of techniques to be employed with the results of such a theoretical analysis. Comity has gone on to become deeply embedded in the American legal mindset and vocabulary, but it has also transformed, adapting to the needs of a different type of sovereignty—that of a super power, unafraid of most things, never mind an extraterritorial assertion of jurisdiction. But if the force of comity has faded as a jurisdictional constraint, that is, on the international plane, its significance at the domestic level has remained substantial. In that respect, comity has gone beyond its nature of conflict of laws principle and has become a fully-fledged tenet of foreign relations law, conveying constitutional concerns of separation of powers and contributing, in addition to its external dimension, to the domestic element of the “allocative function” of this area of the law.

To be sure, the concept remains elusive and difficult to define unambiguously. But, if “the definition of comity may be tenebrous, its importance could not be more clear.”⁴³⁶ This article has sought to show how American courts have invoked the concept for a variety of purposes, some fitting the original theoretical foundations of the doctrine, and some others going beyond; we have tried to highlight the most representative examples of comity-based reasoning to illustrate how entrenched the doctrine is in the way American judges and scholars think; we have elected to partially distance ourselves from the perspective of private international law to frame the discourse on comity in the broader framework of foreign relations law, endeavoring—to our mind, successfully—to make sense of the peculiarity of the American experience; finally, we have discussed how the American experience on comity has proved influential and helped construe the doctrine as a theoretical construct capable of mitigating conflict between regimes fundamentally different from interstate ones, such as the ones between the jurisdictions of international courts and tribunals.

When all is said and done, it is perhaps true that comity plays a smaller role today than it did two hundred years back. Critics of the doctrine have always conceded that it was never really forgotten by the American legal world and in this study we acknowledge as much, but we also demonstrate how its use is as lively as ever. True, comity has undoubtedly “transformed.”⁴³⁷ This transformation, though, is

⁴³⁶ Quaak v. Klynveld Peat Marwick Goerdeler Bedrijfsrevisoren, 361 F.3d 11, 19 (1st Cir. 2004).

⁴³⁷ We borrow the expression from Paul, *The Transformation of International Comity*,

not an indication of the doctrine's demise: in fact, it confirms its relevance, in light of the incapacity of its corollaries, sovereignty and territoriality, to deal with the challenges of the 21st Century—an incapacity that affected the discourse far beyond legal scholarship.⁴³⁸ Comity, in other words, still represents an unparalleled "springboard" from which a number of inferences relating to regime conflict can be drawn.⁴³⁹

* * *

supra note 9.

438 As an example of the recent debates on the nature of sovereignty in mainstream media, see *Boris Johnson is Wrong: In the 21st Century, Sovereignty is Always Relative*, THE ECONOMIST, <http://www.economist.com/blogs/bagehot/2016/02/bojo-breaks-ranks> (last visited Mar. 20, 2016).

439 McLACHLAN, *supra* note 60, at 32.

Contract as Reliance, Promise, and Consent

*By: Caleb B. Koufman, J.D., M.S.Ed.
A Student Note**

* This essay was largely prepared through an independent study supervised by Professor Peter Enrich, Esq. and I am extremely appreciative of his help and guidance.

Table of Contents

I.	Introduction	91
A.	<i>The Teardrop Example</i>	91
B.	<i>Road Map</i>	92
II.	Connective Tissue	93
A.	<i>Background</i>	93
B.	<i>Doctrine of Consideration.....</i>	94
C.	<i>Mistake/Change in Circumstance</i>	99
III.	Theorists.....	103
A.	<i>Grant Gilmore</i>	103
1.	<i>Contract as Reliance</i>	103
2.	<i>Contract as Reliance and Lack of Consideration</i>	105
3.	<i>Contract as Reliance and Mistake/Change in Circumstance.....</i>	106
B.	<i>Charles Fried</i>	108
1.	<i>Contract as Promise.....</i>	108
2.	<i>Contract as Promise and Lack of Consideration</i>	110
3.	<i>Contract as Promise and Mistake/Change in Circumstance.....</i>	111
C.	<i>Randy Barnett</i>	112
1.	<i>Contract as Consent</i>	113
2.	<i>Contract as Consent and Lack of Consideration</i>	114
3.	<i>Contract as Consent and Mistake/Change in Circumstance.....</i>	114
IV.	Application	115
A.	<i>Example</i>	115
B.	<i>Morality in Contract as Reliance, Promise, and Consent</i>	118
C.	<i>Conclusion</i>	121

I. Introduction

A. *The Teardrop Example*

On a Sunday morning during my second year of law school, I drove to my family's home on the north shore of Boston to say goodbye to my brothers before they departed for their respective colleges. My vehicle, a 1996 periwinkle blue Saab hatch-back, nicknamed "the teardrop," was in desperate need of a cleaning because it was covered in pine tar after a trip to a friend's lake house in New Hampshire.

I envisioned that cleaning the car would take a couple hours because the pine tar was everywhere. In my experience, pine tar is difficult to remove, and it is equally as time-consuming to make a dirty car sparkle again. I figured that my youngest brother would jump at the opportunity to make some cash, so I offered Sam \$50 for the job. The amount seemed fair to me because both of us were full-time students and cash is hard to come by when all hours of the day are spent on academic pursuits. I also knew that the local car wash charged \$28 for a standard exterior wash, and I felt that the car wash machine would not be able to remove the tar.

Sam quickly accepted the agreement, and I went about my business, which entailed raiding my parents' fridge. Ten minutes later, Sam walked in and asked for his money. Incredulous, I stormed outside to find the teardrop in tip-top shape. Sam had completely removed the pine tar. Sam informed me there was a solvent in my father's garage that only takes a minute to apply, and it both removes pine tar and buffs up paint jobs with ease. Sam knew about the solvent when we made the agreement.

Despite joking with Sam about how his work was not worth the \$50, and feeling like he took advantage of my ignorance with regard to the cleaning solvent, I felt obligated to pay Sam.

This informal transaction, hereinafter referred to as "the teardrop example," led to certain ruminations about contractual liability. In particular, I wanted to know why I felt that my agreement with Sam deserved to be reinforced. I turned to Grant Gilmore, Charles Fried, and Randy Barnett for answers, because these three scholars' respective theses are preeminent in the field of contract law philosophy.

B. Road Map

Specifically, this note attempts to address the following issues:

- (1) Whether something compels individuals to fulfill promises made during informal private ordering;
- (2) Whether there exists a singular moment when contractual liability begins to exist;
- (3) Whether the doctrine of consideration plays a role in informal, privately ordered agreements;
- (4) Whether the doctrines of mistake and change in circumstance play a role in informal, privately ordered agreements;
- (5) Whether Grant Gilmore's, Charles Fried's, and Randy Barnett's theories, respectively, adequately address Issues 1-4; and,
- (6) Whether Grant Gilmore's, Charles Fried's, and Randy Barnett's theories, respectively, explain the promisor's sense of an obligation to perform in the teardrop example.

The remainder of this note is organized into parts. Part II will begin with a background into contract law and will serve as a connective tissue between contract law's foundational elements and the informal interactions analyzed. Part II(B) will discuss the doctrine of consideration. Part II(C) introduces the doctrines of mistake/change in circumstance. Each of these doctrines will prove integral to the concluding discussion given the informal nature of the example posited at the outset.

Part III delves into the three theories as explained in each author's respective works: (A) Grant Gilmore's *The Death of Contract*,¹ (B) Charles Fried's *Contract as Promise*,² and, (C) Randy Barnett's *Contract as Consent*.³ Each includes background of the given thesis in subsection (i), a discussion about consideration with respect to each thesis in subsection (ii), and analysis of how each deals with mistake and/or change in circumstance in subsection (iii).

Part IV serves as the conclusion. Part IV(A) applies each theory to the teardrop example. Part IV(B) presents brief closing remarks concerned with the intersection of morality and contractual liability. Part IV(C) concludes which scholar's theory most accurately captures the promisor's sense of obligation in the teardrop example.

1 GRANT GILMORE, *THE DEATH OF CONTRACT* (1st ed. 1974).

2 CHARLES FRIED, *CONTRACT AS PROMISE* (2d ed. 2015).

3 RANDY BARNETT, *CONTRACTS* (2010).

II. Connective Tissue

A. Background

Some background into contract law and the field's doctrinal history is warranted in this case because much of the legal theory behind the ensuing discussion draws on the manner in which contract law's treatment of informal interactions, like the teardrop example, has changed over time.

In general, contract law enables individual actors to define their relationships to one another in an effort to achieve a desirable result.⁴ Tort law uses a reasonableness standard to define expectations surrounding the conduct of interacting individuals who did not work those expectations out ahead of time. Contract law affords individuals the opportunity to define expectations before an exchange has occurred. Contractual liability enables individual parties to rely on each other and combine efforts in an attempt to manufacture a particular, desired future.

An important distinction between a contractually obligated promise to perform and a duty in tort law is that, typically, a contracting party intends to legally obligate itself to perform on a promise. Contracting parties create duties. In a sense, tort law functions retroactively, by first considering whether or not a duty existed between parties, and, if so, imposing civil liability.

An alternative view reasons that contract law is just a piece of tort law.⁵ On this reading, reliance plays an important role in contract disputes. Reliance principles can sometimes result in the enforcement of an agreement despite the fact that it may be unclear

4 See Jennifer A. Drobac & Oliver R. Goodenough, *Exposing the Myth of Consent*, 12 IND. HEALTH L. REV. 471, 476 (2015) ("At a more theoretically grounded level of justification, neoclassical economists assume that informed, autonomous actors strike welfare-maximizing bargains through free contracting. These actors exchange bargained-for consent to arrive at efficient, productive deals, often demonstrating 'Pareto efficiency.'") (internal citations omitted).

5 See, e.g., David G. Epstein et. al., *Reliance on Oral Promises: Statute of Frauds and "Promissory Estoppel"*, 42 TEX. TECH L. REV. 913, 925 (2010) ("If promissory estoppel is tort as Professor Gilmore wrote or even 'quasi-contract,' as Judge Lynn recently wrote, then it is arguable that contract law concepts do not apply to causes of action based on promissory estoppel.") (citing Eagle Metal Prods., L.L.C. v. Keymark Enters., L.L.C., 651 F. Supp. 2d 592, 592 (N.D. Tex. 2009) ("The promissory estoppel claim is of a different order from the tort claims. Promissory estoppel is a quasi-contract theory, which seeks to hold a party responsible for promises that induced justifiable reliance on another.")).

that the agreement reflects the contracting parties' respective intents. Indeed, the relationship between contracting parties and the promises that they make is important to understand both within the legal and everyday worlds in which individual parties interact.

B. Doctrine of Consideration

Contract law has evolved, however. While today contracting parties may be able to enforce an agreement scribbled hastily onto a piece of paper,⁶ this has not always been the case. The brief ensuing discussion will provide the historical backdrop needed to appreciate how contemporary contract law has evolved. In this subsection, this history is told through the lens of the doctrine of consideration. In the next subsection, the history is expounded upon through the lens of the doctrine of mistake. The development of both doctrines indicates the thought process of Anglo/American courts as the law has attempted to reflect society's changing values.

In the beginning, contract law valued formal systems. Although not all informal agreements lacked legal enforceability, the first Statute of Frauds, adopted in 1677, required "[e]nhanced formality" such that some agreements needed to satisfy formal requirements in order to benefit from legal enforcement.⁷ In order to distinguish formal promises that required enforcement from informal promises that did not benefit from legal enforcement, courts applied the "doctrine of consideration."⁸ "The consideration, or considerations, for a promise meant the factors which the promisor considered when he promised, and which moved or motivated his promising. Although not a precise equivalent, 'motive' is perhaps

6 See *Lichty v. Merzenich*, 563 P.2d 690, 691–93 (Or. 1977) ("According to Merzenich, when he asked Lichty at what cost he would complete the listed work, Lichty suggested that each of them write a figure on a piece of paper and turn the papers over simultaneously . . . Merzenich further testified that Lichty would not sign the slip because he did not have a contractor's license, but that he gave his word and they shook hands on it. Lichty's testimony was that the figures were estimates, that he made it clear he could not legally act as a contractor at a firm price, and that there was no handshake, but he did not deny the parties' use of the slips with their proposed cost figures . . . Dr. Merzenich's list of work to be completed, which in another setting might have left a contractor's commitment too undefined, here drew meaning from the parties' collaboration on the project for a full year.").

7 BARNETT, *supra* note 3, at 149.
8 *Id.*

about as near as one can get by way of synonym.”⁹

“[G]ood considerations” were seen to introduce a detriment to the promisee; the concept was called the “benefit-detriment” conception of consideration, and it “can still be glimpsed in some of the cases taught in contracts courses [today,]” like in *Blatt v. University of Southern California*.¹⁰ In *Blatt*, the plaintiff:

alleged that as a law student he had been induced, by defendants’ promises of eligibility for membership [into a national honorary legal society] if he graduated in the top 10 percent of his class, to work hard, that he had, in fact, finished in the top 10 percent, and that refusal of membership, based on his not meeting certain allegedly inapplicable law review requirements, would affect his professional and economic interests.¹¹

The *Blatt* court found that the university’s promise of eligibility did not create an enforceable agreement because “[t]here was no benefit flowing” to the university resulting from the student’s promise to maintain that grade point average.¹² Among other criteria, courts required a benefit-detriment relationship between parties involved in an agreement in order to create contractual liability.¹³

In 1880, Dean Langdell published *Summary of the Law of Contracts* as an appendix to his casebook published in that same year.¹⁴ In his summary, Langdell argues that the courts should operate as:

[D]etached umpires or referees, doing no more than to see that the rules of the game were observed and refusing to intervene affirmatively to see that justice or anything of that sort was done[;] . . . if A, without the protection of a binding contract, improvidently relies, to his detriment, on B’s promises and assurances, that may be unfortunate for A but is no fit matter for

9 A.W.B. SIMPSON, *A HISTORY OF THE COMMON LAW OF CONTRACT: THE RISE OF ACTION OF ASSUMPSIT* 321 (1975).

10 BARNETT, *supra* note 3, at 150; *see Blatt v. Univ. of S. Cal.*, 85 Cal. Rptr. 601 (Ct. App. 1970).

11 *Blatt*, 85 Cal. Rptr. 601, 601 (Ct. App. 1970).

12 *Id.* at 606.

13 *See Barnett, supra* note 3, at 150.

14 GRANT GILMORE, *THE DEATH OF CONTRACT*, at 13 (2d ed. 1995).

legal concern.¹⁵

Initially, consideration was based on “the idea that the legal effect of a promise should depend upon the factor or factors which motivated the promise... [requiring information about] why the promise was made.”¹⁶ As will become abundantly obvious, this understanding of consideration fell by the wayside.

The benefit-detriment theory of consideration failed because of reasons exemplified in *Hamer v. Sidway*, where the executor of an uncle’s estate sought to void the uncle’s promise to compensate his nephew in a certain monetary amount on the condition that the nephew refrain from drinking, smoking, and playing billiards for money, until age twenty-one.¹⁷ The *Hamer* court wrote:

It is not essential in order to make out a good consideration for a promise to show that the promisor was benefited or the promisee injured; a waiver on the part of the latter of a legal right is sufficient. . . . Courts will not ask whether the thing which forms the consideration does in fact benefit the promisee or a third party... [because i]t is enough that something is promised, done, forborne, or suffered by the party to whom the promise is made as consideration for the promise made to him.¹⁸

The *Hamer* court recognized that the benefit-detriment theory requires an inquiry into parties’ subjective understandings of value, which the court, and legal theorists, deemed irrelevant.

The first formal organization of contract law in this country is credited to Oliver Wendell Holmes in *The Common Law* (1881) and to Samuel Williston in *Williston on Contracts* (1920).¹⁹ It was at this time that theorists expressed favor for the “bargain theory” of consideration over the benefit-detriment theory. The bargain theory of consideration requires a showing that the consideration was bargained-for.²⁰ The turn of the century marked the height of classical

15 *Id.* at 16.

16 SIMPSON, *supra* note 9, at 321.

17 *Hamer v. Sidway* 124 N.Y. 538 (1891).

18 *Id.* at 538-39, 545. (internal quotations omitted).

19 GILMORE, *supra* note 1, at 6.

20 BARNETT, *supra* note 3, at 154.

legal theory, because classical theorists were conceptualizing the common law contract system, and theorists were trying to build a body of contract law that embodied a more objective stance, making the benefit-detriment theory of consideration inapposite for its emphasis on subjective inquiry. The classical approach to contract theory was “blind to details of subject matter and person.”²¹ The classical approach values formality, and requires objective indicators that courts can use to analyze questions of liability.

The bargain-theory of consideration does not require a subjective inquiry; it instead asks from an objective viewpoint whether there was an exchange or a bargain. The bargain-theory is unconcerned with what motivated parties, and instead considers whether or not a legal right was foregone, which would symbolize a detriment. A foregone legal right in exchange for a promise represents bargained-for consideration.²² Bargained-for consideration theory is the progeny of Holmes’ “reciprocal conventional inducement” theory, which found reliance on a promise “significant only if it is undertaken in exchange for the promise and the promise is given in exchange for the detriment.”²³

The bargain-theory of consideration accounts for both the *Blatt* and *Hamer* holdings. The *Blatt* court explained that, “the promisee [never] suffered actual detriment in foregoing an act, in refusing other [opportunities] or in expending definite and substantial effort or money in reliance on a promise.”²⁴ In essence, the student never forewent a legal right, so a legally enforceable agreement never existed. Similarly, the *Hamer* court found that the nephew provided consideration by foregoing his legal rights to drink, smoke, and play billiards; whether or not a foregone legal right induced a promise is trivial because, “[i]n general a waiver of any legal right at the request of another party is a sufficient consideration for a promise.”²⁵

An effective contract law theory does not require the court to attempt to understand subjective intent because the ability to read minds and assess the value of promises is irrelevant to a system of contractual liability that values objectivity. It might also be noted

21 LAWRENCE FRIEDMAN, CONTRACT LAW IN AMERICA: A SOCIAL AND ECONOMIC CASE STUDY, at 20 (Madison Univ. of Wis. Press ed., 1965).

22 See BARNETT, *supra* note 3, at 151.

23 Vincent A. Wellman, *A Common Mistake About the Common Law*, 92 MICH. B.J. 39, 40 (2013).

24 *Blatt v. Univ. of S. Cal.*, 85 Cal. Rptr. 601, 606 (1970).

25 *Hamer v. Sidway*, 124 N.Y. 538, 545 (1981).

that, as in *Hamer*, where the promisor passed away, this type of subjective inquiry may be impossible.

The drafters of the Restatements of Contract first convened in the 1930's, during the heyday of the classical approach. The Restatement (First) of Contract was drafted in 1932, and the Restatement (Second) of Contract was drafted in 1981.

The Restatement (First) of Contract section 75 explains,

(1) Consideration for a promise is

- (a) an act other than promise, or
- (b) a forbearance, or
- (c) the creation, modification or destruction of a legal relation, or
- (d) a return promise,

bargained for and given in exchange for the promise.²⁶

The Restatement (Second) of Contracts section 71 holds:

(1) To constitute consideration, a performance or a return promise must be bargained for.

(2) A performance or return promise is bargained for if it is sought by the promisor in exchange for his promise and is given by the promisee in exchange for that promise.

(3) The performance may consist of

- (a) an act other than a promise, or
- (b) a forbearance, or
- (c) the creation, modification, or destruction of a legal relation.²⁷

In analyzing whether or not a contract exists, classical theorists want courts to look for the existence of objective criteria, like consideration. Classical theorists want consideration to function as a sorting mechanism that methodically indicates which agreements are worthy of legal enforcement. Emphasis on consideration in the Restatements of Contracts keeps formalist theory alive, to this day.

Despite the nod to formalism in sections 75 and 71, both Restatements include a section 90 that, arguably, denies that

26 RESTATEMENT OF THE LAW OF CONTRACTS § 75 (AM. LAW INST. 1932).

27 RESTATEMENT (SECOND) OF CONTRACTS § 71 (AM. LAW INST. 1981).

formality. The Restatement (First) of Contracts section 90 provides, “[a] promise which the promisor should reasonably expect to induce action or forbearance of a definite and substantial character on the part of the promisee and which does induce such action or forbearance is binding if injustice can be avoided only by enforcement of the promise.”²⁸

The Restatement (Second) of Contracts section 90 explains:

A promise which the promisor should reasonably expect to induce action or forbearance on the part of the promisee or a third person and which does induce such action or forbearance is binding if injustice can be avoided only by enforcement of the promise. The remedy granted for breach may be limited as justice requires.²⁹

Suddenly, consideration and other formal systems are not the essential pieces to all enforceable contracts. Rather, courts may be persuaded by a showing that a performance was induced by a promise that the promisor should have reasonably foreseen to induce reliance. This analysis disregards formal systems, such as consideration.

The result is an undeniable tension between the sections on consideration, and the original and revised section 90, as they take opposing stances on whether or not contractual liability requires the formal systems that were to function as sorting mechanisms.

C. *Mistake/Change in Circumstance*

As formal values fell by the wayside, and section 90 recoveries became attainable, informal contracting became essential to any comprehensive discussion of contractual liability. And with informal but enforceable contracts came mistake and change in circumstance. The teardrop example demonstrates that mistake is an important piece to understanding contractual liability in informal agreements because the promisor in that example was mistaken as to the facts underlying the agreement.

Perhaps part of the reason Sam’s demand for payment was met with some amount of surprise and an initial disinclination to

28 RESTATEMENT OF THE LAW OF CONTRACTS § 90 (AM. LAW INST. 1932).

29 RESTATEMENT (SECOND) OF CONTRACTS § 90 (AM. LAW INST. 1981).

perform was because Sam was aware of circumstances to which all parties to the agreement were not privy. Namely, Sam was aware of the existence of the cleaning solvent, which was not factored into the promisor's calculations when considering the task's difficulty. It is not clear that Sam had a duty to make the promisor aware of this solvent, but it is arguable that the promisor made a mistake.

Although the teardrop example does not involve change in circumstance, the doctrines of mistake and change in circumstance are very similar. In a case involving a mistake or change in circumstance, the facts fail to match the shared assumptions of the parties entering into the transaction. The difference is that, with mistake, the facts mismatched at the time of contracting, and with change in circumstance, the facts mismatch after contract formation. The similarity of the two doctrines allows them to be explored simultaneously in this note.

The doctrine of mistake has experienced an evolution analogous to the doctrine of consideration. In the 1887 decision, *Sherwood v. Walker*, Sherwood arranged for, and confirmed in writing with Walker, the purchase and delivery of a barren cow.³⁰ Prior to his accepting payment and prior to delivery, however, Walker found out that the cow was not barren; rather, it was pregnant.³¹ After Walker made this determination, he refused payment and refused to deliver the cow, and the buyer sued.³² The court found that "the mistake affected the character of the animal for all time . . . for its present and ultimate use . . . [because s]he was not a barren cow, and, if this fact had been known, there would have been no contract."³³ The court reversed and remanded the case, despite the trial court's and the appellate court's respective rulings in favor of the buyer, for further consideration based on that holding.³⁴

In *Sherwood*, formal systems of contract theory were alive and well because the parties executed the agreement in writing, and the parties possessed correspondences to demonstrate their mutual assent. Nonetheless, with the written agreement at issue clearly mistaken as to the heart of the agreement, the court found that writing lacking in enforceability.³⁵

30 *Sherwood v. Walker*, 66 Mich. 568, 569–79 (1887).

31 *Id.* at 572.

32 *Id.*

33 *Id.* at 577–78.

34 *Id.* at 578.

35 *Id.*

Drafters of the Restatement (First) of Contracts confronted this issue in section 504, which states:

[W]here both parties have an identical intention as to the terms to be embodied in a proposed written conveyance, assignment, contract or discharge, and a writing executed by them is materially at variance with that intention, either party can get a decree that the writing shall be reformed so that it shall express the intention of the parties, if innocent third persons will not be unfairly affected thereby.³⁶

According to The Restatement (First) of Contracts, mistakes were only voidable, “where parties on entering into a transaction that affects their contractual relations are both under a mistake regarding a fact assumed by them as the basis on which they entered into the transaction . . .”³⁷ The Restatement (First) of Contracts did not leave much remedy for parties that committed a unilateral mistake.

When it is said that a man’s act is a mistake, the word may refer to the propriety of the act either with reference to the accompanying circumstances, or with reference to the man’s intentions. In the Restatement of this Subject the first meaning is that attached to the word. Mistake in this sense produces no legal consequences unless some act otherwise capable of legal consequences is done while the state of mind exists.³⁸

The Restatement (First) of Contracts section 503 continues:

A mistake of only one party that forms the basis on which he enters into a transaction does not of itself render the transaction voidable; nor do such mistakes of both parties if the respective mistakes relate to different matters; but if the mistakes relate to the

36 RESTAMENT OF THE LAW OF CONTRACTS § 504 (AM. LAW INST. 1932).

37 RESTAMENT (FIRST) OF CONTRACTS § 502 (AM. LAW INST. 1932) (emphasis added).

38 *Id.* § 500 cmt. a.

same matter, the power of avoidance is not precluded because the mistakes of the parties as to that fact are not the same.³⁹

The Restatement (First) of Contracts lacked room for recovery for a party that made a unilateral mistake, except in cases that resembled fraud.⁴⁰

The Restatement (Second) of Contracts made more room for unilaterally mistaken parties to seek recovery. The Restatement (Second) of Contracts section 151 defines “mistake” as, “a belief that is not in accord with the facts.”⁴¹ The Restatement (Second) of Contracts section 153 states:

Where a mistake of one party at the time a contract was made as to a basic assumption on which he made the contract has a material effect on the agreed exchange of performances that is adverse to him, the contract is voidable by him if he does not bear the risk of the mistake under the rule stated in § 154, and

- (a) the effect of the mistake is such that enforcement of the contract would be unconscionable, or
- (b) the other party had reason to know of the mistake or his fault caused the mistake.⁴²

The Restatement (Second) of Contracts section 154 clarifies that:

A party bears the risk of a mistake when

- (a) the risk is allocated to him by agreement of the parties, or

39 RESTatement OF THE LAW OF CONTRACTS § 503 (AM. LAW INST. 1932).

40 See *id.* § 505 (“[I]f one party at the time of the execution of a written instrument knows not only that the writing does not accurately express the intention of the other party as to the terms to be embodied therein, but knows what that intention is, the latter can have the writing reformed so that it will express that intention.”); see also *id.* § 71(c) (“If either party knows that the other does not intend what his words or other acts express, this knowledge prevents such words or other acts from being operative as an offer or an acceptance.”).

41 RESTatement (SECOND) OF CONTRACTS § 151 (AM. LAW INST. 1981).

42 RESTatement (SECOND) OF CONTRACTS § 153 (AM. LAW INST. 1981).

- (b) he is aware, at the time the contract is made, that he has only limited knowledge with respect to the facts to which the mistake relates but treats his limited knowledge as sufficient, or
- (c) the risk is allocated to him by the court on the ground that it is reasonable in the circumstances to do so.⁴³

The development of the doctrine of unilateral mistake is important to this discussion because it reflects the fact that, initially, contract law was adamant about formal systems. Prior to the Restatement (Second) of Contracts, unilaterally mistaken parties lacked substantial opportunity for recovery. The Restatement (First) of Contracts valued formal systems, preserved the idea that formality was the salient issue in analyzing contractual liability, and lacked sympathy for a mistaken party that satisfied the formal requirements. By opening the door to recovery for unilaterally mistaken parties, drafters of the Restatement (Second) of Contracts demonstrate the trend to place less emphasis on formality, and more importance on equity. Therefore, if *Sherwood* involved a good faith transaction in which the buyer thought he was purchasing a fertile cow and the seller accurately knew it barren, prior to 1981, that buyer may have lacked avenues for recovery altogether.⁴⁴

At this point, contract law and its development with respect to the doctrines of consideration and mistake, has been introduced. The following section introduces the theories that are intended to add clarity to this discussion, before making conclusions regarding their application to the teardrop example.

III. Theorists

A. Grant Gilmore

“[Estopel...] is simply a way of saying that, for reasons which the court does not care to discuss, there must be judgment for plaintiff.”⁴⁵

1. Contract as Reliance

Grant Gilmore is credited with capturing and articulating

43 RESTAMENT (SECOND) OF CONTRACTS § 154 (AM. LAW INST. 1981).

44 See *Sherwood*, 66 Mich. at 576.

45 GILMORE, *supra* note 14, at 71.

an American legal movement that indicates a wholesale rejection of the classical turn contract law took at the end of the nineteenth century. Gilmore suggests that contract law is better treated as a subset of torts, and that contract law practitioners and theorists alike should return to analyzing contractual liability in terms of the duties individual parties owe to one another.⁴⁶ The idea that there is a specific/identifiable moment where an agreement becomes legally enforceable is central to classical contract law theory. As such, it is not surprising that the reliance-based theory of contract law that Gilmore observes and articulates in his book, *The Death of Contract*, does not always result in a clear moment where contractual liability suddenly exists.

Gilmore argues that as contract law theorists push for external indicators to be used in analyzing contractual liability, contract law needs section 90 recoveries, which resemble recoveries in tort.⁴⁷ Gilmore explains that the Restatement (First) of Contract section 75 reflects traditional values associated with contractual liability, like bargained-for consideration.⁴⁸ Despite drafters' attempts to preserve classical contract law principles, Gilmore argues that in 1932, with the introduction of the Restatement (First) of Contracts section 90, contract law started to recognize something at least akin to a tort recovery under a reliance-based theory.⁴⁹

According to Gilmore, the American movement away from the benefit-detriment theory towards the bargain-theory of consideration demonstrates the classical approach's planting the seeds of its own demise. Gilmore explains:

Jurisprudential thought was that the inevitable process of legal development was from a starting point at which rules of law are based on or derived from judgments about subjective moral fault or culpability toward an end point at which the original moral content of a given rule will have quite disappeared and the subjective state of mind of the defendant will have become irrelevant. The law moved from 'subjective' to 'objective,' from 'internal' to 'external,'

46 See generally GILMORE, *supra* note 1.

47 *Id.*

48 *Id.* at 23.

49 See, *supra* Part II(B) (discussing differences between each respective revision).

from ‘informal’ to ‘formal.’⁵⁰

Among the benefits of an objective undertaking of contractual liability was the ability to point to a specific moment in time when formal systems of traditional contracting were satisfied, and contractual liability began to exist. However, as the law becomes objective, the law simultaneously becomes indifferent to individuals who reasonably and foreseeably rely on promises but fail to satisfy those formal systems of traditional contracting, and courts wished to enforce reasonable and foreseeable reliance. The introduction of section 90 to the Restatement (First) of Contract addressed this issue and advanced the “death of contract” theory.

Perhaps section 90 represents social desire to find contractual liability in cases that lack contracting formalities but involve a duty to fulfill a promise. Gilmore points out the poignant tension between the second restatement’s sections 75 and 90.⁵¹ With the addition of sections 75 and 90, a contract dispute that never involved certain formal contracting systems, like bargained-for consideration, can result in a recovery prescribed by contract law. In fact, contractual remedies are consistently available to parties that never contracted according to section 71.⁵² Gilmore observes that the presence of contractual liability, absent the consideration required in section 71, indicates an evolution to recognize social constructs, like duties, in the face of formal contracting systems.⁵³

2. Contract as Reliance and Lack of Consideration

A reliance-based theory of contractual liability is rather unconcerned with consideration. The movement towards a reliance-based theory of contractual liability represents trends “[d]uring the past forty years [that indicate] the effective dismantling of the formal system of classical contract theory.”⁵⁴ Gilmore explains, classical contract law theorists struggled to incorporate cases lacking consideration into their understanding of contract liability because many of the cases involved fact patterns lacking in “reciprocal

50 GILMORE, *supra* note 1, at 46.

51 *Id.* at 79–80.

52 See, e.g., N. Litterio & Co. v. Glassman Const. Co., 319 F.2d 736 (1963) (finding no contractual agreement took place, but reversing and remanding to determine whether or not § 90 warrants recovery).

53 See generally, GILMORE, *supra* note 1.

54 GILMORE, *supra* note 1, at 72.

conventional inducement.”⁵⁵

Classical theorists argued “that these were bad cases” and that despite New York’s acceptance of the strict bargain theory of consideration, “many judges, it appeared, were not prepared to look with stony-eyed indifference on the plight of a plaintiff who had, to his detriment, relied on a defendant’s assurances without the protection of a formal contract.”⁵⁶ But, they were not “bad cases;” on the contrary, people’s livelihoods were at stake, and often (and rightly so), courts wanted to find in their favor, despite the inability to point to the objective criteria touted by classicalists.

On a reliance-based theory, an agreement is enforceable when it is apparent that a contracting party reasonably and foreseeably relied on another contracting party’s promise. Whether or not there was consideration is, simply, not the issue to be resolved in order to determine whether or not an enforceable agreement exists under a reliance-based theory. Rather, the salient issue is the reasonableness and foreseeability of the promisee’s reliance. The reliance-based theory is unfettered by problems with consideration, making a finding of contractual liability under these circumstances akin to a tort recovery that identifies a breach of a duty as a salient issue in determining civil liability.

3. Contract as Reliance and Mistake/Change in Circumstance

Similarly, the reliance-based theory is unconcerned with a lack of formal systems in examples involving mistake and change in circumstance. Consider the change in circumstance case, *Harris v. Watson*, in which seamen petitioned the court for extra wages promised by their master when circumstances changed, and the ship was in danger; the master promised the extra wages to induce the seamen to work harder.⁵⁷ Lord Kenyon noted in his decision that a ruling in the seamen’s favor would result in vessels being held hostage for greater pay, and declined to find in the seamen’s favor.⁵⁸ With a very similar fact pattern, *Stilk v. Myrick* extended the *Harris* holding with Lord Ellenborough’s decision that the second agreement for higher wages for seamen was unenforceable, because the seamen were already legally bound to the work (pursuant to

55 *Id.* at 70.

56 *Id.*

57 *Harris v. Watson*, (1791) 170 Eng. Rep. 94.

58 GILMORE, *supra* note 1, at 27–28.

admiralty law at the time, seamen must continue in extraordinary circumstances) and as such there was no legal benefit to the master to change the prior agreement.⁵⁹ In their decisions, the courts appeal respectively to public policy and want of consideration.⁶⁰ Williston argues:

[T]he [Stilk] agreement . . . is invalid for the performance by the recalcitrant contractor is no legal detriment to him whether actually given or promised, since, at the time the second agreement was entered into, he was already bound to do the work; nor is the performance under the second agreement a legal benefit to the promisor since he was already entitled to have the work done.⁶¹

Williston finds the decisions consistent with consideration theory where consideration must be given for modifications to a pre-existing contract.

Gilmore counters, “It seems unlikely that a neutral observer, not yet corrupted by the study of law, would jump to that conclusion.”⁶² Gilmore suggests that insofar as *Harris* and *Stilk* relied on formal systems to resolve the disputes, those decisions may have been unconcerned with equity.⁶³ Gilmore points out that Lord Kenyon appealed to public policy, and (in his opinion written over a decade later) Lord Ellenborough relied on want of consideration.⁶⁴ At the time of the *Harris* decision, courts were still trying to make sense of when an agreement was enforceable using the equitable tools available. By the time of *Stilk*, consideration required giving up a legal right. The two decisions indicate the move into bargained-for consideration. They also demonstrate how courts depended on formal systems to analyze cases involving change in circumstance.

The *Harris* and *Stilk* holdings rely on the theory that the seamen did no more than originally promised. Gilmore explains that “[i]f the ambiguities in the factual situation were resolved in favor

59 *Stylk v. Merick* 170 Eng. Rep. 1168, 1169 (1809).

60 The discrepancy in the respective courts’ rationale is addressed below.

61 SAMUEL WILLISTON, THE LAW OF CONTRACTS § 130, at 276 (Baker, Voorhis & Co.) (1920).

62 GILMORE, *supra* note 1, at 25.

63 *Id.* at 29–30.

64 *Id.* at 29.

of the crew, it is quite possible that, under maritime law, they would have been justified in quitting the service of the ship... [but c]ounsel for the plaintiff made no attempt to investigate these questions."⁶⁵ Gilmore points out that there may have been a safety valve for the seamen inherent in maritime law, but that issue was unresolved.⁶⁶

Perhaps that safety valve could have made a reliance-based recovery available. Were section 90 available to the seamen, a lack of consideration may have been trivial if the seamen could have demonstrated a reasonable and foreseeable reliance on their masters' new promises. In this way, a reliance-based theory of contractual liability effectively disposes of issues relating to change in circumstance and mistake by redirecting the fact-finder's focus away from issues with consideration and onto reasonably foreseeable reliance. By focusing the issue on reliance, a failure to observe formal systems becomes less salient to disposition.

B. Charles Fried

"The moral force behind contract as promise is autonomy: the parties are bound to their contract because they have chosen to be."⁶⁷

1. Contract as Promise

Charles Fried notably argued for a concept of contractual liability based in the intuitive moral principles that effectuate promises. Contract as promise depends on the autonomy of individuals entering into agreements because in order for an agreement to be effectuated, a contracting party must be able to freely make a fair promise, rationally and deliberately, and the "promisor must have been serious enough that subsequent legal enforcement was an aspect of what he should have contemplated at the time he promised."⁶⁸ According to Fried, parties are morally bound to a contract because they have chosen to be, as evidenced by their promises.⁶⁹ For contract as promise, it is at the moment where a rational, deliberate, and serious promise becomes morally salient to a contracting party that contractual liability results.

65 *Id.* at 26.

66 *Id.*

67 FRIED, *supra* note 2, at 57.

68 *Id.* at 38.

69 *Id.* at 57.

Contract as promise trends towards valuing the individual ability to willfully take on obligations. As such, contract as promise attempts to revive a value of classical contract theory in the face of the contract is dead thesis. Fried argues that promissory obligations are entitled to a special legal status, and that promise is the primary indicator of contractual liability.⁷⁰ The result is that individual actors are empowered to contract how they see fit, regardless of social constructs, like duties.

Promissory obligations are not tied to language or to formal legal conventions. Rather, Fried argues that promissory obligation comes from moral principles.⁷¹ Although a party could have a greater sense of moral obligation were the other party's objective signs of assent to the agreement extremely clear, a party could still have a sense of moral obligation to fulfill a promise when formal aspects of contracting are absent, and assent may be questionable.⁷² Consider the common contractor-subcontractor example in which a subcontractor gives a quote to a contractor who relies on that quote and places a bid on a project. Fried explains that, although consideration theory does not require the subcontractor to perform based on the quote, an appeal to moral principles seems to evoke the desire to have that result.⁷³

Contract as promise is very inclusive of informal agreements in terms of legal liability. When promissory obligation indicates existence of contractual liability, any genuine promise becomes enforceable. The unattractiveness of the theory's inclusiveness of promissory agreements is countered by the thesis's straightforwardness.

Acceptance in contract as promise can only result from autonomous action, thereby showing "the moral relation of promising to be voluntary on both sides."⁷⁴ Acceptance can be difficult to recognize under Fried's concept. Under Fried's theory, tacit acceptance functions as acceptance, and acceptance may be inferred.⁷⁵ Consider the example of the neighborhood boy promised \$10 to shovel the snow from a promisor's driveway. If the neighborhood boy walks by the initial promisor without indicating assent, but the next morning, when the promisor is rushing to get to work, the driveway is cleared

70 See generally FRIED, *supra* note 2.

71 *Id.*

72 *Id.*

73 *Id.*

74 *Id.* at 43.

75 *Id.*

of snow, certainly, the promisor owes the neighborhood boy \$10. Fried argues that objective assurances of assent are important when considering whether or not an agreement exists, but also maintains that (in certain circumstances) a promise made and not revoked evokes moral obligation upon full performance by the promisee.⁷⁶ In this light, Fried accounts for examples involving an option contract created by part performance or tender, also referred to as an “offer for a unilateral contract.”⁷⁷

2. Contract as Promise and Lack of Consideration

A successful application of Fried’s theory is relatively indifferent to the doctrine of consideration. In fact, Fried argues that consideration theory fails to function as a consistent sorting mechanism.⁷⁸ Fried explains that there are “two guiding premises of the doctrine of consideration: (A) that only promises given as part of a bargain are enforceable; (B) that whether there is a bargain or not is a formal question only.”⁷⁹ While these two tenets of consideration theory seem to reflect the primary thrust of consideration, Fried explains, they are, undeniably, at odds with each other because (A) restricts the class of arrangements to bargains, and (B) affirms the liberal principle that the free arrangements of rational persons should be respected.⁸⁰ According to Fried, (A) limits enforceable agreements to those that included formal bargaining, and (B) makes formal bargaining accessible to autonomous actors.⁸¹ Fried sees this as a contradiction because autonomous actors sometimes fail to make use of formal systems of contracting, as in the teardrop example, and contract as promise attempts to impose contractual liability on all agreements involving promissory obligation. Whether or not his criticism is true, Fried sees consideration theory as an attack on individual autonomy with respect to contracting because consideration theory can sort out agreements that Fried finds

76 See generally *id.* at 43–45.

77 “Where an offer invites an offeree to accept by rendering a performance and does not invite a promissory acceptance, an option contract is created when the offeree tenders or begins the invited performance or tenders a beginning of it.” RESTATEMENT (SECOND) OF CONTRACTS § 45 (AM. LAW INST. 1981).

78 FRIED, *supra* note 2, at 33.

79 *Id.*

80 *Id.* at 35.

81 See *id.*

enforceable under contract as promise.⁸² The criticism exposes Fried's respect for private ordering.

Contract as promise seeks to account for the social value of contracts while denying the relevance of consideration. The reason contracts are enforceable is not because of the various doctrinal indicators that drafters have identified, but because of the moral underpinnings to a promise, without which, society would be paralyzed. Fried argues that promise, as a convention, is so imbedded into human understanding as to how to act in society that a person's self-interest contributes to a person's instinct to fulfill a promise.⁸³ To some extent, individuals are self-interested when it comes to acting morally, and that sense of morality advances promissory obligations towards fulfillment.⁸⁴ Fried notes that external sanctions play a supplemental role, but focuses his analysis on the idea that contracting parties satisfy self-imposed moral obligations when fulfilling promises.⁸⁵ The success of Fried's concept is relatively unconcerned with bargained-for consideration. When moral promissory obligation invokes a legal requirement to fulfill promises, consideration becomes irrelevant.

3. Contract as Promise and Mistake/Change in Circumstance

Fried does not intend his theory to be taken in a vacuum. Fried discusses the necessity for common law to fill the "gaps" in his theory, and argues that his reliance on common law does not mean his theory fails.⁸⁶ Gaps represent legal quandaries that a given theory cannot explain away.

In order for a party to make an enforceable promise with a desired outcome in mind, that party must have considered the circumstances giving rise to the promise, and the circumstances that will eventuate the desired outcome. When making a promise, a promisor engages in a "promising calculus" that considers the status quo in order to understand how an exchange of intentional promises could eventuate a desirable outcome. But, if the promisor is mistaken about the current state of affairs when engaging in that calculus,

82 See generally, FRIED, *supra* note 2.

83 *Id.* at 14–16.

84 *Id.*

85 *Id.* at 14–17.

86 *Id.* at 58–73.

then the promisor was not able to form the initial intent required to invoke moral obligation. A promisor lacks autonomy when the promisor is unable to understand how circumstances will affect the desired outcome. Without autonomy and an ability to form the initial intent to bring about a potential outcome, contract as promise fails to resolve the tension that results.⁸⁷ Without autonomy, there is no moral promissory obligation. Fried identifies mistake and change in circumstance as a gap in his theory.⁸⁸

Furthermore, as a practical point, when courts determine who should bear the burden that results from a mistake, courts look to the factors identified in the Restatement (Second) of Contracts section 154, which do not consider initial intent (unless the parties contracted for that mistake). A party that did not foresee a mistake during contracting will bear the burden if “he is aware, at the time the contract is made, that he has only limited knowledge . . . but treats his limited knowledge as sufficient” or if “the risk is allocated to him by the court on the ground that it is reasonable in the circumstances to do so.”⁸⁹ The courts are unconcerned with what was in the parties’ minds while contracting. Fried admits, “The further courts are from the boundary between interpretation and interpolation, the further they are from the moral basis of the promise principle and the more palpably are they imposing an agreement.”⁹⁰ As such, contract as promise cannot directly account for mistake and change in circumstance, nor can the theory account for the courts’ ensuing deliberations pursuant to the Restatement (Second) of Contracts section 154. Fried continues, as courts “move further from actual intention the standard of presumed intention tends to merge into the other substantive standards used to solve the problems caused by a failure in the agreement.”⁹¹

C. Randy Barnett

“You are allowed to do what you wish with what is yours (as defined by property law), provided that you do not infringe (as defined by tort law) on the property of others—including the inalienable

87 *Id.* at 58–60.

88 *Id.* at 58.

89 RESTATEMENT (SECOND) OF CONTRACTS § 154(b)–(c) (AM. LAW INST. 1981).

90 FRIED, *supra* note 2, at 61.

91 *Id.*

property rights one has in one's own body. Contract law provides the means by which people can transfer their property to another by their consent.”⁹²

1. *Contract as Consent*

With contract as consent, Randy Barnett proposes to solve many of the problems that Fried admits contract as promise encounters. Barnett explains that courts should enforce private commitments when there exists an objectively “manifested intention” to create a legal relation, which Barnett calls “consent.”⁹³ Barnett explains that an objectively manifested intention to contract (i.e. consent) is the most effective objective indicator of the subjective intent to contract.⁹⁴ Whereas a promise is a commitment to perform or to refrain from performing some activity, to consent to contract is to accept that legal liability will be apportioned upon failure to fulfill a promise (breach).⁹⁵

Contract as consent recognizes the force behind contract as promise, but argues that mere promise exists mostly in the social realm in which people exist. To Barnett, legal enforceability needs a second layer demonstrating the willingness to involve the law.⁹⁶ To determine whether or not a legally enforceable agreement was made pursuant to contract as consent, a confused party to a potential contract may consult both its own feeling as to whether or not it feels obligated to fulfill a promise and any objective indications that both parties to the contract consented to legal liability.⁹⁷ A promise and subsequent objective manifestations of consent to legal liability in relation to that promise result in the existence of contractual liability under contract as consent.⁹⁸

Barnett’s contract as consent restores a second, somewhat formal layer to the framework Fried introduced as contract as promise. Contract as consent does not propose the legal enforceability of many informal promises because an individual must demonstrate an intent to invoke contractual liability in order to consent, and consent is essential to justiciability. The result is that some informal

92 BARNETT, *supra* note 3, at 72.

93 Randy E. Barnett, *Contract is Not Promise; Contract is Consent*, 45 SUFFOLK U. L. REV. 647, 655 (2012).

94 *See id.*

95 *Id.*

96 *See id.*

97 See generally Barnett, *supra* note 92.

98 *Id.*

agreements lacking in objective indicators of consent that parties, nonetheless, desire to invoke legal protection,⁹⁹ lack enforceability. As such, Barnett's concept values both the individual autonomy inherent in private ordering and formal contracting systems.

Although the important social function of promising is inherent in enforceable contracts, in consent Barnett requires more than Fried to evoke contractual liability. Where Fried casts a wide net over agreements qualified as contractual, Barnett argues that net is too wide and narrows the field of enforceable agreements using objective consent to determine contractual liability.

2. Contract as Consent and Lack of Consideration

Requiring consent attempts to make objective tools more available for evaluating contractual liability. Among those objective tools is bargained-for consideration. Barnett explains, "formal bargaining usually manifests an intention to be legally bound."¹⁰⁰ Bargained-for consideration can serve as the objective indicator of consent.¹⁰¹ Therefore, contract as consent represents somewhat of an attempt to revive formal systems in the face of Gilmore's reliance-based theory and Fried's subjectivism. The result is that objective sorting mechanisms play an essential role in contract as consent.

3. Contract as Consent and Mistake/Change in Circumstance

To account for mistake and change of circumstance, Barnett distinguishes between intrinsic and extrinsic circumstances.¹⁰² According to Barnett, there is no legal duty to disclose extrinsic circumstances, which affect the supply of or demand for a particular item.¹⁰³ In this way, if the mistake or change in circumstance was the result of the failure to disclose extrinsic circumstances, then contract as consent may still account for a court's enforcement of this type of agreement, because the amount of knowledge legally required to

99 See, e.g., *McIntosh v. Murphy*, 469 P.2d 177 (Haw. 1970) (finding lack of writing, but enforceable agreement on promissory estoppel/reliance grounds).

100 BARNETT, *supra* note 3, at 153.

101 *See id.*

102 *Id.* at 240–42.

103 *Id.* at 240, 246.

consent to legal liability would have been present.¹⁰⁴

In contrast to extrinsic facts, Barnett explains that there is a presumptive duty to disclose an intrinsic fact, which concerns information about the item itself.¹⁰⁵ The logic behind this necessity parallels the reason that Fried's theory fails to account for holdings that enforce an agreement that failed to account for a mistake or change in circumstance. In order for a party to intend to invoke legal liability on a given agreement, that party must understand the "intrinsic" circumstances attending to that agreement; otherwise, the intention is to invoke legal liability on an agreement that is different than the actual agreement at hand. Intention to invoke liability, whether in contract as promise or consent, requires knowledge of "intrinsic" facts.

Consider again *Harris v. Watson* and *Stilk v. Myrick*. Barnett argues these change in circumstance cases fail because of a lack of consideration and labels them examples of "opportunistic contracting behavior" in which one party is not free to walk away.¹⁰⁶ Despite the lack of information available regarding the underlying facts in *Stilk*, Barnett appears to take a classical view on the *Harris* and *Stilk* decisions, by arguing that the lack of consideration (i.e. consent) rendered those modifications unenforceable.¹⁰⁷ Barnett thinks that a lack of consent sorts out those agreements that are opportunistic, and renders them unenforceable.¹⁰⁸ By portraying these decisions as lack of consideration cases, Barnett skirts the issue of reasonable and foreseeable reliance.

IV. Application

A. Example

The example presented at the onset of this project yields varying results when each theory is applied. Each theory asks

¹⁰⁴ See *id.* at 238–39. See also *Laidlaw v. Organ*, 15 U.S. (2 Wheat.) 178 (1817), in which Chief Justice John Marshall introduces the intrinsic/extrinsic distinction, and opines that the relevant party "was not bound to communicate [extrinsic circumstances]."

¹⁰⁵ BARNETT, *supra* note 3, at 240–42. See, e.g., *Obde v. Schelmeyer*, 56 Wash. 449 (1960) (holding that termite infestation needed to be disclosed by buyer). This supports Barnett's argument that there is a duty to disclose intrinsic facts.

¹⁰⁶ BARNETT, *supra* note 3, at 159.

¹⁰⁷ *Id.*

¹⁰⁸ *Id.* at 161.

different questions and focuses on different dimensions of a proposed agreement. Where the reliance-based theory asks whether or not reasonably foreseeable reliance was induced by a promise, contract as promise asks a moral actor to inquire as to whether or not a promise was made, and contract as consent asks whether a promise was made in conjunction with an indication that the agreement would invoke legal liability.

Turning first to Gilmore's reliance-based theory, contractual liability depends on the amount of reasonable reliance induced by the promise. As a result, the inquiry would delve into the reasonableness of Sam's reliance on my promise. Additional considerations would include what, if any, unfairness would result from either enforcement or non-enforcement of the original agreement.

In this case, Sam's reliance seems, at best, modest. He did not forego alternative opportunities to make money expecting that I would follow through on my promise. Furthermore, he did not go buy the solvent that spared him so much time cleaning. Under a reliance-based reading, if I were to breach the agreement, a court in equity might not award any damages.

Sam's reliance was reasonable only insofar as he did not know about my misinformation. If he was unaware of my lack of knowledge with regards to the cleaning solvent, then perhaps his reliance was reasonably foreseeable, and a court would offer damages consistent with Sam's performance. If Sam knew or should have known about my misinformation, according to a reliance-based reading, the reasonableness of his reliance decreases. A reliance-based reading results in a contractual liability that causally correlates to what the promisee should have known with respect to the promisor's understanding of circumstances. If Sam took advantage of my lack of knowledge, then the reasonableness of Sam's reliance declines.

Fried's contract as promise appears to quickly include my agreement with Sam amongst those benefitting from contractual liability. I made a promise to Sam, and despite my lack of knowledge with regard to the cleaning solvent, I felt obligated to perform. For Fried, that internal dialogue is all that is necessary. However, the nuance that I did not know about the cleaning solvent brings into question whether or not I intended for the particular agreement at issue.

Fried's theory experiences difficulty dealing with the ethical quandaries that result when parties' promises do not anticipate mistake or change in circumstance. Fried points out that "contracts

are largely a deliberate attempt to deal with uncertainty[,]” and parties may contract for mistake and change in circumstance.¹⁰⁹ For instance, I could have told Sam that if he ends up finding a solvent that saves him considerable time, I’d expect to pay less. But I did not express this condition. Fried explains, “The court cannot enforce the will of the parties [when] there are not concordant wills. Judgment must therefore be based on principles external to the will of the parties.”¹¹⁰ Fried argues that there is a boundary between interpreting what was agreed to, and “interpolating terms” that the parties did not agree to, but probably would have.¹¹¹ When a court attempts to “interpolate terms,” and uses gap fillers like the Restatement (Second) of Contracts section 154, Fried’s theory has nothing to say by way of mitigating the unforeseen damages. Quite simply, I did not intend to exchange my promise of \$50 for ten minutes of work. Contract as promise fails to explain why I felt obligated to perform on my promise.

Barnett’s contract as consent requires a finding as to whether or not there was any objective indication of consent by both parties. This agreement was not subject to a writing; Sam and I did not even shake hands. Contract as consent forces parties to consider whether or not a given agreement is of the type upon which those parties want to impose legal liability. Many critics may argue this agreement does not deserve legal liability because it is too informal. Certainly, my mother would have been upset had Sam taken me to court.

If we suppose the example included some objective indication of consent to legal liability, contract as consent asks whether or not the existence of the solvent was an intrinsic or an extrinsic fact, because Sam knew about the existence of the solvent and I did not.¹¹² Presumably, the existence of the solvent affects the supply or demand of car washes because it speaks to the difficulty of the job, as opposed to how one completes the job. The existence of the solvent “concerns information about the item itself[,]” namely, the car wash.¹¹³ So, had Sam known I was ignorant as to the existence of the solvent, Barnett would have Sam disclose the intrinsic information that is the existence of the solvent, making the agreement unenforceable.¹¹⁴

¹⁰⁹ FRIED, *supra* note 2, at 59.

¹¹⁰ *Id.* at 60.

¹¹¹ *Id.*

¹¹² BARNETT, *supra* note 3, at 240.

¹¹³ *Id.* at 241.

¹¹⁴ The discussion in Barnett’s book that speaks to this issue is primarily

But this does not seem right, either. Was it Sam's fault that I was unfamiliar with the items in both of our father's garage? While my mother would have been upset had Sam taken me to court, she also likely would have called me "lazy" and told me that my ignorance was my own fault. (Perhaps you can infer who is the judge and jury in our family.)

Application of the Restatement (Second) of Contracts section 154(b) arguably favors Sam, because by estimating the amount of time it would take, there is a reasonable inference that I was "aware, at the time the contract [was] made, that [I had] only limited knowledge with respect to the facts to which the mistake relates but treat[ed this] limited knowledge as sufficient . . ." As such, it would seem, when I headed to the fridge instead of to investigate the tools available to complete the task, the mistake became my burden. But this still does not explain where my sense of obligation originated from.

B. Morality in Contract as Reliance, Promise, and Consent

A personally felt obligation to perform on a promise stems from an individual sense of morality. Holmes explains, "The law is the witness and external deposit of our moral life."¹¹⁵ The fact that law derives principles from morality is undisputable. A legal system that did not conform with society's sense of morality would fail. This may seem obvious, yet the drafters of the Uniform Commercial Code prescribe a relationship between custom and liability.¹¹⁶

concerned with trading on information. As a result, how necessary disclosure relates to intrinsic and extrinsic facts related to services remains somewhat ambiguous. *See id.*

115 Oliver Wendell Holmes, Jr., *The Path of the Law*, 10 HARV. L. REV. 457, 459 (1887).

116 *See U.C.C. § 1-103 (AM. LAW INST. & UNIF. LAW COMM'N 2014)*

(a) [The Uniform Commercial Code] must be liberally construed and applied to promote its underlying purposes and policies, which are:

(1) to simplify, clarify, and modernize the law governing commercial transactions;

(2) to permit the continued expansion of commercial practices through custom, usage, and agreement of the parties; and

(3) to make uniform the law among the various jurisdictions.

(b) Unless displaced by the particular provisions of the [Uniform Commercial Code], the principles of law and equity, including the law merchant and the law relative to capacity to contract, principal and

Furthermore, each of the three theories explored in this essay informs its respective concept of contract law with morality, but each goes about that process differently. Barnett draws a helpful distinction; “ethics” refers to the private morality governing how one ought to live one’s life, and “justice” is the public moral domain defined by instances when “coercion may be justifiably used against an individual for breaching a duty.”¹¹⁷ Fried’s contract as promise finds its moral thrust in ethics; that is, in the moral motivation one feels after making a promise. Indeed, Benjamin Franklin identifies the sense that one ought to follow through on a promise as a virtue.¹¹⁸

According to Barnett, Fried “focused on the wrong morality.”¹¹⁹ Barnett’s contract as consent derives moral thrust from the public domain, because when parties consent to be legally bound, they open themselves to the will of the public morality machine: namely, the justice system. Barnett argues for a dichotomy between these two domains. While promises have an ethical charge, without a consent to contractual liability, Barnett argues, there should be no imposition of public morality.

Meanwhile, Gilmore’s contract as reliance exists somewhere in between. Contract as reliance draws on both the moral idea that inducing reliance and not following through is ethically wrong, and the social interest in ensuring justice in these scenarios in a public venue.

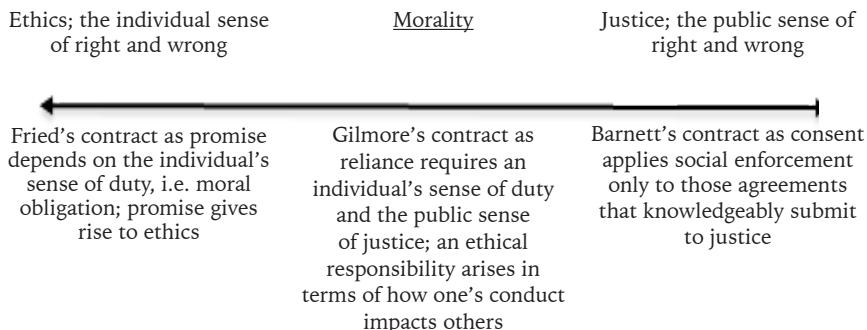
Consider the graphic below, which attempts to illustrate this spectrum:

agent, estoppel, fraud, misrepresentation, duress, coercion, mistake, bankruptcy, and other validating or invalidating cause supplement its provisions.”)(emphasis added).

117 Barnett, *supra* note 92, at 663.

118 BENJAMIN FRANKLIN, THE AUTOBIOGRAPHY OF BENJAMIN FRANKLIN 215–16 (John Bigelow ed., 1868), (describing that the fourth of Franklin’s thirteen virtues is “Resolution. Resolve to perform what you ought; perform without fail what you resolve.”).

119 Barnett, *supra* note 91, at 647.



Although Fried may appear misguided by attributing the effectiveness of his theory to individual moral sentiment, contract as promise highly values “the liberty to engage in private ordering.”¹²⁰ By allowing for the enforceability of all promises, private ordering of all types retains enforceability. Gilmore’s contract as reliance does not value private ordering to such a great extent by recognizing agreements based on reliance, instead of on the basic morally obligatory nature of an exchange of promises between two parties. Meanwhile, contract as consent represents the opposite end of the freedom to contract spectrum because private ordering without specific formalities may fail to be justiciable.

Contract as promise represents an attractive “deontological justification for contract law as against theories grounded essentially in consequentialism (the underlying moral basis of welfare economics) or sociology.”¹²¹ Certainly, Fried captures a desirable notion: that people should follow through on their promises, so all promises should be enforceable. While contract as promise may represent an attractive model, it does not always reflect how the legal industry functions. Fried’s theory in reality would likely result in individuals and business entities, simply, not making promises. Rather, people would be confined to soft promises, like, “I will do my best to deliver on Monday,” or, “If my supplier delivers on time it should be here by tomorrow.”¹²²

In practice, the relationship between the moral authority of promises and legal liability for actors that lack a driving moral

120 Jean Braucher, *The Sacred and Profane Contracts Machine: The Complex Morality of Contract Law in Action*, 45 SUFFOLK U. L. REV. 667, 668 (2012).

121 Jeffrey M. Lipshaw, *Contract As Meaning: An Introduction to “Contract As Promise at 30”*, 45 SUFFOLK U. L. REV. 601, 602 (2012).

122 Braucher, *supra* note 120, at 684–85.

compass is more complicated than Fried explains in his book. Consider large corporations, which are not capable of moral thought. Add “the complication that corporate managers have fiduciary obligations to shareholders [resulting in] managers [who] are not supposed to be free to indulge their personal morals (an ‘agency cost’), [because indulging in personal morals would mean] losing out in the marketplace to other entities with more ruthless managers who better serve their firm’s shareholders.”¹²³ Simply, the efficacy of contract as promise in the real world falters. Contract as reliance and contract as consent do not waver in the face of a realistic lack of morality because they do not place all of their eggs in the private ethics moral “basket.”¹²⁴

Perhaps the lesson is that not one of the three theories is perfect. Not one of them encapsulates the enforceability of informal agreements while accounting for the nuances of mistake and change in circumstance.

C. Conclusion

Returning, one last time, to the teardrop example, there appears no theory sufficient to explain my sense of obligation to pay Sam. Nonetheless, all three theories speak to the nature of my sense of obligation: morality. Given that my mother did not want to see either Sam or me litigating this issue in court, it would appear this transaction is not of those we wished to invoke legal liability. But, there remains a sense of moral obligation, not because of Sam’s modest reliance, but because of my own sense of right and wrong. Thus, Barnett’s contract as consent accounts for my sense of obligation while recognizing there is no desire to seek legal enforcement.

Perhaps Barnett’s theory, despite its disparaging fallbacks and formal values, represents the most practical and applicable thesis to informal interactions. Certainly, some private ordering

123 *Id.* at 676.

124 In her article, *The Sacred and Profane Contracts Machine: The Complex Morality of Contract Law in Action*, Jean Braucher suggests that these problems could be solved were contract as promise to include an element of forgiveness. Braucher, *supra* note 120, at 684–85. Braucher writes, “The law in action perspective suggests the importance of adjustment, release, and forgiveness in contractual relations as an important principle, perhaps as much so as promise-keeping.” *Id.* at 692. Braucher does not, however, discuss whether forgiveness would taint autonomy and invalidate the necessary intentional undertaking inherent in making a promise. *Id.* at 669.

will lack justiciability, but cases like the teardrop example make sense under this reading. No theory is perfect, but, in terms of this paper's example, contract as consent comes closest to a workable explanation.

* * *

ADR In Japan's Financial Markets & the Rule of Law

Tomohiko Maeda & Andrew M. Pardieck*

Abstract

One practitioner recently described the number of cases resolved in private alternative dispute resolution in Japan simply as “unsatisfactory.” That statement aptly describes efforts to mediate disputes arising between financial service providers and their customers up to 2008. But in 2008 things changed: Japan amended each of the statutes regulating its principal financial industries, including the banking, insurance, and securities industries. In doing so, it established Japan’s Financial ADR System. Claimants have used this system in large numbers: in some industries, they are mediating more claims than they are filing in the courts. And they do so voluntarily. There are no mandatory mediation or arbitration clauses that require a customer to arbitrate or mediate instead of litigate. The system is far from perfect: practitioners and scholars find much to dislike. But there is also evidence to suggest that there are meaningful benefits. It offers speedy and inexpensive resolution of claims that are unlikely to be litigated. And the resolution is, on average, reasonable when measured against litigated outcomes. It reaches this result not by limiting the role of legal professionals and access to the courts, as found in previous statutorily-imposed alternative dispute resolution schemes. Instead, it does so by relying on legal professionals and judicially established norms. The numbers suggest Japan’s Financial ADR System warrants a closer look, which is what this article offers.

* Tomohiko Maeda, Professor of Law, Meijo University and former Research Scholar at the University of Hawai'i William S. Richardson School of Law and the University of Michigan Law School. Andrew M. Pardieck, Associate Professor of Law, Southern Illinois University School of Law. The authors thank Professors Yoshifumi Imagawa, Daniel Foote, Mark Levin, and Luke Nottage for their helpful comments, and Brendan Harvey and Jerrod Montgomery for their assistance with this paper. The authors are solely responsible for any errors or omissions, as well as the translations where English language sources are unavailable. The interviews and correspondence with secretariats of the Financial ADR programs discussed in this work were supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (A) Grant Number 25245002.

Table of Contents

I.	INTRODUCTION.....	403
II.	THE ORIGINS OF JAPAN'S FINANCIAL ADR.....	406
	A. <i>The Loss Compensation Scandal & Resulting Regulatory Scheme</i>	406
	B. <i>The Judiciary's Role in Resolving Financial Service Disputes</i>	409
III.	THE INSTITUTIONALIZATION OF FINANCIAL ADR	411
	A. <i>The Institutional Framework for Japan's Financial DDROs</i>	413
	B. <i>A Procedural Outline-DDRO Complaint & Mediation Processing</i>	418
IV.	DDRO PRACTICE AT FINMAC & THE JBA	426
	A. <i>FINMAC-ADR for the Investment Industry</i>	428
	B. <i>The JBA Consultation Office-ADR for the Banking Industry</i>	433
	C. <i>JBA and FINMAC Caseloads</i>	439
V.	IMPLICATIONS FOR THE RULE OF LAW IN JAPAN	444
	A. <i>The Role of Law</i>	445
	B. <i>The Role of Lawyers</i>	451
	C. <i>The Role of Financial ADR</i>	454
	1. <i>Time to Resolution</i>	454
	2. <i>Cost of Resolution</i>	456
	3. <i>Adequacy of Relief</i>	458
	a. <i>Compensation & Comparative Fault</i>	459
	b. <i>Voluntary Participation & Settlement Rates</i>	463
	c. <i>Voluntary Participation & Procedural Fairness</i>	465
VI.	Conclusion.....	468

I. INTRODUCTION

While scholarship on Japanese law has described, at length, Japan's long history of alternative dispute resolution (ADR),¹ there has been less to discuss of late. Setting aside court-sponsored mediation and civil conciliation, Japan's contemporary ADR programs have met with limited success.² There is, however, an exception: in

-
- 1 See, e.g., TAKEYOSHI KAWASHIMA, NIHONJIN NO HOISHIKI [THE LEGAL CONSCIOUSNESS OF THE JAPANESE] (1967) (Japan); John O. Haley, *The Myth of the Reluctant Litigant*, 4 J. OF JAPANESE STUDIES 359 (1978); FRANK K. UPHAM, LAW & SOCIAL CHANGE IN POSTWAR JAPAN (1987); J. Mark Ramseyer & Minoru Nakazato, *The Rational Litigant: Settlement Amounts and Verdict Rates in Japan*, 18 J. LEGAL STUD. 263, 267–70 (1989); Tom Ginsburg & Glenn Hoetker, *The Unreluctant Litigant? An Empirical Analysis of Japan's Turn to Litigation*, 35 J. LEGAL STUD. 31 (2006); Tom Ginsburg & Glenn Hoetker, *The Effects of Liberalization on Litigation: Notes Toward A Theory in the Context of Japan*, 8 WASH. U. GLOBAL STUD. L. REV. 303 (2009).
 - 2 See, e.g., Yasuhiko Tanabe, *ADR no Tsukaimichi* [Usage of ADR], 94 BIJINESU RO- JA-NURU 51 (2016) (Japan) (Private ADR numbers, outside of court-sponsored ADR, are “unsatisfactory”); Shuji Yanase, *The Standards of Judgment for Dispute Resolution in Financial ADR in Japan*, 26 COLUM. J. ASIAN L. 29, 43–44 (2013) (“ADR by private third parties has not been actively used, except in limited specific areas such as traffic accidents.”). In 2004, the Japanese legislature enacted the Act on Promotion of Use of Non-Judicial Dispute Resolution Procedures. Saiban-gai Funsōkaiatsu Tetsuzuki no Riyō no Sokushin ni kansuru Hōritsu [Act on Promotion of Use of Non-Judicial Dispute Resolution Procedures], Law No. 151 of 2004, Horei Teikyo De-ta Shisutemu [Horei DB], <http://law.e-gov.go.jp/htmldata/H16/H16HO151.html> (Japan). The Act obligates national and local governments to make efforts to provide information to the public to promote private ADR. See Chūsai ADR Tōkei Nenji Hō koku-Sho [Arbitration ADR Statistics Annual Report], *Nihon-Bengoshi-Rengokai ADR (Saibangai-Funso-Kaiatsu-Kikan) Senta* [JAPAN FED'N OF BAR ASS'NS ADR CENTER], available at http://www.nichibenren.or.jp/jfba_info/statistics/reform/adr_statistical_yearbook.html (Japan). The Act also allows private ADR service providers to obtain authorization as Certified Dispute Resolution Business Providers by the Ministry of Justice. Act on Promotion of Use of Non-Judicial Disputes Resolution Procedures, Law No. 151 of 2004, arts. 5-11 (Japan). Since the Act came into force in 2007, 149 institutions have earned official certifications. Chūsai ADR Tōkei Nenji Hō koku-Sho, *supra*. Despite these efforts, most of those newly certified ADR programs have seen little use. *Id.* Most notable of the private programs are the thirty-five ADR Centers established and maintained by district bar associations. *Id.* In fiscal year 2012, the bar associations' ADR centers saw 1,046 cases in total compared to 1,284 cases filed in the 125 certified ADR programs. Chusai ADR Tokei Nempo [Statistical Yearbook for Arbitration and ADR], NIHON-BENGOSHI-RENGOKAI ADR (SAIBANGAI-FUNSO-KAIKETSU-KIKAN) SENTA [Japan Fed'n of Bar Ass'ns], available at http://www.nichibenren.or.jp/library/ja/legal_aid/consultation/data/statistical_yearbook2012.pdf (Japan). Excluding the five certified bar association ADR centers and FINMAC, the

2009, Japan adopted a Financial ADR System (*Kin'yuu ADR Seido*) that has been widely used,³ and, according to some, successful.⁴

This article examines Japan's designated alternative dispute resolution centers now operating in various financial services industries, including the securities, banking, and insurance industries. It argues that the new Financial ADR System has expanded access to "legal" remedies for investors and other financial service customers. Research shows that Financial ADR generally provides a less expensive, faster remedy than is available from the courts: cases can be mediated for no or a reduced cost to the customer, and most are settled in less than six months.⁵ Research suggests that the settlements are modest—generally less than fifty percent of claimed damages—but they track the comparative fault awards found in civil litigation, and most cases are settled in accordance with legal norms developed by the courts.⁶ In some cases, they even expand the remedies available, and they offer claimants a means to avoid

other certified ADR centers processed only 496 cases in total. *Id.* See also *Ninsho-Funso-Kaiketsu-Jigyosha no Toriatsukai Kensu* [Number of Cases Handled by Certified Dispute Resolution Service Providers], MINISTRY OF JUSTICE, available at <http://www.moj.go.jp/KANBOU/ADR/images/kensu.pdf> (Japan). None of the private ADR programs in Japan come close to the volume of in-court conciliation. In 2012, there were 48,627 cases filed in the Japanese Summary Courts for civil in-court conciliation (*minji-choutei*). See *Saiko-Saibansho-Jimusokyoku* [General Secretariat of Supereme Court], *Shiho-Tokei-Nempo* (2) *Minji-Gyosei-Hen Heisei 24 Nendo* [Judicial Yearbook vol.2 Civil and Administrational Cases 2012] (Saiko-Saibansho, 2013) at 61, <http://www.courts.go.jp/app/files/toukei/632/006632.pdf> (Japan).

3 Act for the Amendment of the Financial Instruments and Exchange Act (Act No. 58 of 2009). See *Explanations of Laws and Regulations*, FSA NEWSLETTER No. 83 (Financial Services Agency, Japan), 2010, at 10, <http://www.fsa.go.jp/en/newsletter/2010/02.pdf>. The eight Designated Dispute Resolution Organizations (*shitei-funso-kaiketsu-kikan*) had 6,474 complaints filed against financial service providers and 1,626 mediation filings (*funso-kaiketsu-tetsuzuki*) in fiscal year 2012. See *Shitei-Funso-Kaiketsu-Kikan no Funso-Kaiketsu-tou-Gyomu Jisshi-Jokyo* [Implementation Status of Dispute Resolution and Complaint Handling by Designated Dispute Resolution Organizations], FIN. SERVS. AGENCY, http://www.fsa.go.jp/singi/singi_trouble/siryou/20130603/03.pdf (Japan).

4 The Financial Services Agency Council of Advisors (*yushikisha kaigi*) found heavy use of the Financial ADR System following the 2007-2008 financial crisis and positively evaluated the results. *Kin'yuu ADR Seidou no Arikatato ni Tsuite no Gaiyou* [Report of Experts on Financial ADR System], FIN. SERVS. AGENCY (March 8, 2013), <http://www.fsa.go.jp/singi/adr-followup/20130308/01.pdf> (Japan). They suggested it made a "significant contribution to customer protection." *Id.*

5 See *infra* Section V.C.

6 See *infra* Sections V.A, V.C.

severing necessary business relationships.⁷ There is a cost to all of this. Practitioners report that the Financial ADR System dispenses “rough justice,” settling cases quickly for less than if resolved through litigation.⁸ There is a consensus that Financial ADR is not for the big cases, and cases requiring specific findings of fact or novel questions of law.⁹ But, at the same time, they acknowledge that use of the Financial ADR system is voluntary, and cases are being mediated that would not be litigated.¹⁰

Whether the case is ultimately mediated or litigated, Japan’s Financial ADR has expanded the influence of the courts and lawyers.¹¹ Unlike earlier ADR mechanisms, Japan’s Financial ADR allows, and, in some cases, necessitates continued use of legal professionals and the court system.¹² Legal professionals, rather than Japan’s traditional bureaucracy, LDP, and business triumvirate, continue to play a significant role in shaping the legal norms in this area.¹³

7 See *infra* Section V.A.

8 E-mail from Osaka attorney to authors (Sept. 9, 2016) (on file with authors).

9 *Id.*; Yanase, *supra* note 2, at 79; Yutaka Tanaka, *Kin'yuu ADR no Hyouka to Kongo no Kadai*, *Zenkinkyo ADR 5 Shuunen Symposium [The 5th Anniversary Symposium by ADR]* 4 (Nov. 13, 2015), available at http://www.zenginkyo.or.jp/fileadmin/res/news/news280125_1.pdf (Japan).

10 See *infra* Section V.C.

11 See *infra* Part V. In contrast, scholars have documented earlier instances of Japanese elites limiting causes of action and access to courts through implementation of alternative dispute resolution mechanisms. See, e.g., UPHAM, *supra* note 1; Haley, *supra* note 1; Takao Tanase, *The Management of Disputes: Automobile Accident Compensation in Japan*, 24 LAW SOC'Y REV. 651, 655 (1990).

12 See *infra* Part V.

13 See *infra* Sections V.A-V.B. This is part of a larger trend of legal professionals assuming a larger role in Japanese society. See Ginsburg & Hoetker, *supra* note 1, at 49–50. Scholars have previously established an empirical link between the increase in lawyer population and civil litigation. *Id.* Scholars and practitioners have documented the rapid growth of corporate lawyering beginning in the late 1990s and following the judicial system reforms beginning in 2001. See, e.g., Bruce E. Aronson, *My Key Phrase for Understanding Japanese Law: Japan as a Normal Country ... with Context*, 22 MICH. ST. INT'L L. REV. 815, 823–24 (2014); Eric Sibbitt, *The New World of Corporate Lawyering in Japan*, 3 CHI.J. INT'L L. 503 (2002); Kota Fukui & Yusuke Fukui, *Empirical Support for Redefining the Legal Profession and New Roles for Lawyers in Japanese Corporations*, 12–2 AUSTL. J. ASIAN L. 273, 291 (2010). More recently, scholars have noted the important role lawyers have played in both litigation and alternative dispute resolution following the Fukushima Nuclear Disaster. See, e.g., Luke Nottage, *Resolving Claims from the Fukushima Nuclear Disaster* (Jan. 26, 2015), JAPANESE LAW AND THE ASIA-PACIFIC, http://blogs.usyd.edu.au/japaneselaw/2015/01/resolving_nuclear_claims.html.

Part II of this article offers some background to the introduction of Financial ADR in Japan. Part III describes its statutory framework. Financial ADR is authorized by a variety of statutes including the Financial Instruments and Exchange Act (FIEA), the successor to the Securities and Exchange Act, and it is offered through separate Designated Dispute Resolution Organizations (DDROs) affiliated with various self-regulatory organizations or industry associations. It also describes the basic process utilized by the DDROs. Part IV discusses two commonly used DDROs, the organization used by securities broker-dealers and that used by banks, and it describes their processes in more detail. Part V and the Conclusion offer thoughts on the role of Financial ADR and the rule of law in Japan.

II. THE ORIGINS OF JAPAN'S FINANCIAL ADR

A. *The Loss Compensation Scandal & Resulting Regulatory Scheme*

There are several defining events in the development of modern Japanese securities law. The first is, undoubtedly, Japan's adoption of the Securities and Exchange Act following the end of World War II.¹⁴ In 1948, Japan enacted a securities law based almost entirely on the United States' Securities Act of 1933 and Securities Exchange Act of 1934.¹⁵ For the next fifty years, Japan saw almost no securities litigation,¹⁶ and then came the collapse of Japan's economic bubble in 1990, and its Loss Compensation Scandal in 1991.¹⁷

In 1991, newspapers reported that major securities broker-dealers, including the now defunct Yamaichi Shouken, had compensated their most important clients for losses those clients suffered following the collapse of the bubble economy.¹⁸ Though

14 Mitsuru Misawa, *Securities Regulation in Japan*, 6 VAND. J. TRANSNAT'L L. 447, 447–48 (1972–1973).

15 See, e.g., Mitsuru Misawa, *A Comparative Study of US and Japanese Bank Directors' Duty of Disclosure*, 123 BANKING L.J. 39, 64 (2006).

16 See, e.g., Jonathan Macey & Hideki Kanda, *The Stock Exchange as a Firm: The Emergence of Close Substitutes For The New York and Tokyo Stock Exchanges*, 75 CORNELL L. REV. 1007, 1043 (1990); Mark D. West, *The Pricing of Shareholder Derivative Actions in Japan and the United States*, 88 NW. U. L. REV. 1436, 1437, 1445–46, 1493 (1994).

17 Andrew M. Pardieck, *The Formation and Transformation of Securities Law in Japan: From the Bubble to the Big Bang*, 19 UCLA PAC. BASIN L.J. 1, 3–9 (2001).

18 See Mitsuru Misawa, *Loss Compensation in the Japanese Securities Market: Causes, Significance, and Search for a Remedy*, 25 VAND. J. TRANSNAT'L L. 37, 37–41 (1992–1993). Most of these clients were corporate or institutional investors.

ex post compensation of an investor's loss had not been expressly prohibited, it clearly undermined trust in the securities markets.¹⁹ The scandal dominated the news and politics in the early 1990s, and it changed the law.²⁰

In 1991, the Japanese Diet amended the Securities and Exchange Act to impose a total ban on *ex post* loss compensation by securities broker-dealers to their customers.²¹ Though the Ministry of Finance drafted amendments were offered as a hasty response to the scandal,²² the statutory ban remains in effect today, with some

Id. at 38. Some obtained monetary compensation; most obtained assistance in delaying recognition of losses by shuffling bad assets among related corporate entities in order to avoid disclosure—a practice referred to as *tobashi*, or literally the “flying away” of bad assets. Bruce E. Aronson, *The Olympus Scandal and Corporate Governance Reform: Can Japan Find a Middle Ground Between the Board Monitoring Model and Management Model?*, 30 UCLA PAC. BASIN L.J. 93, 107 (2012). These practices led to the failure of Yamaichi Securities, then one of the largest securities companies in Japan, and ultimately triggered a financial crisis. *Id.* at 107–08. See also Mark D. West, *Information, Institutions, and Extortion in Japan and the United States: Making Sense of Sokaiya Racketeers*, 93 NW. U. L. REV. 767, 780 (1999); C. Jeffrey Char, *Reforming Japan's Securities Market - The Loss Compensation Scandal*, 10 INT'L TAX & BUS. L. 173, 173–75 (1993); Masahisa Ikeda, *The Legality of Compensating Investors in the Japanese Securities Market*, 33 HARV. INT'L L.J. 592, 594–95 (1992); Curtis J. Milhaupt, *Managing the Market: The Ministry of Finance and Securities Regulation in Japan*, 30 STAN. J. INT'L L. 423 (1994).

19 Misawa, *supra* note 15, at 47.

20 See, e.g., Ichiro Omi, *The Compensation Scandal in Japanese Finance*, 39 JAPAN Q. 45 (1992); Char, *supra* note 18, at 173–74.

21 Shouken Torihiki Hou nado no Ichibu wo Kaisei Suru Houritsu [Act for Partial Amendment of Securities and Exchange Act], Law No. 96 of 1991 (Japan).

22 *Shouken-Torihiki-Hou Oyobi Gaikoku-Shouken-Gyousha ni Kansuru Houritsu no Ichibu wo Kaisei-suru Houritsu-An: Kaigiroku Sakuin Joho* [Minutes Index for Bill to Amend Securities and Exchange Act and the Act Regarding Foreign Securities Dealers], NATIONAL DIET LIBRARY, <http://hourei.ndl.go.jp/SearchSys/viewShingi.do?i=112101004> (Japan). See also Char, *supra* note 18, at 173–74. The Ministry of Finance started working on the ban in July 1991, less than a month after the start of the scandal. ‘Houritsu de Sonshitsu Hoten Kinshi’ [Ban on loss compensation to be written in the statute], NIHON KEIZAI SHINBUN, morning ed., July 9, 1991, at 3 (Japan); Dai 120 Kai Kokkai Shugiin Okura Linkai Kaigiroku 19 Gou (Hashimoto Ryutaro Hatsugen) [Statement by Minister of Finance, Ryutaro Hashimoto, Hearing of Finance Committee, House of Representative, 120th Diet, Committee Minute No. 19] available at <http://kokkai.ndl.go.jp/SENTAKU/syugiin/120/0140/12007250140019.pdf>, at 6 (Japan). The loss compensation amendment drafted by the Ministry of Finance was proposed to the 121st Diet convened in August and passed by the Diet in early October. Its proponents in the Diet considered the ban on loss compensation an emergency or stop-gap measure. See, e.g., Dai 121

notable exceptions discussed below.²³

A second legacy of the collapse of the bubble economy and the loss compensation scandal was Japan's regulatory focus on loss compensation.²⁴ In 1992, Japan established its Securities and Exchange Surveillance Commission.²⁵ In its first eight years of operation, it issued more recommendations for charges (*kokuhatsu*) against regulated individuals relating to loss compensation than any other category, except insider trading.²⁶ During the same period, it issued more recommendations for remedial action (*kankoku*) by regulated individuals relating to loss compensation than any other category, except discretionary trading contracts.²⁷ The loss compensation scandal and related cases documenting fraudulent manipulation of corporate balance books pushed these issues to the top of the priorities list for Japan's securities regulators.²⁸

Kai Kokkai Shugiin Honkaigi 10 Gou (Tsujii Daiichi Hatsugen) [Statement by Representative Daiichi Tsuji, full House of Representatives, 121th Diet, Minute No. 10] available at <http://kokkai.ndl.go.jp/SENTAKU/syugiin/121/0001/12109200001010.pdf>, at 9 (Japan). This legislative process was abbreviated and expedited compared to the standard lengthy amendment process incorporating the work of various advisory bodies. See Frank J. Schwartz, *ADVICE AND CONSENT: THE POLITICS OF CONSULTATION IN JAPAN*, (1998), Ch. 2. Both legal and economic scholars are critical of the amendment, as drafted and implemented. See Char, *supra* note 18, at 173–74.

23 Kin'yuu Shouhin Torihiki Hou [Financial Instruments and Exchange Act], Law No. 25 of 1948 [hereinafter FIEA], art. 39 (Japan).

24 See Pardieck, *supra* note 17, at 7–24.

25 The Securities and Exchange Surveillance Commission was established as an independent council affiliated with the Ministry of Finance by the 1992 Shouken Torihikitou no Kousei wo Kakuho Suru Tame no Shouken Torihiki Hou tou no Ichibu wo Kaisei Suru Houritsu [Act for Partial Ammendment of Securities and Exchange Act to Ensure Fairness of Securities Transactions], Law No. 73 of 1992 (Japan).

26 Pardieck, *supra* note 17, at 13 chart 1. *Kokuhatsu* is translated in this source as “accusation.” *Id.* at 11.

27 *Id.* at Chart 2. *Kankoku* is translated in this source as “recommendation.” *Id.* at 11.

28 Pardieck, *supra* note 17, at 8–9, 13. More recently, the SESC has shifted its focus from loss compensation to other issues like disclosure violations and market manipulation. See Securities and Exchange Comission Nenji-Kohyo [Annual Reports Index] <http://www.fsa.go.jp/sesc/reports/reports.htm> (Japan); Securities and Exchange Surveillance Comission Fuzoku-Shiryo [SESC Appendix], http://www.fsa.go.jp/sesc/reports/n_11/20000928-1b.pdf (Japan). In 1999, the SESC reported fifteen violations of the loss compensation (*sonshitsu-hoten*) statute out of a total of 109 reported violations resulting in recommendations for administrative action (*kankoku*) against thirty-seven separate individuals and entities. *Id.* In 2000, that number dropped to three

B. The Judiciary's Role in Resolving Financial Service Disputes

These two legacies explain, in part, a third legacy of the collapse of the bubble economy and Japan's financial markets: judicial activism in the development of Japanese securities law.²⁹ The statutory ban on loss compensation, in effect, shut the door to out-of-court settlement between broker-dealers and investors alleging misrepresentation or fraud. It pushed their disputes into the courts.

The statutory ban on *ex post* loss compensation, little changed since its adoption in 1991, prohibits individuals from: "making an offer or promise . . . in connection with a Purchase or Sale or Other Transaction of Securities . . . to a customer . . . that the Financial Services Provider . . . will cover the whole or part of a loss that the customer incurred in connection with the relevant Securities . . ."³⁰

There are limited exceptions. The statute itself offers the first: loss compensation is permitted if the offer or promise of an economic benefit is made to compensate for a wrong confirmed by a regulatory authority. Loss compensation is permitted:

to cover the whole or part of a loss incurred due to problematic conduct (meaning illegal or wrongful conduct by a Financial Services Provider which is specified by Cabinet Office Ordinance as a potential cause of a dispute . . .) . . . if the Financial Services Provider, etc., receives confirmation from the Prime Minister in advance that the loss to be covered was incurred due to problematic conduct . . .³¹

A person seeking this confirmation must submit a written application to the designee of the Prime Minister's Office, in this case the Director-General of the Local Finance Bureau.³² It must state, *inter alia*, the facts of the dispute, the parties involved, the reason for the loss giving rise to the dispute, and the amount of compensation to be provided, along with supporting documentation.³³ To find

reported loss compensation violations, and in 2001, two. *Id.* From 2002 to present, the SESC has issued one or no recommendations relating to loss compensation each year. *Id.*

29 See, e.g., Pardieck, *supra* note 17.

30 FIEA, *supra* note 23, at art. 39(1)(ii) (Japan).

31 *Id.* at art. 39(3).

32 *Id.* at art. 39(5).

33 *Id.*; Kin'yuu Shouhin Torihikigyou to ni Kansuru Naikakufurei [Cabinet

“problematic conduct,” regulators typically require admission of a tortious act or illegal conduct and a causal link between that conduct and the losses.³⁴ In short, the statute allows privately negotiated loss compensation only in cases where the loss is incurred due to legally defined “problematic conduct” with causation confirmed by the government.³⁵

The introduction of the loss compensation prohibition in 1991 and its accompanying requirement of regulatory approval of any settlement of claims created strong incentives for broker-dealers not to settle with investors out of court. While it was common practice for attorneys to negotiate out-of-court settlements for their client’s losses with broker-dealers throughout the 1980s, the 1991 loss compensation ban changed things.³⁶

Ordinance on the Financial Instruments and Exchange Industries], Cabinet Office Ordinance No. 52 of 2007, art. 121 (Japan).

- 34 See KIN’YUU SHOUHIN TORIHIKIHO [KEEPING FINANCIAL INSTRUMENT TRANSACTIONS] 457–458 (Masayuki Kawamura ed., 2014) (Japan).
- 35 FIEA, *supra* note 23, at arts. 39(3), 194-7 (Japan). Kin’yuu Shouhin Torihiki Hou Shikourei [Order for Enforcement of the Financial Instruments and Exchange Act], art. 42(6) (Japan). While the statute contemplates confirmation by the Prime Minister, the Prime Minister delegates authority to the Commissioner of Financial Services Agency, who in turn delegates confirmation authority to the Director-General of each Local Financial Bureau (local branch offices of the Ministry of Finance). FIEA, *supra* note 23, at arts. 194-7(1), (6) (Japan).
- 36 See, e.g., Yoshio Takei, *Shōken-Higai-Kyūsai no Tameno Nyūmon-Kōza* [Introduction to Remedies for Loss from Securities Transactions] (Apr. 14, 2000), available at <http://www2.osk.3web.ne.jp/~syouken/himeji/siryou/nyuumonFrameSet1.html> (Japan). There are no statistics available regarding the number of out-of-court settlements that received administrative confirmation. E-mail from Financial Services Agency Official to authors (Oct. 18, 2016) (on file with authors). An official with the Financial Supervisory Agency indicated that neither the agency nor the local financial bureaus compiled such statistics. *Id.* They have also declined freedom of information requests regarding the confirmation process, including reports of “problematic conduct,” on the grounds that these documents contain trade secrets relating to broker-dealer compliance and oversight. *Id.* There is, however, substantial evidence of a decline in negotiated settlements following the ban. Takei, *supra*. A leading member of the plaintiffs’ bar has written about the changes following the loss compensation scandal. *Id.* He reports litigated, as opposed to negotiated settlements, became much more common and remembers few cases where he negotiated settlements after 1991. *Id.* The statutory ban’s effect on out-of-court negotiation caused the Japanese Federation of Bar Associations (JFBA) to advise against expansion of the ban on loss compensation to futures trades. JAPAN FEDERATION OF BAR ASS’NS, SHŌKEN TORIHIKI-HŌ NO ICHIBU O KAISEI SURU HŌRITSU-AN (KIN’YŪSHŌHIN TORIHIKI-HŌ (IWAYURU TŌSHI SĀ BISU-HŌ)-AN) NO SHŪSEI O MOTOMERU IKEN-

III. THE INSTITUTIONALIZATION OF FINANCIAL ADR

The cabinet ordinance implementing the initial statutory ban on loss compensation recognized additional exceptions to the ban where the loss compensation occurred pursuant to a court decision, court-sponsored mediation, or civil conciliation.³⁷ Starting in 2007, following the adoption of the Financial Instruments Exchange Act (FIEA), the ordinance recognized loss compensation without administrative approval if, *inter alia*, it occurred pursuant to mediation by a self-regulatory organization designated as an investor protection organization.³⁸

In 2008, the Japanese legislature expanded this exception when it introduced the Financial ADR System.³⁹ It incorporated

SHO [A DRAFT LAW THAT REVISES A PART OF THE SECURITIES AND EXCHANGE LAW (THE FINANCIAL INSTRUMENTS AND EXCHANGE LAW) (SO-CALLED INVESTMENT SERVER OPINION SEEKING TO REVISE THE DRAFT)] (Mar. 24, 2006), available at <http://www.nichibenren.or.jp/library/ja/opinion/report/data/060324.pdf> (Japan). Even after enactment of FIEA, which expanded the ban to futures trades, JFBA asserted that the ordinance should exempt settlements negotiated by lawyers in futures trade disputes from administrative approval. JAPAN FEDERATION OF BAR ASS'NS, KAISEI SHŌ HINTORIHIKIJOHŌ DAI 214-JŌ NO 2 (SONSHITSU HOTEN-TŌ NO KINSHI) NO SHŌREI NI KANSURU IKEN-SHO [OPINION ON MINISTERIAL ORDINANCE OF ARTICLE 214-2 OF THE REVISED COMMODITY ACT], (Feb 16, 2007), available at http://www.nichibenren.or.jp/library/ja/opinion/report/data/070216_4.pdf (Japan).

³⁷ Cabinet Office Ordinance No. 52 of 2007, art. 119(1)(3) (Japan).

³⁸ *Id.* at art. 119(4). See also FIN. SERVS. AGENCY, *Gyōsha ga junshu subeki kōi kisei no seibi* [Improvement of Action Regulations that Vendors Should Observe], in KIN'YUU SHOUHIN TORIHIKI HOUSEI NO SEIREI NAIKAKUFUREI-TŌ NO GAIYŌ [CABINET OFFICE ORDINANCE OF FINANCIAL INSTRUMENTS AND EXCHANGE LAW] 7 (July 31, 2007), available at <http://www.fsa.go.jp/news/19/syouken/20070731-7/34.pdf> (Japan). The ordinance established exceptions to the loss compensation ban where a dispute was mediated by a bar association arbitration center; the National Consumer Affairs Center of Japan, or a regional public-purpose organization; a certified dispute resolution center; or a mediation involving ten million yen or less among parties represented by attorneys or 1.4 million yen or less involving judicial scriveners. Cabinet Office Ordinance No. 52 of 2007, art. 119(5-7) (Japan). Where the parties are represented by attorneys or judicial scriveners, the legal professional representing the investor must provide a document to the financial service provider certifying that they have investigated the claim and confirmed the loss was due to problematic conduct of the type recognized by regulators pursuant to Article 39(3) and the relevant SRO has investigated the incident and confirmed the same. *Id.*

³⁹ FIN. SERVS. AGENCY, *Kin'yuu Shouhin Torihiki Hou To no Ichibu wo Kaisei Suru Houritsu: Seirei, Naikakufurei no Pointo* [Law to Amend Part of the Financial Instruments

Financial ADR into the FIEA, as well as sixteen other statutes regulating financial services.⁴⁰ While the Financial Service Agency and financial service providers had discussed these amendments for years, the disruption caused by the financial crisis of 2007-2008 passed this legislation.⁴¹

The amendments relating to the Financial ADR System created the legal framework for the DDROs.⁴² The amendments took effect in October of 2010, just in time for these new organizations to address a wave of disputes arising from thousands of foreign exchange derivatives contracts gone bad following the 2007-2008 global financial crisis.⁴³ In a short period of time, the number of successfully mediated disputes relating to financial services roughly

and Exchange Law: Points of Cabinet Office Ordinance], available at <http://www.fsa.go.jp/news/21/20091222-4/01.pdf> (Japan). The legislature introduced the Financial ADR System as part of a comprehensive package of amendments to the FIEA. *Id.* See also Yanase, *supra* note 2, at 34. They sought to improve the fairness and transparency of the financial markets by introducing new regulations governing credit-rating agencies; strengthen investor protection by introducing a “Financial ADR System” and new derivatives regulations; and increase the fairness and use of the markets by integrating commodities and securities markets and strengthening disclosure regulations. FIN. SERVS. AGENCY, *Kin'yuu Shouhin Torihiki Hou To no Ichibu wo Kaisei Suru Houritsu: Seirei, Naikakufurei no Pointo [Law to Amend Part of the Financial Instruments and Exchange Law: Points of Cabinet Office Ordinance]*, available at <http://www.fsa.go.jp/news/21/20091222-4/01.pdf> (Japan). See also Yanase, *supra* note 2, at 34.

40 See, e.g., FIN. SERVS. AGENCY, *Soumu Kikakukyoku, Shitei Funso Kaiketsu Kikan Muke no Sougou teki na Kantoku Shishin 2-3* (June 2016), <http://www.fsa.go.jp/common/law/guide/kinyuadr.pdf> (Japan); Yanase, *supra* note 2, at 74.

41 Dai 171-kai kokkai zaimu kin'yū iinkai dai 14-gō (Yasano Kaoru Hatsugen) [The 171st Nat'l Assembly Fin. And Fin. Comm. No. 14], No. 14 of 2009, Assemb. 171 (Shūgiin) http://www.shugiin.go.jp/internet/itdb_kraigokua.nsf/html/kaigrokua/009517120090408014.htm (Japan). Changes to other financial dispute resolution systems resulting from the global financial crisis are discussed in Shahla F. Ali, *Globalization and Financial Dispute Resolution: Examining Areas of Convergence and Informed Divergence in Financial ADR*, 2013 J. DISP. RES. (2) 331, 336, 336–46 (2013).

42 See Act for the Amendment of the Financial Instruments and Exchange Act, *supra* note 3.

43 See *Chuushou-Kigyo-Muke Kawase-Deribatibu Torihiki-Jokyo (Bei-Doru/Yen) ni kansuru Chosa no Kekka ni Tsuite (Sokuho-Chi)* [Interim Report of the Result of Investigation over Foreign exchange derivatives (U.S. Dollar/Yen) Sold to Small-to-Medium Businesses] <http://www.fsa.go.jp/news/22/ginkou/20110311-2.html> (Japan). For more background on the global financial crisis, see Dick K. Nanto, *The Global Financial Crisis: Analysis and Policy Implications* 74 (Congressional Research Service Oct. 2, 2009), <https://www.fas.org/sgp/crs/misc/RL34742.pdf>.

quadrupled.⁴⁴ Japan went from resolving 224 financial disputes through ADR sponsored by self-regulatory organizations in 2006,⁴⁵ to settling 944 such suits through its DDROs in 2012.⁴⁶

In the next section, we describe the legal framework for DDROs and the basic procedures they follow.

A. The Institutional Framework for Japan's Financial DDROs

While legislated as amendments to seventeen statutes each covering a specific category of financial services, e.g., banking, insurance, securities, and consumer finance,⁴⁷ the mandate remains the same: “to ensure sound and proper business operations, to develop fair pricing by ensuring fair dealing, to maintain the public’s trust of financial service industry, and to provide customer protection.”⁴⁸

A DDRO is established following the process outlined by the

44 *Gyoukai Dantai no Funsou Kaiketsu Shi'en Tetsuzuki no Riyou Sokushin Nitsuite* [Promotion of the Use of Dispute Resolution Assistance Provided by Self-Regulatory Organizations] 3, http://www.fsa.go.jp/singi/singi_trouble/siryou/20071207/05.pdf(Japan); *Shitei Funsou Kaiketsu Kikan no Funshou Tetsuzuki Jisshi Joukyou* (Heisei 24 nen 4 gatsu 1 nichi – Heisei 25 nen 3 gatsu 31 nichi) [Status Report of Dispute Resolution through DDROs (April 1, 2012 – March 31, 2013)], http://www.fsa.go.jp/singi/singi_trouble/siryou/20130603/02.pdf (Japan).

45 *Gyoukai Dantai no Funsou Kaiketsu Shi'en Tetsuzuki no Riyou Sokushin Nitsuite*, *supra* note 44.

46 *Shitei Funsou Kaiketsu Kikan no Funshou Tetsuzuki Jisshi Joukyou*, *supra* note 44.

47 See FSA, *supra* note 40, at 2–3. Representative examples can also be found in *Kin'yuu Kikan to no aida Toraburu wo Kakaeteiru Riyousha no Minasama he 4* [For Everyone who has Difficulty with Financial Institutions] (Nov. 2016), http://www.fsa.go.jp/policy/adr/adr_pamphlet.pdf (Japan). They include, e.g., the Money Lending Business Act; Banking Act; Financial Instruments and Exchange Act; Business Trusts Act; and Insurance Business Act. *Id.* The FSA initially contemplated establishment of a single, unified dispute resolution organization but opted for industry-specific entities given the complexity and unique characteristics of the individual financial services and the ability to incorporate and take advantage of already existing industry-specific ADR frameworks. See *Kin'yuu ADR Seido Ni Tsuite (Sankou)* [About Fin. ADR Sys.], <http://www.fsa.go.jp/policy/adr/adr.pdf> (Japan). See also Yanase, *supra* note 2, at 74.

48 Yanase, *supra* note 2, at 85. Yanase paraphrased statements of purpose found in the various financial services statutes. *Id.* at 75. For the original wording, see FIEA, *supra* note 23, at art. 1 (Japan); Ginkou Hou [Banking Act], Law No. 59 of 1981 [hereinafter Banking Act], art. 1 (Japan); Kashikingyou Hou [Money Lending Business Act], Law No. 32 of 1983, art. 1 (Japan); Hokengyou Hou [Insurance Business Act], Law No. 105 of 1995, art. 1 (Japan); Shintakugyou Hou [Trust Business Act], Law No. 154 of 2004, art. 1 (Japan).

statute and ordinance governing a particular industry, e.g., the FIEA and related ordinances governing securities transactions.⁴⁹ They generally require organization by competent persons; a showing that the organization has a sufficient financial and technical basis to provide dispute resolution services; and that the composition of its officers and employees are “unlikely to compromise the fair implementation of dispute resolution services.”⁵⁰ The DDRO is required to draft “operational rules” (*gyoumu kitei*) for implementing the dispute resolution services, make them available for public comment, and receive a vote by the affected businesses with less than one-third of the votes registering opposition to the proposed rules.⁵¹ The Prime Minister’s Office must authorize the operational rules.⁵²

The DDRO must be organized as a registered juridical person (*houjin*) that may or may not be incorporated,⁵³ but which is designated by the Commissioner of the Financial Services Agency⁵⁴

49 See, e.g., FIEA, *supra* note 23, at art. 156-39(1) (Japan). While, the insurance DDROs process the most claims, Japan’s primary financial services regulator, the Financial Services Agency, references the FIEA and related ordinances when discussing financial ADR issues applicable to all financial service industries. See FSA, *Shishin*, *supra* note 40 (Japan); FSA, *Shitei Funsou Kaiketsu Kikan no Funsou Kaiketsu tou Gyoumu Jisshi Joukyou*, 1-2 (June 9, 2016), http://www.fsa.go.jp/singi/singi_trouble/siryou/20160609/03.pdf (Japan). Similarly, the discussion here is based, primarily, on the FIEA.

50 FIEA, *supra* note 23, at art. 156-39(1)(v)-(vi) (Japan). See also Banking Act, *supra* note 48, at art. 52-62(1) (Japan). The FIEA defines competency by precluding individuals from serving as an officer of the DDRO if they have been designated an adult ward, adjudicated bankrupt, subject to a term of imprisonment, an officer of a corporation whose designation was rescinded under the FIEA, or a person sentenced to a fine for violating the FIEA. *Id.*

51 FIEA, *supra* note 23, at art. 156-39(1)(vii)-(viii) (Japan). The operational rules must include, *inter alia*, matters relevant to the contents of the Basic Contract, the conclusion of the Basic Contract, and the implementation of Dispute Resolution Services; matters relevant to dues and fees paid; and matters relevant to the processing of complaints. *Id.* at art. 156-44(1)(i)-(viii); Banking Act, *supra* note 48, at art. 52-67 (Japan).

52 FIEA, *supra* note 23, at art. 156-44(8) (Japan). Banking Act, *supra* note 48, at art. 52-67(8) (Japan).

53 FIEA, *supra* note 23, at arts. 156-39(1)(i), 156-40(2) (Japan). Banking Act, *supra* note 48, at arts. 52-63(2), 52-67(1) (Japan).

54 The Financial Services Agency is an administrative agency affiliated with the Prime Minister’s Cabinet Office responsible for the inspection and supervision of financial institutions. FIN. SERVICES AGENCY, THE ROLE OF THE FSA, 3 (2017), <http://www.fsa.go.jp/en/about/pamphlet.pdf>. It is considered an “external organ” of the office, and a successor agency to the Ministry of Finance and Financial Reconstruction Commission. *Id.* at 2. The Securities

to process complaints and provide dispute resolution services for certain types of financial services.⁵⁵ If a DDRO is available for a particular category of financial services, a financial service provider must conclude a Basic Contract for the Implementation of Dispute Resolution with the DDRO.⁵⁶ The authorizing statute, e.g. the FIEA, enumerates the mandatory contents for this Basic Contract,⁵⁷ and the Basic Contract must be included in the DDRO's Operational Rules.⁵⁸

The Basic Contract must include terms establishing:

- (i) That a financial services provider or its customer may commence "Complaint Processing Procedures" or "Dispute Resolution Procedures" on application;⁵⁹
- (ii) That the DDRO or mediator "may request the member person or firm comply with DDRO procedures and the provider "must not refuse such request without just cause for doing so."⁶⁰
- (iii) That the DDRO or mediator "may request" the member person or firm "make a report

and Exchange Surveillance Commission along with the Certified Public Accounts and Auditing Oversight Board currently operate as independent entities external to, but affiliated with, the Financial Services Agency. *Id.* at 13–14.

55 FIEA, *supra* note 23, at arts. 156-38 to 156-40, 194-7 (Japan).

56 *Id.* at arts. 37-7(1), 156-38(13). Banking Act, *supra* note 48, at art. 52-79 (Japan). If a DDRO is not available for a service category, a financial service provider must make arrangements for resolving disputes through Certified Dispute Resolution Procedures and its equivalents specified in the Ordinance, which usually means a mediation program exempted from administrative certification of settlement. *Kin'yu-Shohin-Torihiki-Gyo ni Kansuru Naikakufu-Rei* [Cabinet Office Ordinance on Financial Instruments Business] Cabinet Office Ordinance No. 52 of 2007, art. 115-2 (2), <http://law.e-gov.go.jp/htmldata/H19/H19F10001000052.html> (Japan).

57 FIEA, *supra* note 23, at art. 156-44(1)(i) (Japan).

58 *Id.*

59 *Id.* at art. 156-44(2); Banking Act, *supra* note 48, at art. 52-67(2) (Japan).

60 FIEA, *supra* note 23, at art. 156-44(2)(ii) (Japan). FINMAC's rules state that business operators "must cooperate" with the mediators and FINMAC staff in promoting resolution of the complaint. FINMAC, OPERATIONAL RULES CONCERNING COMPLAINT RESOLUTION SUPPORT AND MEDIATION art. 7 para. 1, https://www.finmac.or.jp/english/pdf/en_05.pdf [hereinafter FINMAC, OPERATIONAL RULES] (official English translation).

or . . . submit books and documents or any other articles” in the dispute resolution proceedings;⁶¹

(iv) That the mediator may prepare a “settlement proposal” for resolving a dispute and “recommend” settlement,⁶² and if there is no prospect for reaching an agreed settlement the mediator may present a “Special Conciliation Proposal” along with justifications for the same.⁶³

Statutory requirements for the Basic Contract go on to include requirements that the member inform the DDRO of related litigation and take specified steps to ensure its customers are aware of the dispute resolution service.⁶⁴

The DDRO is also intended to serve a compliance role, with authority to investigate business members’ performance of obligations assumed in negotiated settlements and to make recommendations regarding the same.⁶⁵ Breach of this Basic Contract without cause, following a hearing by the DDRO, may result in the publication of the breach to the public and reporting the same to the Prime Minister’s Office.⁶⁶ The DDRO may expel a member business for noncompliance with the contract.⁶⁷

61 FIEA, *supra* note 23, at art. 156-44(2)(iii) (Japan). FINMAC rules state the basic contract or “Master Agreement” must include matters relating to member businesses’ obligation to “give explanation and submit materials, accounting books, or other goods” and its “obligation to submit a written answer” to the complaint. FINMAC, *OPERATIONAL RULES*, *supra* note 60, at art. 5-3(3).

62 FIEA, *supra* note 23, at art. 156-44(2)(iv) (Japan).

63 *Id.* at art. 156-44(2)(v).

64 *Id.* at art. 156-44(2)(vii-x). Banking Act, *supra* note 48, at art. 52-67(1)(vii)-(ix) (Japan). FSA, *Shishin*, *supra* note 40, at 14 (Japan).

65 FIEA prescribes disclosure to the public and report to the Prime Minister as sanctions for breach of the Basic Contracts by member service providers. FIEA, *supra* note 23, at art. 156-45 (1) (Japan). The DDRO establishes its authority to investigate service providers’ compliance with the settled outcomes in its Basic Contracts. See, e.g., FINMAC, *OPERATIONAL RULES*, *supra* note 60, at arts. 5-3(8), 57; Nipponginkō kyōkai [Japanese Bankers Ass’n], Gyōmu Kitei [Operational Rules] [hereinafter JBA, Operational Rules] art. 46, attachment art. 8, <http://www.zenginkyo.or.jp/abstract/adr/adrabout/rules/>.

66 FIEA, *supra* note 23, at art. 156-45(1) (Japan).

67 *Id.* at art. 156-56. Kin’yuu Shouhin Torihikigyou to ni Kansuru Naikakufurei [Cabinet Ordinance on the Financial Instruments and Exchange Industries], Cabinet Office Ordinance No. 52 of 2007, art. 14(1) (Japan). Sankou, *supra* note 47. FINMAC’s Articles of Association provide additional grounds

The DDRO may be part of an industry association, self-regulatory organization or an independent corporate entity.⁶⁸ The legal framework leaves it to the representative industry association to set up the DDRO program for its members.⁶⁹ As discussed in more detail in Part IV, the investment industry chose to incorporate a separate, independent non-profit juridical entity to provide dispute resolution services.⁷⁰ In contrast, other industry associations, including the Japanese Bankers Association, set up internal DDRO programs.⁷¹

After establishment, DDROs are required to report each year to the Financial Services Agency, and the Agency is to hold a hearing regarding DDRO activities.⁷² The Financial Services Agency has established a “Financial ADR Office” to oversee the DDROs.⁷³ The Financial ADR Office is authorized to request information, conduct on-site inspections, and hold hearings as necessary to respond to disclosed problems.⁷⁴ Following investigation, the Financial Supervisory Agency may sanction a DDRO by requiring the DDRO to submit a report analyzing the issue and detailing remedial steps.⁷⁵ On a finding of material problems impacting the DDRO’s ability to function fairly and appropriately, the Financial ADR Office may also

for expulsion. FINMAC, ARTICLES OF ASSOCIATION OF FINANCIAL INSTRUMENTS MEDIATION ASSISTANCE CENTER, art. 12, https://www.finmac.or.jp/english/pdf/en_04.pdf [hereinafter FINMAC, ARTICLES OF ASSOCIATION] (official English translation). Those grounds include violation of its rules; an action that “breaches laws and regulations, or is offensive to public order and morals”; or actions that damage the reputation of FINMAC or run counter to the purport of its establishment. *Id.*

68 FSA, *supra* note 40, at 13. FIEA, *supra* note 23, at art. 156-39(1) (Japan). Banking Act, *supra* note 48, at art. 52-62(1) (Japan).

69 See *Dai 171 Kai Kokkai Zaimu Iinkai Dai 15 Go 2 Pe-ji (Jun’ichi Naitou Hatsugen)* [Statement by Deputy Director General of Finance Bureau, Jun’ichi Naitou, Hearing of the Committee for Financial Affairs, House of Representatives, 171st Diet, Committee Minutes Vol. 15] at 2, <http://kokkai.ndl.go.jp/SENTAKU/syugiin/171/0095/17104140095015.pdf> (Japan).

70 FINMAC, ARTICLES OF ASSOCIATION, *supra* note 67, at arts. 4, 5.

71 See, e.g., JBA, *Kujou Shori Tetsuduki oyobi Funsou Kaiketsu Tetsuduki no Gyoumu Kitei* [Complaint Resolution Procedures and Dispute Resolution Procedures Operating Rules] [hereinafter JBA, Operating Rules] art. 1, available at <http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/rules/rule02.pdf> (Japan).

72 FIEA, *supra* note 23, at art. 156-57 (Japan). Cabinet Office Ordinance No. 52 of 2007, art. 15 (Japan).

73 FSA, *Kin’yu Toraburu Kaiketsu Seido Suishin Shitsu*, <http://www.fsa.go.jp/policy/adr/> (Japan).

74 FIEA, *supra* note 23, at art. 156-58(1) (Japan). FSA, *supra* note 40, at 3, 9.

75 FIEA, *supra* note 23, at art. 156-58(1) (Japan). FSA, *supra* note 40, at 8.

issue a disciplinary order (*choukai meirei*),⁷⁶ business improvement order (*gyoumu kaizen meirei*), business suspension order (*gyoumu teishi meirei*),⁷⁷ designation rescission order (*shitei no torikeshi*),⁷⁸ and other administration sanctions, including more regular reporting requirements.⁷⁹

B. A Procedural Outline—DDRO Complaint & Mediation Processing

The statutory framework for DDROs contemplates that they provide both complaint processing and mediation services.⁸⁰ It also identifies consultation services as an ancillary operation of the DDRO.⁸¹ The Financial Services Agency notes that neither complaint nor dispute are defined terms and they should be interpreted broadly.⁸² Through fact investigation and advice to users, the DDRO is required to encourage independent resolution by the parties of the dispute.⁸³

The DDROs may encourage but not force extra-judicial resolution of the dispute—there is no mandatory or binding ADR.⁸⁴ The parties may choose to file suit in a civil court, to petition the DDRO to process a complaint; or initiate mediation, and they may

76 FSA, *supra* note 40, at 8–9.

77 FIEA, *supra* note 23, at art. 156-5(1) (Japan).

78 *Id.* at art. 156-60(1) & 156-61.

79 *Id.* at art. 156-58(1) & 156-61(1) & 156-59(1)-(2). FSA, *supra* note 40, at 8–9.

80 FIEA, *supra* note 23, at art. 156-38(10-11) & 156-44(2)(ii) (Japan). This structure is similar to the complaint, mediation, and adjudication structure found in financial ombudsman systems in other jurisdictions. Shahla F. Ali & Antonio Da Roza, *Alternative Dispute Resolution Design in Financial Markets—Some More Equal Than Others: Hong Kong’s Proposed Financial Dispute Resolution Center in the Context of the Experience in the United Kingdom, United States, and Australia*, 21 PAC. RIM L. & POL’Y J. 485, 510 (2012); Shahla F. Ali, *CONSUMER FINACIAL DISPUTE RESOLUTION IN A COMPARATIVE CONTEXT: PRINCIPLES, SYSTEMS, AND PRACTICE* (2013).

81 FIEA, *supra* note 23, at art. 156-38(11) (Japan). FSA, *supra* note 40, at 18. FIEA, *supra* note 23, at art. 156-45(2) (Japan) also states that a DDRO must endeavor to provide information, consultation, and other support to member persons, firms, or other persons to preemptively prevent complaints.

82 FSA, *supra* note 40, at 18.

83 FIEA, *supra* note 23, at art. 156-49 (Japan). FSA, *supra* note 40, at 19.

84 Yanase, *supra* note 2, at 37, 82. See, e.g., Tsuneo Osawa, *ADR o Gaikansuru, BIJINESU RO- JA-NURU* [BUSINESS LAW JOURNAL], (2010), at 27 (Japan). See also Yasuhiko Tanabe, *ADR no Tsukaimichi* [Usage of ADR], 94 BIJINESU RO- JA-NURU 51, 54 (2016) (Japan).

choose to terminate the DDRO proceedings at any time.⁸⁵

Complaints may be initiated via phone, mail, or in person and are processed without charge.⁸⁶ When a customer files an application for resolution of a complaint, the DDRO is directed to provide the customer with “necessary advice” and “investigate the underlying facts,” as well as notify the member firm of the complaint and request that it be “expeditiously” processed.⁸⁷

The DDRO may request the business provide an explanation, verbally or in writing, and submit materials relating to the complaint, and the business may not object without “justifiable reason.”⁸⁸ The DDRO may ask for the “opinion” of the business, instruct it to enter into a “face-to-face negotiation” with the customer and report the result of the negotiation.⁸⁹ Following notification by the DDRO, the business is to “immediately” contact the claimant, “respond . . . in good faith,” and “make efforts to resolve” the complaint.⁹⁰ The DDRO is obligated to prepare a record of complaints processed, including the date of filing, the names of the customer and business member, the chronology of the complaint resolution process and its result.⁹¹

If the complaint cannot be resolved through negotiation,

85 See, e.g., FINMAC, DETAILED RULES RELATING TO THE OPERATIONAL RULES CONCERNING COMPLAINT RESOLUTION, annex 8, https://www.finmac.or.jp/english/pdf/en_06.pdf [hereinafter FINMAC, DETAILED RULES] (official English translation).

86 Kin'yuu Shouhin Torihikigyou to ni Kansuru Naikakufurei [Cabinet Ordinance on the Financial Instruments and Exchange Industries], Cabinet Office Ordinance No. 52 of 2007, art. 12(2) (Japan). See, e.g., FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 12; JBA, Operating Rules, *supra* note 71, at art. 28. See also Kujou Shori no Tetsuzuki oyobi Funsou Kaiketsu Tetsuzuki to no Jisshi ni Kansuru Gyomu Kitei [Business Regulations Concerning the Implementation of Complaint Handling Procedures and Dispute Settlement Procedures], JAPANESE BANKERS ASS'N, art. 7, <http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/rules/rule01.pdf>.

87 FIEA, *supra* note 23, at art. 156-55(4)(v-vi), 156-49 (Japan). FINMAC, OPERATIONAL RULES, *supra* note 60, at 13. See also JBA, Operating Rules, *supra* note 71, at art. 9-10.

88 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 14. See also JBA, Operating Rules, *supra* note 71, at art. 9-10.

89 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 13 para. 2. See also JBA, Operating Rules, *supra* note 71, at art. 9-10.

90 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 13 para. 3. See also JBA, Operating Rules, *supra* note 71, at art. 9.

91 Cabinet Office Ordinance No. 52 of 2007, art. 10(1) (Japan). See also FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 45.

usually within two months, the DDRO interviews the claimant, explains the mediation process and determines whether the claimant wishes to proceed with mediation.⁹² The DDRO then provides to the applicant documentation describing fees to be paid by the applicant as well as the procedural rules from commencement to termination.⁹³

Mediation at a DDRO begins with a party filing a “motion for mediation.”⁹⁴ DDRO staff send the respondent(s) notification of the filing and requests to submit the documents usually generated by the transactions in dispute.⁹⁵ The DDRO may request the business provide an explanation verbally or in writing about facts relevant to the mediation and/or submit documents relating to the complaint. The business may not object absent “justifiable reason.”⁹⁶

Financial service providers have a legal obligation, based on their Basic Contract with the DDRO, to answer the claims filed and to participate in DDRO mediation proceedings.⁹⁷ The application for mediation is provided to the opposing party who then submits a written answer detailing the party’s “defense,” including a “[w]ritten answer to the purport of the motion” and the “[d]efense or plea against the points of the [d]ispute.”⁹⁸ The party may also submit at that time documentary evidence related to the answer.⁹⁹

The DDRO maintains a roster of mediators,¹⁰⁰ and the statute provides that it is to “appoint” a mediator or mediators following an application for dispute resolution.¹⁰¹ The statutes implementing

92 FINMAC, OPERATIONAL RULES, *supra* note 60, at arts. 13, 15-2, 16. See also JBA, Operating Rules, *supra* note 71, at art. 9.

93 FIEA, *supra* note 23, at art. 156-50(8) (Japan). JBA, Operating Rules, *supra* note 71, at art. 11. Though the rules permit the parties to seek mediation without going through the complaint resolution process, the JBA Consultation Office reports that only ten to twenty percent of the claimants do so. E-mail from Japanese Bankers Ass’n Secretariat to authors (Aug. 12, 2016) [hereinafter E-mail from JBA Secretariat to authors] (on file with authors).

94 See, e.g., FINMAC, DETAILED RULES, *supra* note 79, at annex 3; JBA, *Assen n’o Moushitatesho* [Mediation Application Instructions], http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/format_visitor/format_visitor11.pdf (Japan).

95 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 30(2). JBA, Operational Rules, *supra* note 65, at art. 25(1).

96 See, e.g., FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 37.

97 *Id.* at arts. 5-3(3), 27.

98 FINMAC, DETAILED RULES, *supra* note 85, at annex 7; FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 35 para. 1.

99 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 35 para. 1.

100 FSA, *supra* note 40, at 16.

101 FIEA, *supra* note 23, at art. 156-50(2) (Japan).

the Financial ADR System leave the number of mediators utilized in each case to the DDROs: some utilize a single mediator and some a panel.¹⁰²

The FIEA and related Cabinet Ordinance establish minimum qualifications for mediators working for the DDRO.¹⁰³ Mediators are to be “persons of the highest moral character”¹⁰⁴ selected from five different categories of persons: attorneys with more than five years of experience;¹⁰⁵ individuals engaged in work related to consumer protection for ten years or more at a registered juridical person handling complaints related to financial services;¹⁰⁶ consumer

102 *Id.* The majority, five DDROs including the JBA, utilize a panel of three mediators with mixed backgrounds. A minority, FINMAC and the Japan Financial Services Association, utilize a single attorney mediator. The Japanese Trust Companies Association departs from both majority and minority practice by appointing a panel of five mediators, consisting of two lawyers, two consumer advocates, and one employee of the association. See YUTAKA TANAKA, Q&A KIN'YUU ADR NO TEBIKI [Q&A GUIDEBOOK TO FINANCIAL ADR] 10–12 (2014) (Japan); ‘Shitei Funsou Kaiketsu Kikan Muks no Sougou Teki na Kantoku Shishin’ nado he no Taiou Joukyou [“Comprehensive Supervision Guidelines for Designated Dispute Resolution Organizations”] FIN. SERVS. AGENCY 1, 30 (June 15, 2013), http://www.fsa.go.jp/singi/singi_trouble/siryou/20150615/05.pdf (Japan).

103 FIEA, *supra* note 23, at art. 156-50(3) (Japan). See Fin. Servs. Agency, Soumu Kikakukyoku [General Affairs Planning Bureau], *Kin'yuu ADR Seido Ni Tsuite* [*About the Financial ADR System*], in 21 HEISEI KIN'YUU SHOUIN TORIHIKI HOU KAISEI TO NI KAKAWARU SEIREI, NAIKAFUREI NO GAIYOU (2009) (Japan).

104 FIEA, *supra* note 23, at art. 156-50(3) (Japan).

105 *Id.* at art. 156-50(3)(1). The act refers specifically to attorneys. *Id.* The Cabinet Ordinance also permits mediation by one who has worked for five years or more as a judge, assistant judge, public prosecutor, attorney, or professor of law. *Kin'yuu Shohin Torihiki Ho Dai 5 Sho no 5 no Kitei ni yoru Shitei Funso Kaiketsu Kikan ni kansuru Naikakufu-Rei* [Cabinet Office Ordinance on Designated Dispute Resolution Organizations under the Provisions of Chapter V-5 of the Financial Instruments and Exchange Act], Cabinet Office Ordinance No. 27, art. 11, para. 3 (Horei teikyo de-ta shisutemu [Horei DB]), <http://law.e-gov.go.jp/htmldata/H21/H21F10001000077.html> (Japan). See also Fin. Servs. Agency, Soumu Kikakukyoku [General Affairs Planning Bureau], *Kin'yuu ADR Seido Ni Tsuite* [*About the Financial ADR System*], in 21 HEISEI KIN'YUU SHOUIN TORIHIKI HOU KAISEI TO NI KAKAWARU SEIREI, NAIKAFUREI NO GAIYOU (Japan).

106 *Kin'yuu Shouhin Torihikigyou to ni Kansuru Naikakufurei* [Cabinet Ordinance on the Financial Instruments and Exchange Industries], Cabinet Office Ordinance No. 52 of 2007, art. 11(3)(ii) (Japan). This category includes current or former secretariats of self-regulatory organizations. See Fin. Servs. Agency, *Kinyū-ADR-Seido-no Forōappu-ni Kansuru Yūshikisha Kaigi*, [Summary of Arguments at the Council of Advisers Regarding Evaluating the

advocates who have worked in designated consumer counseling occupations for more than five years;¹⁰⁷ and judicial scriveners if the dispute involves a demand of less than 1.4 million yen.¹⁰⁸ Finally, the statute authorizes other persons specified by Cabinet Ordinance as having equivalent qualifications, e.g., judges, prosecutors, and law professors with more than five years' experience.¹⁰⁹

The DDRO rules exclude mediators if there is a conflict of interest or "any other circumstances that are likely to hinder the fair implementation" of the dispute resolution procedures.¹¹⁰ By ordinance, conflicts are defined to include relationships with "interested" persons including a spouse or former spouse of a party; a party's relative within the fourth degree of kinship; a party's guardian; a party's representative or agent in the dispute; and those having received income from one of the parties within the prior three years.¹¹¹

Meditators assume a duty of confidentiality, to refrain from divulging or using for personal benefit any confidential information learned in the course of the mediation.¹¹² For the purposes of criminal law and other penalties, the mediator is deemed to be an official engaged in public service.¹¹³

After appointment, the DDRO sends the application to the appointed mediator(s).¹¹⁴ If they determine that there is a conflict, that they lack the necessary expertise, the application was filed for improper purposes, or the case should be heard by another dispute resolution organization, they may decline to proceed and notify the

Financial ADR System] 7-8 (Mar. 8, 2013), <http://www.fsa.go.jp/singi/adr-followup/20130308.html> (Japan).

107 Cabinet Office Ordinance No. 52 of 2007, art. 11(3)(iii) (Japan). Cabinet Ordinance Article 11(2)(i-iii) defines these occupations to include a consumer counseling specialist certified by the National Consumer Affairs Center of Japan, an advisory specialist for consumer affairs certified by the Japan Industrial Association, and a consumer consultant certified by the Japan Consumer's Association. *Id.* at art. 11(2)(i-iii).

108 *Id.* at art. 11(2)(iv)-(v).

109 *Id.*

110 FIEA, *supra* note 23, at art. 156-44(4)(ii) (Japan).

111 Cabinet Office Ordinance No. 52 of 2007, art. 11(1) (Japan).

112 FIEA, *supra* note 23, at art. 156-41(1) (Japan). FINMAC, OPERATIONAL RULES, *supra* note 60, at arts. 42, 47 paras. 1-2. The operational rules further establish rules for the DDRO and mediators to "reliably retain any confidential information." FIEA, *supra* note 23, at art. 156-44(4)(xiv) (Japan).

113 *Id.* at art. 156-40(2).

114 *Id.* at art. 156-50(4).

applicant providing their reason for doing so.¹¹⁵

If the mediator(s) accept the appointment, the DDRO notifies the parties of the name(s) of the mediators and their acceptance date(s).¹¹⁶ The parties may then challenge the appointment of a mediator by filing a “Motion for Challenge.”¹¹⁷ Mediators appointed by the DDRO other than the challenged mediator then determine whether there is cause to strike the challenged mediator by a majority vote.¹¹⁸

If the mediators proceed, they have authority to “hear the opinions of the parties” and request that they submit written reports, books, documents, or other relevant articles.¹¹⁹ Such requests may extend to documents often protected from discovery in the course of civil litigation.¹²⁰ Some DDROs permit in camera review of this information by mediators; some discourage in camera review and recommend disclosure of all documents produced in mediation to all parties.¹²¹

The mediator(s) will typically solicit statements and supporting documents from each party prior to the hearing.¹²² The mediator(s) may request the parties or witnesses submit to an

115 *Id.* at art. 156-50(4)(5).

116 See, e.g., FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 30 para. 2.

117 *Id.* at art. 32 para. 3; FINMAC, DETAILED RULES, *supra* note 85, at annex 6.

118 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 33 para. 3.

119 FIEA, *supra* note 23, at art. 156-50(6) (Japan).

120 The Code of Civil Procedure protects from discovery documents generated “exclusively for internal use,” e.g., daily reports from sales persons and internal sales manuals; mediators may request the same. Minji Soshou Hou [Minshouhou] [C. CIV. PRO.], Law No. 109 of 1996, art. 220(iv)(d) (Japan), <http://law.e-gov.go.jp/htmldata/H08/H08HO109.html>; Hironori Nishikino, *Kin'yuu ADR no Chakuganten: Seidou/Tetsuzuki no Tokuchou to Taioutokino Juu Gensoku*, 1946 KIN'YUU HOUMU JIJOU 34 (2012) (Japan).

121 Nishikino, *supra* note 120, encourages financial service providers to request “mediator’s eyes only” arrangements. E-mail from Tokyo attorney to authors (Aug. 20, 2016) [hereinafter E-mail from Tokyo attorney to authors] (on file with authors) and E-mail from JBA Secretariat to authors, *supra* note 93, suggest this happens rarely.

122 See, e.g., ‘Shitei Funsou Kaiketsu Kikan Muke no Sougou Teki na Kantoku Shishin’ nado he no Taikyou Joukyou [Status of Compliance with “Comprehensive Supervision Guidelines for Designated Dispute Resolution Organizations”] 2015 FIN. TROUBLE CONTACT ADJ. COUNCIL 1, 9, 151 (June 15, 2015), http://www.fsa.go.jp/singi/singi_trouble/siryou/20150615/05.pdf (Japan). Regarding the authority of mediators to request necessary materials, see JBA, Operating Rules, *supra* note 71, at art. 26(5); FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 37 para. 1.

interview on a set date, and the parties are obligated to attend.¹²³ The proceedings are not open to the public; however, the mediators may permit attendance by “appropriate persons” as agreed by the parties.¹²⁴

Some DDROs structure the mediation to include both joint and individual sessions with the mediator(s); others utilize only separate sessions.¹²⁵ With the former, the parties typically offer opening statements in a joint session, and if the customer is appearing pro se, the mediator(s) will often ask questions of the customer to facilitate his or her statement of the claim.¹²⁶ Following this joint session, the parties are asked to leave the mediation room, and mediator(s) then meet individually with the parties.¹²⁷ Following a series of individual sessions, if the mediator(s) are able to forge an agreement, the parties will reconvene in order for the mediator(s) to confirm the terms of the settlement.¹²⁸ The mediator(s) will terminate the hearing at the point at which they believe it is not possible to reach an agreement.¹²⁹

123 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 36.

124 FIEA, *supra* note 23, at art. 156-50(7) (Japan).

125 Compare Tomonori Ishizuka, *Kin'yuu ADR Benricho: Moushitate kara Kaiketsu madeno Tadashii Riyohoga Wakaru* [Quick Reference for Financial ADR: Understanding the Correct Path from Filing to Disposition], KIN'YU JIJO KENKYUKAI 168–169 (2012) (Japan), with JBA, Operating Rules, *supra* note 71, at art. 30 (2), and the General Insurance Association of Japan, Kujo-Shori oyobi Funso-Kaiketsu-Tetsuzuki-to ni Kansuru Gyomu-Kitei [Operational Rules for Complaint Handling and Dispute Resolution], art. 29, <http://www.shintaku-kyokai.or.jp/profile/pdf/soudankitei.pdf> (Japan), with FINMAC, OPERATIONAL RULES, *supra* note 60.

126 Ikenaga et al., *Zadankai: Kin'yuu ADR no Genzai* [Round-Table Talk: Current Situation of Financial ADR], 1946 KIN'YUHOMU JIJO 6, 17–19, 25–26 (Japan).

127 E-mail from Financial Instruments Mediation Assistance Center Secretariats (Sep 9, 2013) [hereinafter E-mail from FINMAC Secretariats to authors] (on file with authors). See also Kenichi Yamaguchi, *Koseichuritsu Kanijinsoku Nattoku Tomeisei* [Fairness, Neutrality, Expediency, Content, and Transparency], 12 FINMAC 6 (2014) (Japan).

128 E-mail from FINMAC Secretariats to authors, *supra* note 127. Yamaguchi, *supra* note 127.

129 E-mail from FINMAC secretariats to authors, *supra* note 127. FINMAC's Operational Rules provide a multifaceted test: if the mediator “judge[s] that a settlement cannot be reached” and any one of the following grounds are met, the mediator “shall” terminate the mediation: a party “clearly indicates its intention not to settle”; settlement cannot be reached immediately and there is concern that the settlement’s “disadvantages” to the parties outweigh “the benefits to be obtained”; or a party is absent from two consecutive mediations or three total mediations “without a justifiable reason.” FINMAC,

When there is no prospect for settlement, mediator(s) may prepare and issue a Special Conciliation Proposal.¹³⁰ The Special Conciliation Proposal requires a financial service provider to accept the proposal, unless the customer declines to accept the proposal or the financial service provider files suit within one month of learning the customer has accepted the proposal.¹³¹ In other words, if the customer agrees to the Special Conciliation Proposal, it is binding, unless contested by filing suit in civil court.¹³²

Filing an application for dispute resolution with a DDRO tolls the period of prescription.¹³³ If suit is filed in civil court within one month from notice of termination of the mediation, it is deemed filed in court as of the date the applicants' filing for mediation with the DDRO.¹³⁴ Where the customer accepted a Special Conciliation Proposal, the financial service provider has one month to file a suit to contest the Proposal.¹³⁵ If litigation is pending at the time of the application, the parties may jointly petition the court to suspend the proceedings for a fixed period of no longer than four months.¹³⁶

If mediation is successful, payment from the financial service provider to its customer, pursuant to the DDRO proceedings, is exempted from the loss compensation regulations requiring

OPERATIONAL RULES, *supra* note 60, at art. 38(1). Further, if a mediator “judges” that a settlement cannot be reached, he or she “may” terminate the mediation if a party files suit or seeks civil conciliation; a party files for arbitration or mediation with another organization; or there are other grounds that may make mediation inappropriate. *Id.* at art. 38(2).

130 FIEA, *supra* note 23, at art. 156-44 (2)(v) (Japan).

131 *Id.* at art. 156-44(6)(i-ii). FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 40-2 paras. 1-2.

132 FIEA, *supra* note 23, at art. 156-44 (6) (Japan).

133 *Id.* at art. 156-51. Banking Act, *supra* note 48, at art. 52-74(1) (Japan). While there are differences, a period of prescription is generally analogous to a statute of limitations in common law countries. See, e.g., Chase Sec. Corp. v. Donaldson, 325 U.S. 304, 313 (1945); Marc S. Firestone, *Prescription-What You Don't Know Can Hurt You-Louisiana Adheres to A Three Year Limit on the Discovery Rule*, 58 TUL. L. REV. 1547, 1549 (1984).

134 FIEA, *supra* note 23, at art. 156-51(1) (Japan).

135 *Id.* at art. 156-44 (6)(ii). This filing of a lawsuit indicates a lack of prospects for settlement, upon which mediator(s) should terminate the mediation procedures. While FINMAC's operational rules do not mention rejecting a Special Conciliation Proposal as grounds for terminating the mediation, the JBA rules state that rejection of a Special Conciliation Proposal automatically terminates the process. FINMAC, OPERATIONAL RULES, *supra* note 60; JBA, Operational Rules, *supra* note 65.

136 FIEA, *supra* note 23, at art. 156-52(10) (Japan).

regulatory approval.¹³⁷ DDROs may assist with enforcement of mediated settlements by investigating claims of non-payment, notifying the Financial Supervisory Agency of the same, and expelling the non-complying financial service provider from the DDRO for breach of the basic contract.¹³⁸

According to the Financial Supervisory Agency, the Financial ADR System's goals are to provide dispute resolution that is "fast, simple, and flexible in accordance with the parties' circumstances and the nature of the case."¹³⁹ We discuss whether it reaches those goals in the next section: first, by outlining the procedures used at two widely used DDROs, and then by discussing the implications.

IV. DDRO PRACTICE AT FINMAC & THE JBA

Disputes alleging a violation of the FIEA are resolved by two DDROs: the Financial Instruments Mediation Assistance Center (FINMAC) and the Japanese Bankers Association's Consultation Office (JBA).¹⁴⁰

The five self-regulatory organizations overseeing the securities and futures industries in Japan established FINMAC in 2009 as a non-profit entity for resolving disputes with investors.¹⁴¹ It started operation as a Certified Investor Protection Organization

137 Kin'yuu Shouhin Torihikigyou to ni Kansuru Naikakufurei [Cabinet Ordinance on the Financial Instruments and Exchange Industries], Cabinet Office Ordinance No. 52 of 2007, art. 119(1)(iv) (Japan).

138 FIEA, *supra* note 23, at art. 156-45(1) (Japan).

139 Sankou, *supra* note 47.

140 FINMAC, *Outline of FINMAC, FINANCIAL INSTRUMENTS MEDIATION ASSISTANCE CENTER*, <https://www.finmac.or.jp/gaiyou/> (last visited Dec. 29, 2016) (Japan); JBA, *Zenkoku Ginkou Soudanshitsu toha*, <http://www.zenginkyo.or.jp/adr/about/> (last visited Dec. 29, 2016) (Japan).

141 FINMAC, *PURPOSE OF ESTABLISHMENT OF FINMAC 1*, https://www.finmac.or.jp/english/pdf/en_01.pdf (last visited Jan. 26, 2018) [hereinafter FINMAC, *PURPOSE OF ESTABLISHMENT OF FINMAC*] .

in 2010,¹⁴² and began operating as a DDRO in April of 2011.¹⁴³ The JBA's Consultation Office is an internal office within the Japanese Bankers Association, the trade association for Japanese banks.¹⁴⁴ In October 2008, the office acquired its designation as a Certified Investor Protection Organization¹⁴⁵ and started operation in April 2010.¹⁴⁶ It began work as a DDRO in October of the same year.¹⁴⁷

While DDRO proceedings are confidential, both FINMAC and the JBA publish anonymized outlines of cases.¹⁴⁸ Attorneys with

142 *Id.*; FINMAC, HISTORY OF THE FINANCIAL INSTRUMENTS MEDIATION ASSISTANCE CENTER 1 https://www.finmac.or.jp/english/pdf/en_03.pdf (last visited Jan. 26, 2018). The Japan Securities Dealers Association; the Investment Trusts Association, Japan; the Japan Investment Advisors Association; the Financial Future Association Japan; and the Japan Commodity Futures Association constitute the original five member associations. Later in 2011, the Japan Commodity Futures Association dissolved and the Type II Financial Instruments Firms Association joined FINMAC. *Id.*; FINMAC, *Heisei-21-Nendo Jigyo-Hokokusho [The Business Report for FY 2009]* 4, http://www.finmac.or.jp/sosiki-gaiyou/pdf/jigyou_H21.pdf (Japan). See also Financial Instruments and Exchange Act, Law No. 25 of 1948, arts. 79-7 to 79-19 (Japan).

143 FINMAC, PURPOSE OF ESTABLISHMENT OF FINMAC, *supra* note 141, at 3.

144 JBA, *supra* note 140.

145 JAPANESE BANKERS ASS'N, TOUSHISHA HOGO SHISHIN (July 22, 2008), <http://www.zenginkyo.or.jp/fileadmin/res/abstract/clinic/reference02/certification01.pdf> (Japan). Certified Investor Protection Organizations are industry associations that are not operating as self regulatory organizations, but are certified by the Financial Supervisory Agency to mediate disputes between customers and financial service providers in a particular industry. FIN. SERVS. AGENCY, NEW LEGISLATIVE FRAMEWORK FOR INVESTOR PROTECTION 11 (2006), available at <http://www.fsa.go.jp/en/policy/fiel/20061010.pdf> (Japan). The 2006 amendments to the FIEA established this certification. *Id.* at 2. Apart from the JBA, the FSA has recognized the Life Insurance Association of Japan, the General Insurance Association of Japan, and the Trust Companies Association of Japan as Certified Investor Protection Organizations. E-mail from Kyoto attorney to authors (Jan. 23, 2017) [hereinafter E-mail from Kyoto attorney to authors] (on file with authors).

146 JBA, *Heisei 20 Nendo 'Ginko-Torihiki Sodanjo' Toriatsukai-Jokyo 'Assen Linkai' Un'ei-Jokyo [State of Implementation for FY 2008 of Banking Counseling Center and Mediation Committee]* 10, http://www.zenginkyo.or.jp/fileadmin/res/abstract/clinic/reference01/conditions01_2008.pdf (Japan); Fin. Servs. Agency, *Nintei-Toshisha-Hogo-Dantai no Nintei ni Tsuite [Declaration of Certification of Certified Investor Protection Organization]* (2008), <http://www.fsa.go.jp/news/20/syouken/20081001-2.html> (Japan).

147 JBA, *Zenkoku Ginkou Kyokai Funso-Kaiketsu-to-Gyomu no Jisshi-Jokyo* 1, http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/conditions/conditions01_2203_1.pdf (Japan).

148 FIEA, *supra* note 23, at art. 156-50(9) (Japan). This article requires the DDRO

experience representing clients or serving as mediators in DDRO proceedings have also published articles providing insight into procedure and practice at the two major DDROs.¹⁴⁹ Both sources, along with information from interviews with practitioners and staff at FINMAC and JBA are discussed here.

A. *FINMAC-ADR for the Investment Industry*

FINMAC resolves disputes between financial service providers and their customers in three different legal capacities.¹⁵⁰ First, it serves as a DDRO for disputes involving “Type I Financial Instruments Businesses,” which are primarily securities companies and financial futures traders.¹⁵¹ Second, as a Certified Investor

to prepare a record detailing the date of the application, the names of the parties and mediators, the “particulars” of the dispute resolution process, the results, including the reasons for termination and date thereof, and any “particulars necessary for clarifying the contents of the implemented Dispute Resolution Procedures” as specified by Cabinet Ordinance. *Id.*

- 149 See, e.g., TOKYO BENGOSHIKAI BENGOSHI-GYOMU-BU KIN'YUU FUNSO KENKYUKAI, *Q&A Kin'yuu ADR Katsuyo Gaidobukku: Kaiketsudekiru Shoken Ginko Hoken no Toraburu* [Q&A Guidebook for Utilizing Financial ADR: You can Resolve Troubles about Securities, Banking, and Insurance!] (Nihon Kajoshiki Shuppan, 2012), at 85 (Japan); Hironori Nishikino, *Kin'yuu ADR Taio no Chakugantен: Seido Tetuzuki no Tokucho to Taioji no 10 no Gensoku* [How to Deal with Financial ADR as an Attorney for Financial Service Provider: Its Distinctive Features and Ten Principles to Remember], 1946 KIN'YUU HOMU JIJO 34 (Japan). See also TANAKA, *supra* note 102; Ikenaga et al., *supra* note 126; Ishizuka, *supra* note 125.
- 150 The Operational Rules Concerning Complaint Resolution Support and Mediation of FINMAC defines its authority. FINMAC, OPERATIONAL RULES, *supra* note 60. All of its mediation cases are processed pursuant to its authorization as a Certified Dispute Resolution Business Operator, *Id.* at art. 4 para. 4, all complaints and disputes regarding Type I Financial Instruments Business are processed as a DDRO, *Id.* at art. 4 para. 5, and all complaints and disputes involving Type II Financial Instruments Business who entered into a Basic Contract with FINMAC are processed as a Certified Investor Protection Organization. *Id.* at art. 4 para. 6.
- 151 *Id.* at art. 4 para. 5. Type I Financial Instruments Businesses include financial service providers conducting any of the following: the purchase/sale of securities or (domestic and foreign) market transactions in derivatives; intermediation, brokerage, or agency thereof; intermediation, brokerage, or agency for entrustment thereof; brokerage for the clearing of securities; the secondary distribution of securities or exclusive offers to sell; the handling of a public offering or secondary distribution of securities, or the handling of a private placement of securities or exclusive offer to sell to professional investors; over-the-counter derivative transactions or intermediation, brokerage, or agency thereof; brokerage for the clearing of over-the-counter derivatives; underwriting of securities; purchase/sale of securities or

Protection Organization,¹⁵² it has authority to mediate disputes with “Type II Financial Instruments Businesses,” which are primarily direct sellers of investment trusts and mutual funds.¹⁵³ Finally, its certification by the Minister of Justice as a Dispute Resolution Organization gives FINMAC accreditation to mediate civil disputes generally.¹⁵⁴ With this certification, filing for mediation with FINMAC tolls the period of prescription for the claim,¹⁵⁵ and a claim mediated

intermediation, brokerage, or agency thereof via propriety trading system; or securities management business. FIEA, *supra* note 23, at art. 28 (1) (Japan).

152 FIEA, *supra* note 23, at arts. 79-7 to 79-19 (Japan).

153 FINMAC, *OPERATIONAL RULES*, *supra* note 60, at art. 4, para. 6. Type 2 Financial Instruments Businesses include financial service providers conducting any of the following: public offering or private placement of such securities as beneficiary certificates of investment trusts managed under instructions of settlor, beneficiary certificates of foreign investment trusts, (domestic and foreign) mortgage securities, collective investment scheme equity, and beneficiary certificates of investment trusts; purchase/sale of securities with limited negotiability or (domestic and foreign) market transaction of derivatives; intermediation, brokerage, or agency thereof; intermediation, brokerage, or agency for entrustment thereof; brokerage for the clearing of securities, etc.; the handling of a public offering or secondary distribution of securities, or the handling of a private placement of securities with limited negotiability or exclusive offer to sell, etc. to professional investors; or market transaction of derivatives not connected with securities. FIEA, *supra* note 23, at art. 28(2) (Japan).

154 Ministry of Justice, *Kaiketsu Sapo-tto: Ninsho Funso Kaiketsu Sa-bisu*, <http://www.moj.go.jp/KANBOU/ADR/> (last visited May 24, 2017) (Japan).

155 Act on Promotion of Use of Non-Judicial Dispute Resolution Procedures, *supra* note 2, at art. 25. FINMAC is the only DDRO certified, pursuant to the Promotion of Use of Alternative Dispute Resolution Act, as a Certified Dispute Resolution Service Provider. See Ministry of Justice, *Kaiketsu Sapoto Ichiran [Directory of Certified Dispute Resolution Organizations]*, <http://www.moj.go.jp/KANBOU/ADR/jigousya/ninsyou-index.html> (Japan). FINMAC’s unique position derives from its origins. Prompted by enactment of the FIEA in 2006, the five self regulatory organizations (SROs) in the investment industry established a unified organization for dispute resolution. FINMAC, *supra* note 140. Each had authority to resolve disputes individually, but the statutory ban on loss compensation necessitated accreditation as a Certified Dispute Resolution Organization to gain exemption from the administrative confirmation requirement for loss compensation. *Id.*; Fin. Servs. Agency, *Jishukisei-Kikan Ichiran [Directory of Self-Regulatory Organizations]*, <http://www.fsa.go.jp/koueki/koueki10.html> (Japan); FIEA, *supra* note 23, at arts. 77 to 77-3 & 78-6 to 78-8 (Japan). After enactment of the Financial ADR amendments, DDROs no longer needed Ministry of Justice certification to gain exemption from the administrative confirmation requirement. The investment industry SROs, however, continue to advocate for the use of SRO-sponsored ADR in the legislative process, as opposed to transubstantive private ADR organizations, because of the technical knowledge required to

by FINMAC is exempted from administrative confirmation of the settlement otherwise required by the ban on loss compensation.¹⁵⁶

FINMAC mediates claims using a single mediator, rather than a panel,¹⁵⁷ and its roster of mediators currently includes thirty-eight individuals—all attorneys.¹⁵⁸ FINMAC’s website provides a list of their names and geographical region, but it provides no additional information regarding their backgrounds.¹⁵⁹

FINMAC mediators most commonly hear claims asserting violations of “the duty to explain” and suitability principle; claims involving unauthorized trading; and claims alleging improper execution of the trade.¹⁶⁰ FINMAC mediators typically meet with the parties in an initial joint session, followed by separate sessions with the individual parties.¹⁶¹ When the mediator finds settlement possible, the mediator proposes a settlement plan.¹⁶² If the mediator

understand the disputed transactions. See Dai 35 Kai Kyougikai ni okeru Giron wo Fumaeta Kin’yuu ADR no Houkousei ni Kansuru Gyoukai Dantai no Iken [See 35th Council Industry Organization’s Opinion on the Direction of Financial ADR], at 3–5 (2008), http://www.fsa.go.jp/singi/singi_trouble/siryou/20080514/03.pdf (Japan).

156 Kin’yuu Shouhin Torihikigyou to ni Kansuru Naikakufurei [Cabinet Ordinance on the Financial Instruments and Exchange Industries], Cabinet Office Ordinance No. 52 of 2007, art. 119(1)(vii) (Japan).

157 See, e.g., ‘Shitei Funsou Kaiketsu Kikan Muke no Sougou Teki na Kantoku Shishin’ nado he no Taikyou Joukyou, *supra* note 122 (Japan); Soudan nado no Nagare, FINMAC, <https://www.finmac.or.jp/flow/#1> (last visited Dec. 27, 2016) (Japan).

158 *Assen In Meibo*, FINMAC, <http://www.finmac.or.jp/meibo/> (last visited Aug. 5, 2016) (Japan).

159 *Id.*

160 See FINMAC no Gyoumu no Genkyoutou Ni Tsuite [About the Current State of Work of FINMAC] 2013, <http://www.moj.go.jp/content/000111909.pdf> (Japan). The “duty to explain” is an independent duty recognized under general principles of tort law requiring explanation, as opposed to disclosure of potential risks and benefits of a financial transaction. See, e.g., Pardieck, *supra* note 17, at 54–67.

161 E-mail from FINMAC Secretariats to authors, *supra* note 127. Yamaguchi, *supra* note 127, at 6 (Japan).

162 FINMAC, Operational Rules, *supra* note 60, at art. 40. Yamaguchi, *supra* note 127, at 6. A proposal for compensation of nearly all the damages claimed typically reads as follows: “Based on arguments and documents provided from both parties, we can confirm the fact that the respondent gave the claimant misleading materials on solicitation. Therefore, it is apparent that the respondent is liable. It is reasonable to deduct the amount of the interest the claimant already received.” See Tokutei Dai-Isshu-Kin’yu-Shohin Torihiki-Gyomu Funso-Kaiketsu-Jirei (Heisei 23 nen 10-12 gatsu) [Cases from Dispute Resolution for Specified Type I Financial Instruments Transaction Services:

can determine the basic facts from the arguments and the documents provided, the proposals generally apportion “fault” between the parties.¹⁶³ Otherwise, their proposals often point out the relative weakness of arguments and evidence of both sides.¹⁶⁴ In either case, there is a strong evaluative component to the mediation.¹⁶⁵

With individual investors, the mediators’ evaluations often discuss the parties’ relative abilities and experience, both of which have been cited as important factors in court decisions.¹⁶⁶ While an

October to December, 2011], Case No. 4 of 2011, http://www.finmac.or.jp/tokei-siryo/pdf/sitei-assen/2011_10-12_conflict.pdf (Japan). See also Eriko Matsuno, ‘*Nattoku no Iku Kaiketsu*’ wo Mezashite, 3 FINMAC 17 (Japan).

163 See Yamaguchi, *supra* note 127, at 6; Funsoukaiketsu Jireishuu (Shihanki Goto) [Quarterly Case Reports], 2017, FINMAC, http://www.finmac.or.jp/tokei-siryo/index_06.html (Japan). A typical proposal for partial compensation of damages finds “fault” with both parties: “The parties disagree about the explanation given during solicitation, and it is difficult for us to grasp the facts. While it is questionable that the respondent (broker) fulfilled its duty to fully explain the transaction, we cannot deny claimant’s own negligence either. Therefore, we deem it reasonable for parties to concede and settle on the proposed amount...[¥460,000 following a demand of ¥2 million]. FINMAC, REPORT OF MEDIATED CASES REGARDING TYPE I FINANCIAL INSTRUMENTS BUSINESS 1 (Oct – Dec, 2011), http://www.finmac.or.jp/tokei-siryo/pdf/sitei-assen/2011_10-12_conflict.pdf (Japan).

164 A proposal for partial compensation of damages which refers to evidential strength reads as follows: “While we cannot readily recognize a breach of the duty to explain by the respondent (broker), the respondent should admit to having processed at least some transactions without permission and to mistaken orders, as he cannot produce strong evidence refuting claimant’s allegations. Therefore, we deem it reasonable for parties to concede and settle on the proposed amount [¥1.05m following a demand for ¥5.06m].” FINMAC, Report of Mediated Cases Regarding Type I Financial Instruments Business (Oct–Dec, 2011) 7, http://www.finmac.or.jp/tokei-siryo/pdf/sitei-assen/2011_10-12_conflict.pdf (Japan). For general approach by mediators, see Matsuno, *supra* note 162; Yamaguchi, *supra* note 127.

165 “While some mediators steadfastly avoid expressing any opinion about the strength or viability of a disputant’s position, others incorporate such judgments into their standard arsenal of conflict resolution techniques [M]ediators at the extreme end of the evaluative spectrum adopt strategies ‘intended to direct some or all of the outcomes of the mediation.’ At the other end of the continuum, highly facilitative mediators adopt strategies ‘intended simply to allow the parties to communicate with and understand one another.’” Ellen A. Waldman, *The Evaluative-Facilitative Debate in Mediation: Applying the Lens of Therapeutic Jurisprudence*, 82 MARQ. L. REV. 155, 155-6 (1998).

166 Saikō Saibansho [Sup. Ct.] July 14, 2005, 59 MINSHU No. 6, 1323 (English summary available from http://www.courts.go.jp/app/hanrei_en/detail?id=801) (Japan); Osaka High Court, June 24, 1997, 1620 HANREI

investor with considerable experience is blamed for his or her lack of independent research, the financial service provider is faulted for its lack of confirmation of the “actual understanding” of inexperienced, aged, or investors with diminished capacity.¹⁶⁷

FINMAC mediators will end the mediation if they find settlement implausible.¹⁶⁸ In most reports of failed mediations, mediators cite the large gap in settlement demands or alleged facts, an inability to grasp disputed facts, or the parties’ refusal to compromise.¹⁶⁹

If a FINMAC mediator crafts a special conciliation proposal, and it is rejected by the financial service provider who then files suit, the financial service provider must deposit with FINMAC the amount that they would pay under the special conciliation proposal, which the business may request be returned following the first oral argument before the court.¹⁷⁰

FINMAC’s procedural rules prioritize expedited resolution.¹⁷¹ Their rules require complaints be processed within two months,¹⁷²

JIHO 93 (which was later confirmed by Saikō Saibansho [Sup. Ct.] Oct 12, 1999, 14 SHOKEN-TORIHIKI-HIGAI HANREI-SEREKUTO 301) (Japan). See also Tomoyo Matsui, *Tekigousei-Gensoku Ihan to Fuhokoi-Sekinin* (1) [*Breach of Suitability Principle and Tortious Liability* (1)] (Japan); Hideki Kanda & Hiroyuki Kansaku (eds.), *KIN'YUSHOHIN-TORIHIKI-Ho HANREI HYAKUSEN* [Selected and Annotated Hundred Major Precedents Regarding FIEA] 40 (Yuhikaku, 2013) (Case note on the Supreme Court decisions) (Japan); Reiko Nakasone, *Waranto to Setsumei-Gimu* [*Warrants and the Duty to Explain*] *KIN'YUSHOHIN-TORIHIKI-Ho HANREI HYAKUSEN*, *supra*, at 52 (Case note on the Osaka High Court decision) (Japan); Yoshifumi Imagawa, *Toshi-Songai to Kashitsusosai-Riyu no Mondaiten* [*Issues of Investors’ Damage and Factors Weighed in Comparative Fault Judgement*], 37-2 *KOBEGAKUIN-HOGAKU* 55, 78–79 (Japan).

167 See Tokutei Daiishu-Kin'yushohin-Torihiki-Gyomu Funso-Kaiketsu-Tetsuzuki-Jirei (Heisei 24 nen 1-3 gatsu) [Quarternary Case Reports from Mediation as DDRO (Jan. - Mar. 2012)] 24, http://www.finmac.or.jp/tokei-siryo/pdf/sitei-assen/2012_01-03_conflict.pdf (Japan). See also Yasushige Hagio, *Seigino Mikata !?* [*Mediator as a Knight in Shining Armor !?*], 13 FINMAC 4 (2014) (Japan).

168 FINMAC, *OPERATIONAL RULES*, *supra* note 60, at art. 38 para. 1(1)-(2).

169 See, e.g., Tokutei Daiishu-Kin'yushohin-Torihiki-Gyomu Funso-Kaiketsu-Tetsuzuki-Jirei (Heisei 24 nen 1-3 gatsu), *supra* note 165, at 2, 4, 6, 10.

170 FINMAC, *OPERATIONAL RULES*, *supra* note 60, at art. 40-2 para. 2(3-4).

171 FINMAC, *Kujō-KAIKETSU-SHIEN-TO ASSEN-NI KANSURU GYŌMU KITEI* [Operational Rules for Complaint Process Assistance and Mediation], http://www.finmac.or.jp/kujou/pdf/kisoku02_20150529.pdf (Japan).

172 *Id.* at 15.

and mediation be completed within four months after filing.¹⁷³ Most cases are resolved following one or two hearings.¹⁷⁴

B. The JBA Consultation Office—ADR for the Banking Industry

JBA mediation practice differs from its counterpart at FINMAC in several ways. First, while the FINMAC rules provide for a single mediator who is a licensed attorney to mediate each case,¹⁷⁵ the JBA follows the practice of the majority of DDROs and appoints a panel of three mediators, one from each of the categories discussed above.¹⁷⁶

Attorneys mediating for the JBA are selected based on recommendations from established members of the bar and interviews by an outside advisory panel (*un’ei kondankai*).¹⁷⁷ They chair the mediation panel.¹⁷⁸ Mediators working as consumer advocates are typically recommended by officials from consumer advocacy organizations and are typically certified Consumer Life Consultation Specialists (*shouhisha seikatsu soudan iin*).¹⁷⁹ Mediators coming from the financial services industry are not recommended by member banks but chosen by the secretariat.¹⁸⁰

The JBA discloses on its website the attorneys approved

173 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 43.

174 Reports show that most cases are concluded after no more than two oral sessions. FINMAC, *Assen Kujō Sōdan-no Jisshi Gaikyō* [Implementation Status of Mediation, Complaint and Consultation Business in 2014], (2014), http://www.finmac.or.jp/tokei-siryo/pdf/year/finmac_jyoukyou_2014.pdf (Japan). In 2014, out of 108 cases disposed, 86 cases ended with only one session and 10 cases were terminated with two sessions. *Id.*

175 See ‘Shitei Funsou Kaiketsu Kikan Muke no Sougou Teki na Kantoku Shishin’ nado he no Taikyou Joukyou, *supra* note 122.

176 See TANAKA, *supra* note 102, at 10–11. This structure, a panel including a legal specialist, industry representative and consumer representative, is similar to that utilized in Japan’s Labor Tribunal System. Ryuichi Yamakawa, *Systems and Procedures for Resolving Labor Disputes in Japan*, 34 COMP. LAB. L. & POL’Y J. 899, 904 (2013). There, labor disputes may also be resolved out-of-court by a tribunal of three, including a professional judge and two part-time lay members, one a labor-representative and the other a management representative. *Id.*

177 E-mail from JBA Secretariat to authors, *supra* note 93.

178 *Id.*

179 *Id.* See also National Consumer Affairs Center of Japan, *Shouhi Seikatsu Soudan Iin Shikaku Shiken/Shouhi Seikatsu Sennmon Soudan Iin Shikaku Nintei Seido*, <http://www.kokusen.go.jp/shikaku/shikaku.html> (last visited Aug. 23, 2016) (Japan).

180 E-mail from JBA Secretariat to authors, *supra* note 93.

to chair the mediation panels.¹⁸¹ In interviews, JBA staff indicated that they do not disclose the remainder of its mediator roster to the public out of fear of harassment.¹⁸² At the same time, they suggest that mediators from the financial industry, in fact, often view bank actions more critically and customer actions more leniently than the appointed consumer advocate mediator.¹⁸³

The JBA checks for conflicts with the financial institution at the point the mediator is chosen, followed by monthly conflict checks until resolution of the matter.¹⁸⁴ The mediators themselves are also required to submit, in writing, a statement that they have no conflicts with regard to the parties or transactions in question.¹⁸⁵

Both FINMAC and JBA rules allow parties to challenge the impartiality of a mediator.¹⁸⁶ Both rely, in practice, on the DDRO to check for conflicts of interest.¹⁸⁷ There is no procedural basis to explore the mediator's background: the parties must rely on publically available information, and, to date, no party has filed a motion to strike the chosen mediators.¹⁸⁸

Given the difficulties of scheduling three mediators, the JBA rules provide for a comprehensive one-day mediation, and JBA practice is for mediators and parties to prepare for the hearing with extensive written exchanges.¹⁸⁹

JBA mediation is, as a result, more document intensive than FINMAC mediation.¹⁹⁰ When a claim is filed, the JBA secretariat identifies the disputed issues and organizes related documents for

181 *Id.*

182 *Id.*

183 *Id.*

184 *Id.*

185 *Id.*

186 FINMAC, *OPERATIONAL RULES*, *supra* note 60, at art. 33. JBA, *Operational Rules*, *supra* note 65, at art. 22(2).

187 E-mail from JBA Secretariat to authors, *supra* note 93.

188 *Id.*

189 See FIN. SERVS. AGENCY, *supra* note 102, at 151. Japanese Bankers Ass'n, Funsō-Kaiketsu-Tetsuzuki 'Assen-Linkai' no Nagare [Flowchart of Mediation Process], <http://www.zenginkyo.or.jp/abstract/adr/mediation/mediation-flow/> (last visited Jan. 6, 2016) (Japan).

190 The JBA expects a detailed complaint from the customer at the beginning of the process. See, Gen'ichi Osaka, 'Shitei-Funso-Kaiketsu-Kikan' to Shiten Gyomu-Kaishi go Hantoshikan no Un'yo Jisseki to 'Assen-Linkai' Riyou ni Atatte no Ryui-Jiko [Performance during the First Half-Year Period as DDRO and Remainders for Users of Mediation Committee], 1926 KIN'YU-HOMU-JIJO 30, 34–35 (Japan). For comparison, see Ikenaga et al., *supra* note 126, at 15–16.

the mediators to decide if the dispute is eligible for mediation.¹⁹¹ During the mediation process, JBA mediators prompt parties to submit briefs in advance of the hearing, to offer counterarguments to the complaint or answer, and to send written responses to inquiries to clarify the issues.¹⁹²

JBA officials report that its mediators spend approximately three to four hours at mediation.¹⁹³ Pursuant to the JBA Operational Rules, mediators will meet with the parties separately, unless the mediator determines it is necessary for the parties to meet jointly.¹⁹⁴ Given the written arguments preceding it, mediators use the mediation mainly to confirm details.¹⁹⁵ JBA mediators, as a general principle, require the parties and not counsel to respond to their questions.¹⁹⁶ They request counsel seek permission to speak before doing so and work to ensure that counsel does not become the principle conduit for information.¹⁹⁷

The financial institution determines whether the bank employee(s) involved in the dispute attend the mediation.¹⁹⁸ If they do not, the mediators require the bank employee's supervisor to familiarize him or herself with the details of the transaction and attend the mediation.¹⁹⁹ JBA mediations are, as a result, more like depositions of the parties than opportunities for advocacy.²⁰⁰ The focus is, again, on evaluating the claim.²⁰¹

As a general principle, materials provided to and statements

191 JBA, Operational Rules, *supra* note 65. TANAKA, *supra* note 102, at 67.

192 See FIN. SERVS. AGENCY, *supra* note 102, at 9; JBA, Operating Rules, *supra* note 71, at art 25(3) & 26(6).

193 E-mail from JBA Secretariat to authors, *supra* note 93.

194 Compare JBA, Operating Rules, *supra* note 71, at art. 30 (2), and the General Insurance Association of Japan, *Kujo-Shori oyobi Funso-Kaiketsu-Tetsuzuki-to ni Kansuru Gyomu-Kitei* [Operational Rules for Complaint Handling and Dispute Resolution], at art. 29 (3) (Japan), <http://www.shintaku-kyokai.or.jp/profile/pdf/soudankitei.pdf>, with FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 36.

195 TANAKA, *supra* note 102, at 94–95.

196 *Id.* E-mail from JBA Secretariat to authors, *supra* note 93.

197 JBA, Operational Rules, *supra* note 65, at art. 30 (4).

198 TOMONORI ISHIZUKA, KIN'YUU ADR BENRICHOU: MOUSHITATE KARA KAIKETSU MADENO TADASHII RIYOHOGA WAKARU [QUICK REFERENCE FOR FINANCIAL ADR: UNDERSTANDING THE CORRECT PATH FROM FILING TO DISPOSITION] 168–69 (2012) (Japan).

199 JBA, Operational Rules, *supra* note 65, at art. 30 (3).

200 See FIN. SERVS. AGENCY, *supra* note 102, at 9.

201 See Waldman, *supra* note 157, at 155–56.

made to the mediators are shared with the opposing party.²⁰² As an exception, the mediators may review *in camera* “sensitive” documents such as *ringisho* (internal approval documents) or credit rating information.²⁰³ However, JBA officials suggest that because such information cannot be explained to the opposing party, it cannot form a basis for the mediation proposal.²⁰⁴ As a result, in its mediator training it states *in camera* reviews are to be avoided.²⁰⁵

As with FINMAC mediation, aggrieved customers typically allege a breach of the “duty to explain” and a breach of the duty to refrain from soliciting unsuitable transactions.²⁰⁶ In reviewing the documents submitted, mediators at the JBA routinely focus on whether the bank analyzed the customer’s financial situation before solicitation of the disputed transactions.²⁰⁷ The more thorough the background check, the less likely the banks are to be found guilty of breach of the suitability doctrine.²⁰⁸

At the same time, the JBA’s transactional suitability analysis differs from that employed by FINMAC. Both JBA mediators and the parties routinely prepare detailed financial analyses of assets, diversification, hedging and business needs to determine the suitability of the given transactions.²⁰⁹ In many of the foreign

202 E-mail from JBA Secretariat to authors, *supra* note 93.

203 A *ringisho* has been defined as a “circular letter within a corporation” to record major decisions. Carl F. Goodman, *The Evolving Law of Document Production in Japanese Civil Procedure: Context, Culture, and Community*, 33 BROOKLYN J. INT’L L. 125, 173-4 n.324 (2007) (citing Yasuhei Taniguchi, *The 1996 Code of Civil Procedure of Japan—A Procedure for the Coming Century?*, 45 AM. J. COMP. L. 767, 778 (1997)). Such documents have, in the past, been categorized as “internal correspondence” not subject to disclosure under rules of civil procedure. *Id.*

204 E-mail from JBA Secretariat to authors, *supra* note 93.

205 *Id.*

206 See, e.g., Quarterly Case Reports, *supra* note 163.

207 Mediators at JBA focus on the bank’s proper analysis of the customer’s needs and financial situation based on objective materials before and upon solicitation in cases involving derivatives. See TANAKA, *supra* note 102, at 149, 152, 158, 200, 203–204. They address the business customer’s needs for risk-hedging, including foreign exchange derivatives. *Id.*

208 *Id.*

209 See, e.g., Japanese Bankers Ass’n, *Assen-Moshitate-Jian no Gaiyo to Sono Kekka (Heisei 23 nen Dai 1 Shihan-Ki: Deribatibu Kankei {Kawase Kei} [Outlines and Results of Mediation Cases Terminated in the First Quarter of FY 2011: Foreign Exchange Derivatives Related]*, http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/conditions/classify/classification_2301_1.pdf (last visited Dec. 18, 2017) (Japan). In settled cases, mediators routinely cited insufficient analysis of customers’ needs and financial status by banks as part of their reasoning in the settlement proposal. *Id.*

exchange derivative cases, JBA mediators found dispositive analysis of the customer's business needs for the foreign currency and its exposure to fluctuating exchange rates compared to the volume of currency options traded.²¹⁰ This suitability analysis has assumed particular importance compared to FINMAC mediation where the broker's duty to explain is emphasized, and hedging, diversified portfolios, and other suitability factors are rarely mentioned in the mediation reports.²¹¹

Commentators have argued that JBA practice focuses too narrowly on financial analysis by banks and, in doing so, disregards other elements of suitability and the duty to explain.²¹² The JBA Secretariat describes its own practice as less focused on documents than civil litigation and more focused on a careful questioning to determine the customer's understanding of the transaction and the bank's explanation.²¹³ At the same time, they indicate that they look for objective facts including the percentage of assets invested in high-risk financial products, and the knowledge and experience of the investor in formulating a mediation proposal.²¹⁴

The focus in JBA mediation on suitability can be explained,

210 See TANAKA, *supra* note 102, at 152, 195, 198, 200, 203–04.

211 For JBA mediation cases mentioning these points, see Japanese Bankers Ass'n, *Assen-Moshitate-Jian no Gaiyo to Sono Kekka (Heisei 23 nen Dai 1 Shihan-Ki:Toshi-Shintaku Kankei)* [Outlines and Results of Mediation Cases Terminated in the First Quarter of FY 2011: Investment Trust Related] at 5–8, 10, http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/conditions/classify/classification_2301_2.pdf (Japan). Meanwhile, the authors found only two cases at FINMAC that mention the amount of the given investment in proportion to investor's total assets for the same fiscal year. See FINMAC, *FINMAC Funso-Kaiketsu-Jirei (Heisei 23 Nen 10-12 Gatsu)* [Case Reports for Mediation (Oct.-Dec. 2011)], at 7, <http://www.finmac.or.jp/tokei-siryo/pdf/assen/131017hp.pdf> (Japan); FINMAC, *FINMAC Funso-Kaiketsu-Jirei (Heisei 24 Nen 1-3 Gatsu)* [Case Reports for Mediation (Jan.–Mar. 2012)], at 1 http://www.finmac.or.jp/tokei-siryo/pdf/assen/h24_01_03hp.pdf (Japan).

212 While scholars have evaluated this JBA practice positively, practitioners representing investors have criticized it as "too narrowly focused." Compare Hiroko Aoki, *Chūshōkigyo-muke suwappu-jiken no kinyū ADR ni okeru atukai-ni tuite* [Research Memo on Handling of Dollar Swap Contract Sold to Small-to-Medium Enterprise Customers at Financial ADR] 85–10 HŌRITSU JIHŌ 67, 68–69 (2013) (Japan), with Toshiro Ueyanagi's statement in Takashi Asada et al., *Deribatibu Torihiki-ni Kansuru Saiban-Rei-wo Kangaeru (Zadankai) Ge* [Round-Table-Talk: Examining Recent Judicial Decisions about Derivative Transactions Part 2 of 2], 1986 KINYŪ HŌMU JIJŌ 72, 82 (2014) (Japan).

213 E-mail from JBA Secretariat to authors, *supra* note 93.

214 *Id.*

in part, by its ability to identify hedge ratios and other suitability factors from the accounting and transaction documents produced.²¹⁵ They provide an objective measure, reliance upon which avoids difficult factual determinations that must be made in a single hearing regarding who said what during the solicitation process.²¹⁶

The focus in JBA mediation on suitability can also be explained by the detailed financial information many banks possess about their customers and the analysis that a bank typically conducts in lending to those same customers.²¹⁷ While there is a firewall separating banking and investment functions that limits sharing of undisclosed information about the customers, the Banking Act Ordinance allows a financial institution's banking and investment branches to share undisclosed information about a customer if they give the customer an opportunity to opt out.²¹⁸

It is this dual role of the bank as lender and broker, and their knowledge of their customer, that has caused some JBA mediators to condemn banks for proposing unsuitable derivative transactions that threatened the viability of the customer's business.²¹⁹ They have gone on to propose that the banks discount prearranged early termination charges in currency option cases where banks did not conduct a proper financial analysis prior to solicitation.²²⁰

The Basic Contract between the JBA Consultation Office and its 200 member banks requires the banks to participate in the process and permits mediators to recommend that a bank accept a mediation proposal.²²¹ The Secretariat reports that, as a general

215 *Id.*

216 See Aoki, *supra* note 212; Ikenaga et al., *supra* note 126, at 17.

217 Kinichi Ogata, *Endaka ni Naru hodo Shakkin-Zuke: Chuusho-Kigyo no Himei to Ginko no Sekinin* [Strong Yen, Piling-up Loans: Distress Calls from Small-to-Medium Businesses and Banks's Responsibilities], 4196 SHUKAN EKONOMISUTO 44, 46 (2011) (Japan) (citing comment from an anonymous senior board member of a major bank admitting that they should have kept their customer from sustaining significant losses from foreign exchange derivatives transactions with multiple banks because of their knowledge of the customers as a "main bank.").

218 Banking Act, *supra* note 48, at arts. 10(2), 132-3-2 (Japan). Kin'yuu Shouhin Torihikigyou to ni Kansuru Naikakufurei [Cabinet Ordinance on the Financial Instruments and Exchange Industries], Cabinet Office Ordinance No. 52 of 2007, art. 132 (2) & 153(2) (Japan).

219 See Japanese Bankers Ass'n, *Assen-Jian no Gaiyo to Sono Kekka* [Outlines and Outcomes of Mediated Cases], <http://www.zenginkyo.or.jp/abstract/adr/conditions/year/> (last visited Jan. 7, 2017) (Japan).

220 *Id.*

221 JBA, Operational Rules, *supra* note 65 at art. 34, app. art. 2.

principle, the mediation proposal is to be accepted.²²² And the JBA procedural rules require that a participating bank, if it does not accept the mediation proposal, provide an explanation why, in writing, to the mediation panel.²²³

As with FINMAC, JBA mediators terminate the mediation if they believe settlement is unlikely.²²⁴ They may also draft a special conciliation proposal. Unlike FINMAC, JBA does not require members who reject a special conciliation proposal and file suit to deposit the settlement amount payable under the special conciliation proposal.²²⁵ But it also reports only two instances when a bank has refused to accept a special conciliation proposal.²²⁶

C. *JBA and FINMAC Caseloads*

The statistics are incomplete, but the numbers available suggest that Financial ADR has gained widespread acceptance: JBA filings have surpassed the number of corresponding cases filed in Japanese courts, while FINMAC sees fewer but still a substantial number of filings.²²⁷

222 E-mail from JBA Secretariat to authors, *supra* note 93.

223 JBA, Operational Rules, *supra* note 65, at art. 33.

224 *Id.* at art. 33(3).

225 Compare FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 40-2 para. 2(3), with JBA, Operational Rules, *supra* note 65, at art. 37 (2).

226 E-mail from JBA Secretariat to authors, *supra* note 93. One such case subsequently resulted in a published opinion and the court dismissed the claim from the customer. See *Osaka Chisai* [Osaka Dist. Ct.] Nov. 6, 2015, 1484 KIN'YU SHOUJI HANREI [KINHAN] 37 (Japan).

227 This conclusion is based on an analysis of data from the following sources: Japan Securities Dealers Ass'n, *Gyomu-Hokoku* [Annual Business Report for 2010-2015], <http://www.jsda.or.jp/katsudou/gaiyou/houkokusho.html> (Japan) and Japanese Bankers Ass'n, *Zenginkyo Funso Kaiketsu-to Gyomu no Jisshi Jokyo* [Annual Report of Implementation of Dispute Resolution], <http://www.zenginkyo.or.jp/abstract/adr/conditions/year/> (Japan). While there is a rough parity in the number of lawsuits and mediation filings involving securities brokers, bank customers with complaints about securities sales showed strong preference for DDRO mediation. The difference may be explained partially by relationships and perceptions. Bank customers more often seek to resolve the dispute yet continue their relationship with the bank. See Ikenaga et al., *supra* note 126, at 24 (statement by Sakurai) (Japan); 'Bakuchi Mitaina Shokin wo Urareta' Chuusho Ginko ni Ikari Deribatibu Sonshitsu [Small-to-Medium Businesses Furious against Banks for Their Derivatives Loss], ASAHI SHINBUN, morning ed., June 7, 2012, at 1 (Japan). In contrast, Japan often views brokers as "shady" types, with a dispute ending the relationship and individual investors looking for the protection offered by regulators or the courts. See, e.g., Gaffe-prone Aso brands brokerage workers'shadytypes',

Table 1 Numbers of Mediated Securities Disputes Involving Banks²²⁸

	2010	2011	2012	2013	2014	2015
Civil Litigation	52	52	104	74	30	24
In-Court Mediation	13	22	25	28	9	5
JBA Mediation	211	913	693	168	124	58

Table 2 Numbers of Mediated Securities Disputes Involving Dealer-Brokers²²⁹

	2010	2011	2012	2013	2014	2015
Civil Litigation	310	377	360	341	222	177
In-Court Mediation	75	45	35	31	31	27
SRO Committee Investigation	375	402	361	597	485	387
FINMAC Mediation	239	308	208	128	101	118

THE JAPANTIMES (Aug. 31, 2016), http://www.japantimes.co.jp/news/2016/08/31/business/taro-aso-labels-brokerage-workers-shady-types/?utm_source=Daily+News+Updates&utm_campaign=0051b37a8a-Thursday_email_updates01_09_2016&utm_medium=email&utm_term=0_c5a6080d40-0051b37a8a-331573585#.V8c1tfkrJhE; Nihon-Bengoshi-Rengo-Kai Shohisha-Mondai Taisaku Iinkai [Consumer Affairs Committee of Japan Federation of Bar Ass'ns], KIN'YU-SHOHIN-TORIHIKI-HIGAI KYUSAI NO TEBIKI [GUIDEBOOK TO REDRESS DAMAGE FROM FINANCIAL INSTRUMENT TRADES] 2-24 (Minjiho Kenkyukai, 6th ed. 2015) (Japan).

228 See Japan Securities Dealers Ass'n, *Gyomu-Hokoku* [Annual Business Report for 2010-2015], <http://www.jsda.or.jp/katsudou/gaiyou/houkokusho.html> (Japan); Japanese Bankers Ass'n, *Zenginkyo Funso Kaiketsu-to Gyomu no Jisshi Jokyo (Heisi X Nendo)* [Annual Report of Implementation of Dispute Resolution], <http://www.zenginkyo.or.jp/abstract/adr/conditions/year/> (Japan).

229 See Japan Securities Dealers Ass'n, *Gyomu-Hokoku* [Annual Business Report for 2010-2015], <http://www.jsda.or.jp/katsudou/gaiyou/houkokusho.html> (last visited Jan. 7, 2017) (Japan).

Research shows JBA mediation filings far surpassing the number of cases filed in civil court, and FINMAC filings reaching approximately sixty percent of the number of cases filed in civil court.²³⁰ Other scholars have estimated the total number of currency swap contract disputes between banks and their customers caused by the 2007–2008 financial crisis: investors brought some 2,000 cases to the JBA for complaint or dispute resolution, while only 100 cases were filed in civil court.²³¹

A 2014 Ministry of Justice report examining the total number of private ADR claims filed in 2011 offers another point of comparison.²³² It shows 1,352 filings, pursuant to Japan's general ADR statute, compared to civil court filings ranging from 235,508 in 2009 to 142,487 in 2014.²³³

The Ministry of Justice report suggests that, as a general matter, parties with a dispute favor the courts over private ADR mechanisms.²³⁴ The FINMAC and JBA numbers suggest the opposite: investors routinely choose to mediate instead of litigate.²³⁵ Given that investors have a choice, the number of cases filed with DDROs suggests a measure of success. This raises the question why? Why have these DDROs succeeded in capturing large numbers of disputes?

First, timing is everything. Both DDROs started operation just in time to address disputes arising in the aftermath of the 2007–2008 financial crisis and the rapid rise of Japanese yen on foreign exchanges.²³⁶ In 2004, Japanese banks began encouraging

230 See *supra* Part IV.C (Table 1 Numbers of Mediated Securities Disputes Involving Banks).

231 Aoki, *supra* note 212, at 68–69.

232 HÔMUSHÔ [MINISTRY OF JUSTICE], ADR HOU NI KAN SURU KENTOUKAI HOUHOKUSHO [REPORT MEETING ON ADR] 4 (2014), <http://www.moj.go.jp/content/000121361.pdf> (Japan).

233 Yasuhiko Tanabe, *ADR no Tsukaimichi* [Usage of ADR], 94 BIJINESU RO-JA-NURU 51, n. 1 (2016) (Japan). The difference is less significant when compared to court-sponsored civil conciliation. There were 55,855 filings in district and summary courts in 2012 for civil conciliation. *Id.*

234 See MINISTRY OF JUSTICE, *supra* note 232.

235 See *supra* Part IV.C. (Table 1 Numbers of Mediated Securities Disputes Involving Banks & Table 2 Numbers of Mediated Securities Disputes Involving Dealer-Brokers).

236 Early in 2011, Japanese newspapers reported that the FSA requested major Japanese banks to take effective measures to protect their customer from bankruptcy. See Kenji Shimizu, *Kawase Deribatibu: Endaka Shinko de Tagaku Sonshitsu Chuusho Kigyo ni Ote-Ko Yushi Kin'yuchō Yosei* [Foreign Exchange Derivatives

small-to-medium businesses to enter into U.S. dollar swap contracts and other financial derivatives.²³⁷ They did so reportedly as part of their efforts to adapt to the Takenaka Plan, the policy package implemented in 2002 by then Minister of Financial Services Heizou Takenaka to restructure Japanese banks.²³⁸ Commentators suggest that the government's regulatory initiatives to increase the financial strength and profitability of Japanese banks caused the banks to seek short-term profits from fees arising from complex derivative transactions.²³⁹

These complex transactions turned into disputes following the financial crisis. In early 2010, many disputes involved long-term currency option contracts.²⁴⁰ Most contracts, at the time, were designed to benefit the customer only if the US dollar was higher than a fixed amount, and they gave banks leveraged gains in times of a weak US dollar.²⁴¹ During the financial crisis, the dollar dropped

Giving Huge Loss: Major Banks Ready to Give Bale-Out Finance to Small-to-Medium Businesses Prompted by Request from FSA], MAINICHI SHINBUN, morning ed., Jan. 20, 2011, at 6 (Japan). This request and response coincided with the sharp rise in filings with the JBA and FINMAC. Shimazu also reported that banks considered compensating their customers via JBA mediation. *Id.* at 6.

237 *Id.*

238 See Press Release, Fin. Servs. Agency, Keizai Zaisei Un'ei to Kouzou Kaikaku ni Kan suru Kihon Houshin (2002), <http://www.fsa.go.jp/news/newsj/13/kinyu/f-20020625-1.html> (Japan).

239 Toru Fujimori, *En'yasu-Tosan wo Fusegenai Ginko no Kawase Deribatibu no Torauma [Traumatic Memories of Foreign Exchange Derivatives Prevent Banks from Saving their Customers from Bankruptcy under Weak Yen]*, 4368 SHUKAN EKONOMISUTO 46, 47 (2014) (Japan); Yoshio Shima, *Chuusho Kigyo no Kawase Deribatibu Mondai to Toshika Hogo ni Kansuru Kosatsu [Discussion about Foreign Exchange Derivatives Issues for Small-to-Medium Businesses and Investor Protection]*, 18 RONSO (TAMAGAWA-DAIGAKU KEIEI-GAKUBU KIYO) 15, 24–29 (2012) (Japan); *Chuusho-Kigyo no Hatan Zoka ha Hisshi: Ginko ga Hameta Kawase Deribatibu no Wana [Rise of Bankruptcy of Small-to-Medium Businesses Inevitable: Banks Lured Them into a Mousetrap]*, DAIYAMONDO ONRAIN (2010), <http://diamond.jp/articles/-/10258?page=2> (Japan).

240 Kota Funahashi, *Deribatibu Tosan, Kyuzo Endaka de Sason Kin'yuchō, Ginko wo Chosa [Rapid Rise of Bankruptcy Because of Derivatives, Causing Loss in Foreign Exchange under Strong Yen: FSA Investigates Banks]*, ASAHI SHINBUN, morning ed., Dec. 12, 2010, at 1 (Japan); Ogata, *supra* note 217, at 44.

241 For examples of the contract structures disputed in civil litigations, see Toru Usui, *Tuka-Opushon-Torihiki no Hanbai-Kan'yu ni kansuru Setumei-Gimu-Ihan Tekigosei-Ihan to Tokyo Chihan Hei 25.2.22. [Breach of Duty to Explain and Suitability Principle in Solicitation of Currency Option Trades: Case Note on Tokyo District Court, Feb. 22, 2013]*, 1984 KIN'YU-HOMU-JIJO 115, 115–117 (Japan). See also Fujimori, *supra* note 239; Ogata, *supra* note 217.

and the yen gained more than sixty percent.²⁴² Japanese small-to-medium sized businesses are estimated to have lost 140 billion yen as a result.²⁴³

While regulators, consumer advocates, and academics had pushed for a comprehensive ADR system for financial transactions for a decade,²⁴⁴ the 2007–2008 financial crisis pushed a codified Financial ADR proposal through the Diet, overcoming opposition from Japanese businesses.²⁴⁵ And the new Financial ADR System

-
- 242 According to Bank of Japan statistics for the Tokyo central foreign exchange market, the US dollar fell from ¥123.88 on June 25, 2007 to ¥75.86 on October 28, 2011. *Shuyou Jikeiretsu Toukei Deta Hyou [Principal Times Series Statistical Data Table]*, BANK OF JAPAN, http://www.stat-search.boj.or.jp/ssi/mtshtml/fm08_d_1.html (last visited Apr. 13, 2018).
- 243 In September 2010, there were 40,500 foreign exchange derivatives contracts in effect between Japanese banks and some 19,000 businesses. See *Chuushou-Kigyo-Muke Kawase-Deribatibu Torihiki-Jokyo (Bei-Doru/Yen) ni kansuru Chosa no Kekka ni Tsuite (Sokuho-Chi)* [Interim Report of the Result of Investigation over Foreign exchange derivatives (U.S. Dollar/Yen) Sold to Small-to-Medium Businesses] (Mar. 11, 2011), FIN. SERVS. AGENCY, <http://www.fsa.go.jp/news/22/ginkou/20110311-2.html> (Japan). Most generated losses for the investor. *Id.* The 25,000 contracts sold by the major banks yielded the total net loss referenced above. *Id.*
- 244 Following the Financial System Council (*Kin'yū-Shingikai*) advisory seeking to streamline and strengthen complaint processing and dispute resolution in financial industries, the Financial ADR Liaison Meeting (*Kin'yū Toraburu Renrakuchōsei Kyōgikai*) began working on the financial ADR system in September 2000. See FIN. SERVS. AGENCY, SUMMARY OF THE 1ST FINANCIAL TROUBLE COORDINATION COUNSEL MEETING, (2000), http://www.fsa.go.jp/singi/singi_trouble/gijiyosi/f-20000925-1.html (Japan). It published its Model Rules for Financial ADR in April 2004, but they gained little support. See *Kin'yuu Bunya no Gyokai-Dantai Jishukisei-Kikan ni okeru Kujo Funso Kaiketsu-Shien no Moderu* [Model for Complaint and Dispute Resolution Assistance by Industry Association and Self-Regulatory Organization], FIN. SERVS. AGENCY (Apr. 25, 2014), http://www.fsa.go.jp/singi/singi_trouble/houkoku/f-20020425a_hou.pdf (last visited Jan. 7, 2017) (Japan).
- 245 See Yasano Kaoru Hatsugen, *supra* note 41; *Kin'yuu Bunya ni okeru Saibangai no Kuujou/Funsou Kaiketsu Shiten Seidou (Kinnyuu ADR) no Seibi ni Kakaru Kongou no Kadai ni tsuite (Zachou Memo)* [Future issues related to the improvement of financial complaints and dispute resolution support system (financial ADR) in the financial field (Chairperson Memo)], FIN. SERVS. AGENCY (July 1, 2008), http://www.fsa.go.jp/singi/singi_trouble/houkoku/20080624.html (Japan). Even at this stage, most SROs remained opposed to a statutorily-based ADR scheme. See Dai 35 Kai Kyoungikai ni okeru Giron wo Fumaeta Kin'yuu ADR no Houkousei ni Kansuru Gyokai Dantai no Iken [Based on the Discussion at the 35th Council Industry Organization's Opinion on the Direction of Financial ADR], *supra* note 155. The global financial crisis similarly caused Hong Kong regulators to create a Financial Dispute Resolution Center. Shahla F. Ali &

met an immediate demand.

Second, the financial DDROs cover a broader range of applicants. Unlike earlier industry association-sponsored mediation programs, DDROs provide mediation services to non-member businesses in the service category it covers if the business signs the Basic Contract and pays dues established in the DDRO's operational rules.²⁴⁶ Businesses do not have to pay the association membership fees in order to take advantage of this service; they are instead charged a yearly fee by the DDRO.²⁴⁷ This explains, in part, the sharp increase in mediation filings with JBA's Consultation Office on its designation as a DDRO.²⁴⁸

The bottom line is that increased demand following the 2007–2008 financial crisis resulted in FINMAC and the JBA Consultation Office processing large numbers of claims.²⁴⁹

V. IMPLICATIONS FOR THE RULE OF LAW IN JAPAN

At first glance, the establishment of Japan's Financial ADR, and its successful start-up, appear to be another example of what scholars have articulated as a consistent Japanese response to conflict—the creation of institutionalized mediation that removes the dispute from the courts.²⁵⁰

The establishment of Financial ADR, however, differs from other ADR mechanisms formed in response to earlier social crises. Financial ADR does not exclude the involvement of the courts and legal professionals. Rather, the Financial ADR System has expanded

Antonio Da Roza, *supra* note 80, at 485 (“In the aftermath of the financial crisis in 2008, Hong Kong’s regulators proposed to create a Financial Dispute Resolution Center (“FDRC”) for the purposes of deploying and centralizing alternative dispute resolution techniques in Hong Kong’s financial markets.”).

246 FIEA, *supra* note 23, at art. 156-44(5)(ii) (Japan). The dues established must not be “extremely unreasonable.” *Id.* FINMAC describes a goal of its establishment as “individually targeting Type II Financial Instruments Business Operators which have not joined in any of the five associations.” FINMAC, PURPOSE OF ESTABLISHMENT OF FINMAC, *supra* note 133, at 1.

247 See, e.g., FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 6-2.

248 Masao Tsuji, a member of the secretariat staff of the JBA, predicted the increased caseload, partly caused by expansion of its jurisdiction to non-member banks. Masao Tsuji, *Ginkō-Torihiki-Sōdanjo Assen-Inkai [Efforts on Dispute Resolution and Preparation for Designation as SROs: Japanese Bankers Association]*, 1877 KINYŪ-HŌMU-JIJŌ 66 (2010) (Japan).

249 See *supra* notes 224-229.

250 See, e.g., Upham, *supra* note 1; Haley, *supra* note 1; Takao Tanase, *The Management of Disputes: Automobile Accident Compensation in Japan*, 24 LAW SOC'Y REV. 651, 655 (1990).

the domain of the law, lawyers, and legal remedies. Each is discussed in turn.

A. *The Role of Law*

As described in earlier scholarship, Japanese courts have played an important role in the development of legal norms regulating securities fraud.²⁵¹ That continues today. While DDROs may have surpassed the court system in the volume of disputes processed, it is the courts that continue to establish the legal norms used to resolve financial service disputes.²⁵²

As set out above, there is no mandatory mediation or arbitration. Practitioners report that FINMAC officials will, during the consultation process, encourage investors claiming damages in excess of ¥10 million (\$100,000) to hire an attorney and file suit instead of mediate.²⁵³ The Chair of the JBA's Mediation Committee has stated that the civil litigation system works well and it is unwise for the JBA's ADR system to seek to resolve all disputes involving financial products offered by the banks.²⁵⁴ The former chair of the Japan Financial ADR/Ombudsman Study Group states that disputes exceeding a certain monetary threshold should be resolved through the courts and advocates excluding them from the Financial ADR System.²⁵⁵

At this point, the customer chooses to pursue DDRO-sponsored alternative dispute resolution, or not. Within each DDRO, if the claim cannot be resolved through the complaint process, the customer may initiate mediation proceedings unilaterally, but the financial service provider must obtain the consent of the customer to do so.²⁵⁶ If FINMAC cannot confirm the consent of the customer, they must dismiss the petition, and the customer may withdraw their consent to mediation at any time.²⁵⁷ While a customer may withdraw a motion for mediation at any time, a business may not file a motion to terminate mediation unless the customer agrees in

251 See Pardieck, *supra* note 17. For the sake of convenience we assume an exchange rate of ¥100 to \$1 in this paper.

252 See *infra* notes 237 – 270.

253 See E-mail from Osaka attorney to authors, *supra* note 8.

254 Tanaka, *supra* note 9.

255 Yanase, *supra* note 2, at 79.

256 FINMAC, *OPERATIONAL RULES*, *supra* note 60, at art. 26 para. 4; JBA, Operational Rules, *supra* note 65, at arts 26.

257 FINMAC, *OPERATIONAL RULES*, *supra* note 60, at art. 26 paras. 5-6; JBA, Operational Rules, *supra* note 65, art. 32(1).

writing.²⁵⁸ The DDRO rules are set up to ensure that mediation is voluntary and the investor retains the option to file suit.

As a result, the courts continue to develop legal norms. The derivatives disputes again provide an example. Although the Financial System Reform Act of 1998 allowed Japanese banks to begin over-the-counter sales of derivatives, it was not until 2004 that banks systematically promoted such contracts to their customers.²⁵⁹ When they went south following the 2007–2008 financial crisis, customers sued, and the courts confronted new bases for alleging violation of the suitability doctrine and duty to explain: plaintiffs began alleging an abuse of a dominant position by the banks over plaintiff-customers who were reliant on them for financing.²⁶⁰ Courts have rejected these arguments in some cases, finding facts supporting no more than the plaintiff's desire to maintain a good working relationship with the bank;²⁶¹ regulators have acted in other cases banning financial institutions from selling financial derivatives after finding that they forced corporate borrowers “to buy interest rate swaps to get loans.”²⁶²

The result is that the courts continue to resolve financial services disputes and establish legal norms. They do so in the larger cases and in those cases addressing novel issues of law or fact. Once the law is well established, the DDROs then apply it in routine,

258 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 39 paras. 1, 3; JBA, Operational Rules, *supra* note 65, at art. 32 (4).

259 Kin'yuu shisutemu kaikaku no tame no kankei houritsu no seibi nado ni kansuru houritsu [Laws related to the Development of Laws Regarding Financial System Reform], Law No. 107 of 1998, available at Shugiin, Dai 142 kai Kokkai Seitei Houritsu no Ichiran [House of Representatives, Session No. 142 Listing of Statutes], available at http://www.shugiin.go.jp/internet/itdb_housei.nsf/html/housei/h142107.htm (Japan).

260 See, e.g., Tokyo Chisai [Tokyo Dist. Ct.] Feb. 22, 2013, Dai (wa) no. 18501, 1420 KIN'YUU SHOUJI HANREI 40 (Japan). See also Toru Usui, *Tsuuka Opushion Torihiki no Hanbai Kanyuu ni Kan Suru Setsumei Gimu, Tekigousei Gensoku Ihan nado—Tokyo Chihan Hei* 25/2/22, 1984 KIN'YUU HOUMU JIJOU 114 (2013) (Japan).

261 Usui, *supra* note 260, at 120.

262 SMBC to be suspended from derivatives selling, JAPAN TIMES (Online Version), Apr. 28, 2006, <http://www.japantimes.co.jp/news/2006/04/28/business/smbc-to-be-suspended-from-derivatives-selling/#.V2hZqPnhA2x>. See also Kantoukujou no Hyouka Kamoku to Shotetsuzuki (Touroku Kin'yuu Kikan) [Supervisor's evaluation items and various procedures (Registered financial institution)], FIN. SERVS. AGENCY, <http://www.fsa.go.jp/common/law/guide/kinyushohin/08.html> (Japan).

smaller cases.²⁶³

The statutes establishing the Financial ADR System do not establish the norms to be applied in rendering a decision.²⁶⁴ Legal arguments, however, have come to define the disputes. Both FINMAC and the JBA systematically research and educate their mediators and staff on recent judicial decisions.²⁶⁵ The FINMAC secretariat edits an annual report of judicial decisions involving suitability and the duty to explain that it distributes to all mediators.²⁶⁶ They also provide mediators reports from mediated cases.²⁶⁷ The JBA, similarly, shares related judicial precedents and discusses “hard” cases in its programs for mediators.²⁶⁸

Interviews with DDRO officials offer an example of the influence of the courts. A Supreme Court decision from 2013 addressing the risks associated with interest rate swap transactions imposes a duty to explain that is defined by the characteristics of the individual investor.²⁶⁹ According to JBA officials, this decision does not dictate the result in a mediation, but it offers a framework for the mediators to inquire into the investor’s background and abilities during the mediation and resolve the matter based on that analysis.²⁷⁰ FINMAC officials similarly acknowledge that its mediators incorporate the comparative fault jurisprudence of the courts in their assessment of cases.²⁷¹

The Financial Services Agency has stated as its goal that the Financial ADR System “*based on the legal framework*, resolve disputes

263 *Report of Experts on Financial ADR System*, *supra* note 4, at 4. Tanaka, *supra* note 9, at 5.

264 Yanase, *supra* note 2, at 40.

265 *Soudan'in no Kenshuu oyobi Funsoukaiketsu-In no Jouhou Kyouyuu tou-no Joukyou* [Status of training of counselor and information sharing of dispute resolution committee members], FIN. SERVS. AGENCY (May 29, 2014), http://www.fsa.go.jp/singi/singi_trouble/siryou/20140529/04.pdf (Japan). FIN. SERVS. AGENCY, *supra* note 102, at 3, 145.

266 FIN. SERVS. AGENCY, *supra* note 102, at 145.

267 *Id.* at 150. See also Yanase, *supra* note 2, at 94 (“[T]he content of individual cases and the resolution provided by designated dispute resolution institutions are, even if names of individual disputing parties are not publicized, expected to play an important role as precedents.”).

268 FIN. SERVS. AGENCY, *supra* note 102, at 6.

269 Saikō Saibansho [Sup. Ct.] Mar. 7, 2013, 151 SAIKŌ SAIBANSHO MINJI HANREISHŪ [MINSHŪ] 243 (2013) (Japan).

270 E-mail from JBA Secretariat to authors, *supra* note 93.

271 E-mail from Fin. Instruments Mediation Assistance Ctr. Secretariat to authors (Mar. 24, 2017) [hereinafter E-mail from FINMAC Secretariat to authors] (on file with authors).

in a manner the parties feel acceptable.”²⁷² Members of the plaintiff’s bar describe experiences that confirm reliance on judicial norms: “if the courts would not find illegality, the financial institutions will not agree to settle.”²⁷³ Members of the defense bar have stated that preparing for ADR is like preparing for court, you gather evidence and you prepare “legal arguments.”²⁷⁴

The financial institutions focus on legal norms for a variety of reasons. A finding of “problematic behavior,” as required by the FIEA, invites internal disciplinary action by the financial institution against the employee(s) involved. This disciplinary action, in turn, is subject to challenge by the employee, and the financial institution looks for a clearly articulated legal basis to support the action.²⁷⁵ Similarly, financial institutions concerned about derivative actions by stockholders for “illegal expenditures” couch arguments in terms of clearly articulated legal standards justifying payments related to loss compensation.²⁷⁶

With representation by counsel common among business claimants, their arguments are formulated according to judicially recognized standards.²⁷⁷ Even without counsel, the DDROs organize individual customers’ arguments based on legal theory.²⁷⁸ The anonymized reports, for example, list the duty to explain and duty to refrain from soliciting unsuitable transactions as principal points for

272 *Kin’yuu ADR Seido Ni Tsuite (Sankou)*, *supra* note 47 (emphasis added).

273 E-mail from Tokyo attorney to authors, *supra* note 121.

274 See Tanabe, *supra* note 233, at 55. At the same time, they acknowledge that unlike the courts handing down a decision based on established law, ADR enables resolution by a subject matter specialist based on more flexible examination. *Id.*

275 Mikio Kanai & Kentaro Tani, *Zadankai-Kin’yuu ADR no Genzai [Roundtable Meeting – Finance ADR]*, 1946 KIN’YUU HOUMU JIJOU 6, 21 (Japan).

276 Toru Ushui, et al., *Zadankai-Kin’yuu ADR no Genzai*, 1946 KIN’YUU HOUMU JIJOU 6, 21 (2012) (Japan). For an example of a derivative lawsuit arising from loss compensation, see Saikō Saibansho [Sup.Ct.] July 7, 2000, 1767 SAIKŌ SAIBANSHO MINJI HANREISHU [MINSHU] 6, 54 (Japan) (an English summary is available from http://www.courts.go.jp/app/hanrei_en/detail?id=492). Note that the case involved loss compensation before the statutory ban, which necessitated plaintiffs basing their claim on antitrust law. *Id.*

277 Toru Usui, et al., *supra* note 276, at 15–17. The arguments cited in the anonymized case reports also illustrate this. See Japanese Bankers Ass’n, *Outlines and Outcomes of Mediated Cases*, *supra* note 219 (Japan); *Funso-Kaiketsu-Jirei (Shihanki-goto) [Quarterly Report of Mediation Cases]*, FINMAC, http://www.finmac.or.jp/tokei-siryo/index_06.html (last visited Jan. 7, 2017) (Japan).

278 FINMAC, *Quarterly Reports*, *supra* note 277.

most cases alleging improper solicitation.²⁷⁹ Although it is unclear who is responsible for this “legal” sorting, it is likely that DDRO secretariat staff and mediators sort and reconstruct the customers’ cases with reference to legal standards.²⁸⁰

Practitioners, regulatory agencies, and the DDROs all indicate that the legal norms articulated in judicial decisions guide settlement through ADR, playing a foundational role in resolving financial service disputes.

At the same time, the “rule of law” applied in the context of Financial ADR can differ from that found in the courts. The former Chair of the Japan Financial ADR/Ombudsman Study Group has stated his expectation that the adjudicators “will not depart from the law in many cases,” but, at the same time, some cases may warrant “flexible settlement proposals which would be impossible in courts.”²⁸¹ He suggests that mediators “may depart from the application of statutory laws or judicial precedents” in order to apply *jouri*, which he defines as the “common sense prevailing in society.”²⁸²

The JBA Secretariat states that mediators may evaluate cases based not only on established legal norms but also on business codes of conduct or industry practice.²⁸³ Practitioners have indicated that in currency swap disputes, mediators at both JBA and FINMAC may

279 See *Kakushū Tōkei-Shiryō* [Statistical Reports], FINMAC, <http://www.finmac.or.jp/tokei-siryo/> (last visited Jan 20, 2016) (Japan); *Funsō Kaiketsu-tō Gyōmu-no Jisshi Jōkyō* [Business Dispute Resolution Implementation Status], JBA, <http://www.zenginkyo.or.jp/abstract/adr/conditions/year/> (last visited Jan 20, 2016) (Japan).

280 During the complaint handling process, representatives of FINMAC hear directly from investors and clarify facts related to the transactional history. See FIN. SERVS. AGENCY, DAI 49 KAI KIN'YUU TORABURU RENRAKU CHOUSEI KYOUGIKAI: SHITEI FUNSO KAIKETSU KIKAN MUKE NO SOUGOU TEKI NA KANTOKU SHISHIN NADO HE NO TAIYOU JOUKYOU [THE 49TH FINANCIAL TROUBLESHOOTING COORDINATION COUNCIL: STATUS OF COMPLIANCE WITH COMPREHENSIVE SUPERVISION GUIDELINES FOR DESIGNATED DISPUTE RESOLUTION ORGANIZATIONS] 150-1 (2015), http://www.fsa.go.jp/singi/singi_trouble/siryou/20150615/05.pdf (Japan). Once the investor files for mediation, secretariat staff offer advice regarding drafting the complaint and which documents will be considered evidence. *Id.*

281 Yanase, *supra* note 2, at 83.

282 Yanase, *supra* note 2, at 49. In comparison, in the United Kingdom, its Financial Ombudsman decides cases using a “fair and reasonable” standard. Financial Services and Markets Act 2000, c. 8, § 228(2) (U.K.). *See also* Ali, *supra* note 41, at 345.

283 E-mail from JBA Secretariat to authors, *supra* note 93.

propose partial payment to the investor when they find the financial service provider's practice was imprudent, hasty, or aggressive, even if it was not illegal or would not constitute a tort.²⁸⁴

In other instances, mediators are reported to consider industry trends in evaluating conduct and proposing settlement. For example, it is common knowledge in the semi-conductor industry that there is a four-year business cycle, called the silicon cycle, and mediators are reported to have examined, in cases involving businesses in that industry, whether the banks appropriately considered this business cycle in soliciting foreign exchange derivative transactions.²⁸⁵

The JBA has also suggested that its mediators will consider the appropriateness of a bank's follow-up after it solicits a customer to invest in a certain financial product.²⁸⁶ While the law may not impose a duty to warn the investor to take action following a change in market conditions, many banks do, many customers expect the same, and a failure to do so may provide a basis for a mediated settlement.²⁸⁷ Other mediators have specifically found no violation of the law, yet proposed partial compensation (ten to thirty percent of the claimed damages) based on the broker's failure to follow good business practices and confirm the customer's actual understanding of the structure and risks of the solicited securities.²⁸⁸

284 TANAKA, *supra* note 102. See also Tatsushi Kishimoto, *Assen no Myomi?* [Charm of Mediation?], 14 FINMAC 6 (2014) (Japan) (article by practicing attorney serving as chair of the JBA's Mediation Committee).

285 E-mail from JBA Secretariat to authors, *supra* note 93.

286 *Id.*

287 See, e.g., JAPANESE BANKERS ASS'N, *ASSEN NO MOUSHITATE JIAN NO GAIYOU TO SONO KEKKA* (HEISEI 23 NENDO DAI 1 HANKI) TOUSHI SHINTAKU KANKEI [SUMMARY AND RESULTS OF THE COMPLAINTS FILED (FIRST QUARTER OF FY 2011) INVESTMENT TRUST RELATIONSHIP] (2011), http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/conditions/classify/classification_2301_2.pdf (Japan); E-mail from JBA Secretariat to authors, *supra* note 93.

288 FINMAC releases more detailed descriptions of its settlement proposals than the JBA. FINMAC quarterly case reports often cite cases where mediators clearly denied the brokers liability under tort law, but still recommended concessions, i.e. payment of compensation, pointing out their lack of confirmation of customers' actual understanding. FINMAC, *Quarterly Reports*, *supra* note 277. For specific examples of such recommendations, see Case Nos. 5, 12, and 16 from the report of DDRO mediations for the first quarter of 2013. FINMAC, *FINMAC FUNSOU KAIKETSU TETSUDUKI JIREI* (HEISEI 25-NEN 1-3 TSUKI) [FINMAC DISPUTE SETTLEMENT PROCEDURE (JAN. – MAR. 2013)] (2013), http://www.finmac.or.jp/tokei-siryo/pdf/sitei-assen/2013_01-03_conflict.pdf (Japan).

The DDRO rules permit a mediator to prepare a settlement proposal “in consideration of the equities of the two parties concerned.”²⁸⁹ And there is evidence that they have exercised this discretion in some cases to deviate from established legal standards.²⁹⁰ At the same time, it is clear that legal norms drive the process. They are clearly the starting point. The parties negotiate in the “shadow of the law”—in no small part because in Financial ADR legal professionals drive the process.²⁹¹

B. The Role of Lawyers

Another point of departure from earlier ADR processes is the continued involvement of attorneys in the proceedings. They serve both as mediators and advocates.²⁹² While past mediation and conciliation panels drew from the ranks of the important and influential,²⁹³ Financial ADR mediators are legal specialists and subject matter experts.²⁹⁴ Those tasked with creating a settlement plan and mediating a settlement are not exerting social pressure based on social status but instead asserting a legal judgment or, in some cases, consciously departing from a legal norm and relying on a business practice.

In some instances, attorney mediators are expressly required.²⁹⁵ All FINMAC mediations are conducted by a single mediator who must be an attorney.²⁹⁶ All JBA mediation panels must be chaired by an attorney.²⁹⁷ The FIEA and the Financial Services Agency Guidelines both expressly require the participation of at least one attorney on DDRO panels that include mediators who are officers of an industry association.²⁹⁸ The DDROs operational rules

289 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 40. See also JBA, Operating Rules, *supra* note 71, at art. 34.

290 TANAKA, *supra* note 102, at 35. E-mail from JBA Secretariat to authors, *supra* note 93.

291 See Robert Cooter, Stephen Marks & Robert Mnookin, *Bargaining in the Shadow of the Law: A Testable Model of Strategic Behavior*, 11 J. LEGAL STUD. 225 (1982).

292 See FSA, *Kin'yuu ADR Seido Ni Tsuite*, *supra* note 47.

293 See, e.g., Lawrence W. Beer, *Post War Law on Civil Liberties in Japan*, 2 UCLA PAC. BASIN L. J. 98, 106–07 (1983); Joel Rosch, *Institutionalizing Mediation: The Evolution of the Civil Liberties Bureau in Japan*, 21 L. & SOC'Y REV. 242, 248–49 (1987); Yanase, *supra* note 2, at 32.

294 See, e.g., FINMAC, *supra* note 157.

295 FIEA, *supra* note 23, at art. 156-44(4)(iv) (Japan).

296 FINMAC, OPERATIONAL RULES, *supra* note 60, at arts. 22 para. 2, 25.

297 JBA, Operating Rules, *supra* note 71, at art. 20(4).

298 FIEA, *supra* note 23, at art. 156-50(3) (Japan). FIN. SERVS. AGENCY, *Shitei*

must also establish “measures for receiving the advice of an attorney when the Dispute Resolution Mediator is not an attorney and resolution requires expert legal knowledge.”²⁹⁹ The implication is that attorneys serve to protect the impartiality of the panel and provide necessary information about relevant legal norms.

At the same time, attorneys play a role as advocates. As with older ADR programs in Japan, the DDROs are designed to enable application without representation.³⁰⁰ While the majority of individual claimants do appear pro se, the majority of business claimants do not.³⁰¹

The statistics available are limited, but FINMAC numbers for the year 2012 show that 88.5% of businesses pursuing claims were represented by counsel.³⁰² The JBA reports that, while almost all individual claimants appear pro se, approximately fifty percent of all the business claimants are represented by counsel.³⁰³ Statistics for all DDROs published by the Financial Services Agency’s Council of Advisors for the Financial ADR System showed in the first half of 2012, sixty-three percent of corporate customers retained counsel.³⁰⁴

The prevalence of counsel in mediations involving business customers may be explained by the businesses’ greater sophistication. It may also be a recognition of the complexity of the issues. As discussed above, most business claimants have sought compensation for damages arising from investment in derivatives, and the mediators have performed a transactional suitability analysis by examining hedge ratios, diversification, and other factors beyond the ken of

Funso Kaiketsu Kikan Muke no Sougou teki na Kantoku Shishin, *supra* note 40.

299 See, e.g., Banking Act, *supra* note 48, at art. 52-67(2)(iv) (Japan).

300 For example, the FSA, FINMAC and JBA websites provide simplified flowcharts and explanations, complete with cartoons, for lay persons to file complaints and initiate mediation. See FIN. SERVS. AGENCY, KIN’YUU KIKAN TONO AIDA DE TORABURU WO KAKAETE IRU RIYOU SHA NO MINNA SAMA HE [TO USERS WHO HAVE TROUBLE WITH FINANCIAL INSTITUTIONS] (2016), http://www.fsa.go.jp/policy/adr/adr_pamphlet.pdf (Japan); Soundan ando no Nagare, *supra* note 157 (Japan); JAPANESE BANKERS ASS’N, ZENKOKU GINKOU KYOUKAI SOUNDAN SHITSU [NATIONWIDE BANKERS ASSOCIATION CONSULTATION ROOM INFORMATION LEAFLET], http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/about/adr_leaf.pdf (Japan).

301 E-mail from FINMAC Secretariats to authors, *supra* note 127; E-mail from JBA Secretariat to authors, *supra* note 93.

302 Out of 173 individual claimants, 153 retained attorneys in FY 2012. E-mail from FINMAC Secretariats to authors, *supra* note 127.

303 E-mail from JBA Secretariat to authors, *supra* note 93.

304 Fin. Servs. Agency, *supra* note 106.

most lay persons. At the same time, the mediators are applying legal standards, developed piecemeal through judicial decisions, rather than a comprehensive set of administrative regulations to determine whether there was “illegal or wrongful conduct.”

All of this creates a need for legal and financial expertise, which some in the Japanese bar have developed and advertised.³⁰⁵ The Japan Federation of Bar Associations repealed its advertising ban in 2000.³⁰⁶ With the increased size of the Japanese bar, an increasing number advertise niche practices—many in financial products litigation.³⁰⁷ For close to a decade many of these attorneys focused on consumer loan litigation.³⁰⁸ With the running of the statute of limitations and the end of the consumer loan litigation, they have turned to securities litigation.³⁰⁹

This active involvement of attorneys and other legal specialists throughout DDRO proceedings again results in financial service cases being decided in the shadow of law.³¹⁰

305 See *infra* notes 307, 309.

306 See Eiichiro Kamizuma, ‘Bengoshi no Gyomu-Kokoku ni kansuru Kitei’ oyobi ‘Bengoshi oyobi Bengoshi-Hojin narabini Gaikoku-Tokubetsu-Kaiin no Gyomu-Kokoku ni kansuru Shishin’ no Kaisetu [Commentary on Rules for Advertisement by Attorneys and Guideline for Advertisement by Lawyers, Legal Professional Corporations, and Registered Foreign Lawyers], 14 Libra, no. 3, Mar. 2014, at 8, http://www.toben.or.jp/message/libra/pdf/2014_03/p02-29.pdf (Japan). The Japan Federation of Bar Associations imposed a total ban on advertisement from 1955 until 1987. *Id.* Its rule on advertisement in 1987 imposed strict restrictions on media and content, positioning advertisement as an exception rather than the rule. In 2000, it amended its articles of association (*kaisoku*), so that attorneys can advertise freely as long as the advertisement does not conflict with its Code of Ethics (*Bengoshi Gyomu Kitei*). *Id.*

307 For example, a web search of *kawase deribatibu*, currency exchange derivatives in Japanese, showed six law office webpages in its first 10 results. See search query of “*kawase deribatibu*”, GOOGLE SEARCH, <http://www.google.com> (enter “*kawase deribatibu*” in search field) (retrieved on Jan 15, 2016).

308 Saikō Saibansho [Sup.Ct.] Jan. 13, 2006, 60 SAIKŌ SAIBANSHO MINJI HANREISHU [MINSHU] 1 (Japan). See *Saiban no Jinsoku-Ka ni Kakaru Kensho ni Kansuru Hokokusho (Dai 3 Kai)* [3rd Report about Verification of Expediting of Court Procedures] 24–25, SAIBANSHO SAIBANREI JOHO [SAIBANSHO WEB], http://www.courts.go.jp/about/siryo/hokoku_03_hokokusyo/index.html (Japan).

309 See FIN. SERVS. AGENCY, *supra* note 300; FINMAC, *supra* note 300; JAPANESE BANKERS ASS’N, *supra* note 300 (showing examples of currency exchange derivative advertisements).

310 Scholars have referred to this as “Rights-based ADR.” Nottage, *Resolving Claims from the Fukushima Nuclear Disaster*, *supra* note 13. In this regard, the emphasis on legal professionals applying legal norms in Financial ADR is similar to that found at the Fukushima Dispute Resolution Centre:

C. The Role of Financial ADR

To argue that legal norms and lawyers play a pivotal role in Financial ADR, though, leaves unanswered a central question—is it fair? Are the remedies provided equitable according to some standard? There is evidence suggesting that they are—certainly not in every case—but on the whole.³¹¹

Because alternative dispute resolution is not required, the Financial ADR System supplements the civil justice system by providing speedy, inexpensive, and, by some measures, reasonable remedies for classes of investors unlikely to sue: those with limited time, limited resources, or relationships to preserve.

1. Time to Resolution

First, the Financial ADR System provides expedited relief, a factor acknowledged to be particularly important to elderly investors.³¹² FINMAC statistics for the years 2014 and 2015 suggest their mediation services are now predominately used by the elderly. In 2014, 40.8% of its individual claimants were over seventy-five years old.³¹³ In 2015, 31.9% were over the age of seventy-five and 66.9% over the age of sixty-five.³¹⁴ JBA reports for 2014 show that 30.2% of its individual users were over the age of seventy; 45.5% over the age of sixty.³¹⁵ In 2015, nineteen percent were over seventy; thirty-seven percent over sixty.³¹⁶

“[T]he Fukushima Dispute Resolution Centre process involves several novel features compared to earlier ADR schemes established with Japanese government support. In particular, it is operated by (private) lawyers, rather than primarily by government officials (as in environmental pollution or noise pollution cases). Secondly, the Centre applies guidelines and other principles derived from tort law, including an extensive accumulation of court judgments, rather than invoking non-legal norms based on the mediators’ (often “didactic”) notions of broader community expectations.” Nottage, *Resolving Claims from the Fukushima Nuclear Disaster*, *supra* note 13.

311 See discussion *supra* notes 295-310 and *infra* notes 312-366 and accompanying text.

312 See Tanaka, *supra* note 9, at 4.

313 *Id.* at 8.

314 *Id.*

315 JAPANESE BANKERS ASS’N, *Zenkoku Ginkou Kyoukai Funsou Kaiketsu nado Gyounmu no Jisshi Joukyou* 8 (2014), http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/conditions/conditions01_2600.pdf (Japan).

316 JAPANESE BANKERS ASS’N, *Zenkoku Ginkou Kyoukai Funsou Kaiketsu nado Gyounmu no Jisshi Joukyou* 8 (2015), http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/conditions/conditions01_2700.pdf (Japan).

The elderly in Japan have turned to Financial ADR, and they see a difference in time to resolution compared to the courts. As of 2015, the civil courts of first instance in Japan averaged 12.8 months from filing to disposition.³¹⁷ Mean time for mediation in 2014 was 2.1 months for FINMAC and 5.7 months for the JBA.³¹⁸ Financial Services Agency statistics for all the DDROs indicate that sixty-nine percent of all complaints filed in 2014 were resolved within three months of filing.³¹⁹ Seventy-three percent of all mediation claims filed with DDROs were resolved within six months of filing in 2014.³²⁰

Settlements within the Financial ADR System also facilitate enforcement and, in doing so, expedite compensation. While mediated settlements by private ADR organizations result in a settlement enforceable only as a breach of contract,³²¹ DDRO brokered settlements are subject to enforcement by the claimant³²² and by the DDRO, which can publish findings of noncompliance and report the same to the Prime Minister.³²³ Given the financial service provider's contractual obligations with the DDRO, and the DDROs reporting requirements to the FSA, the cost of noncompliance—

317 SUPREME COURT, SAIBANSHO DE-TA BUKKU 2016 [COURT STATISTICS 2016] 71, <http://www.courts.go.jp/about/databook/index.html> (Japan).

318 FIN. SERVS. AGENCY, *supra* note 102, at 1, 144.

319 FIN. SERVS. AGENCY, *Shitei Funsou Kaiketsu Kikan no Funsou Kaiketsu tou Gyoumu Jisshi Joukyou* 5 (2016), http://www.fsa.go.jp/singi/singi_trouble/siryuu/20160609/03.pdf (Japan) (stating that it was seventy-three percent in 2015). See also MATSUO NOBUKORE, KIN'YUU ADR NO HYOUKA TO KONGO NO KADAI 6 (2015), http://www.zenginkyo.or.jp/fileadmin/res/news/news280125_1.pdf (Japan).

320 FIN. SERVS. AGENCY, *supra* note 319, at 10. In 2014, 28% were resolved in less than 3 months and 45% in less than six. *Id.* In 2015, 27% were resolved in less than 3 months, and 45% in less than six. *Id.*

321 Tsuneo Osawa, *ADR wo Gaikansuru*, 24 BIJINESU RO-JA-NURU 27, 30 (2010) (Japan).

322 See Minpou [Civil Code], Law No. 89 of 1896, art. 696 (Japan). The parties can render a settlement agreement eligible for direct compulsory execution through notarization. Minjishikkouhou [Civil Execution Act], Law No. 4 of 1979, art. 22(v) (Japan); Tanabe, *supra* note 226 at 55; KAZUHIKO YAMAMOTO & FUMI YAMADA, *ADR Chuousai Hou [ADR and Arbitration Law]* 179 (2d ed. 2015) (Japan). The parties can also file their agreement with the court as a "Settlement Prior to the Filing of an Action," which is eligible for compulsory execution. See Minji Soshouhou [Code of Civil Procedure], Law No. 109 of 1996, arts. 275, 267 (Japan); Civil Execution Act, Law No. 4 of 1979, art. 22(vii) (Japan); YAMAMOTO & YAMADA, *supra* note 322, at 179.

323 See FIEA, *supra* note 23, at art. 156-45 (Japan); Osawa, *supra* note 321, at 30.

failure to timely pay a mediated settlement—is high.³²⁴

2. Cost of Resolution

The filing fees and costs associated with DDRO resolution also explain the use of the Financial ADR System. DDROs are either part of an industry association or sponsored by one.³²⁵ This financial sponsorship raises questions about impartiality, but it also enables the DDROs to provide complaint processing and mediation services at a reduced or no charge to the investor.³²⁶

The DDRO branch of the JBA imposes no fee on bank customers filing complaints or petitions for mediation.³²⁷ Instead, it charges a ¥100,000 (\$1,000) case fee on the member bank when an application for mediation is accepted.³²⁸ The fee is doubled if a non-member bank is involved.³²⁹

FINMAC charges its members a basic yearly fee of ¥100,000.³³⁰ In a dispute, it will charge its Type I securities company members a per day mediation user fee of ¥50,000 (\$500) and its Type 2 members, selling mutual funds and trusts, graduated fees depending on the number of complaints filed.³³¹

The FINMAC complainant must pay a filing fee, a “mediation

324 The JBA reports that to date, no bank has refused to comply with an agreed mediation plan, i.e., pay compensation based on the plan. JBA officials find it hard to imagine they would not follow their contractual obligations with the association. E-mail from JBA Secretariat to authors, *supra* note 93.

325 See FIN. SERVS. AGENCY, *Shitei Funso Kaiketsu Kikan Muke no Sougou teki na Kantoku Shishin*, *supra* note 40; FIEA, *supra* note 23, at art. 156-39(1) (Japan); Banking Act, *supra* note 48, at art. 52-62(1) (Japan).

326 Article 156-42(2) of the FIEA provides that the DDRO “may receive dues or fees, or any other remuneration for performing Dispute Resolution Services” from persons and firms with whom it has concluded a Basic Contract or made other contractual arrangements. FIEA, *supra* note 23, at art. 156-42(2) (Japan).

327 JBA, Operating Rules, *supra* note 71, at art. 28. See also JAPANESE BANKERS ASS'N, KUJOU SHORIT TETSUDUKI OYOBI FUNSOU KAIKETSU TETSUDUKI NO JISSHI NI KANSURU UN'EI YOURYOU [ADMINISTRATIVE PROCEDURES RELATING TO THE IMPLEMENTATION OF COMPLAINT RESOLUTION PROCEDURES AND DISPUTE RESOLUTION PROCEDURES] art. 17 <http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/rules/rule02.pdf> (Japan).

328 JBA, ADMINISTRATIVE PROCEDURES, *supra* note 327 at art. 17.

329 *Id.*

330 See FIEA, *supra* note 23, at art. 156-44(5) (Japan); FINMAC, ARTICLES OF ASSOCIATION, *supra* note 65, at art. 9; FINMAC, DETAILED RULES, *supra* note 79, at art. 4 para. 1.

331 FINMAC, OPERATIONAL RULES, *supra* note 60, at art. 6-3.

motion fee,” and those filing fees are graduated depending on the amount of damages claimed.³³² But FINMAC mediation filing fees are a fraction of court filing fees. For example, claims under ¥1 million (\$10,000) are assessed a ¥2,000 (\$20) fee.³³³ By comparison, a ¥1 million claim in civil court costs ¥10,000 (\$100).³³⁴ A ¥100 million (\$1 million) claim filed with FINMAC requires payment of a ¥17,000 (\$170) filing fee.³³⁵ The same claim if filed in civil court requires a ¥50,000 (\$500) filing fee.³³⁶ Compared to financial ADR services offered in some other countries, FINMAC costs are high.³³⁷ Yet, compared to filing suit in Japan, there are substantial savings.³³⁸

At the same time, some see, or perceive, savings in legal fees. The majority of individual investors file their complaints with DDROs pro se.³³⁹ Rightly or wrongly, they see DDRO mediation as

332 *Id.* at annex 2.

333 *Assen Moushitate no Ryoukin Hyou [Fee Schedule for Mediation Petitions]*, FINMAC <http://www.finmac.or.jp/ryokin/> (last visited Aug. 5, 2016) (Japan).

334 *Saikō Saibansho* [Sup. Ct.], *Tesuuryougaku Hayami Hyou [Simplified Fee Schedule]*, http://www.courts.go.jp/vcms_lf/315004.pdf (last visited Aug. 5, 2016) (Japan). The comparison is not exact, as the Supreme Court sets fees using smaller increments. *Id.* Fees start with a 1,000 yen filing fee for claims up to 100,000 yen, with an additional 1,000 yen charged per increase of 100,000 in damages claimed. *Id.*

335 FINMAC, *supra* note 333.

336 *Saikou Saibansho* [Sup. Ct.], *Tesuuryougaku Hayami Hyou [Simplified Fee Schedule]*, *supra* note 334. Filing a 500 million yen or more claim with FINMAC necessitates payment of 50,000 yen; the same would require 170,000 yen in civil court. See FINMAC, *Ryoukin Hyou [Fee Schedule for Mediation Petitions]*, *supra* note 333; *Saikou Saibansho* [Sup. Ct.], *Tesuuryougaku Hayami Hyou [Simplified Fee Schedule]*, *supra* note 334.

337 In the United Kingdom and Australia, consumers can file complaints with the Financial Ombudsman Service free of charge. Ali, *supra* note 41, at 344. In Singapore, consumers can seek case management, i.e., claim processing or mediation, free of charge. *Id.* at 345. In comparison, in the United States, FINRA charges a customer a graduated filing fee depending on the amount of damages claimed, ranging from \$50 for a claim under \$1,000 to \$2,250 for a claim over \$5 million, graduated hearing fees, per session, ranging from \$50 to \$450 for a single arbitrator, along with various motion fees. *Summary of Arbitration Fees*, FINRA, <http://www.finra.org/arbitration-and-mediation/summary-arbitration-fees> (last visited Jan. 12, 2017). Estimated fees for arbitrating a claim of \$100,000 is \$975 for the filing fee, and \$1350, for a single day hearing. *Arbitration Fee Calculator*, FINRA, <http://apps.finra.org/ARbitrationMediation/ArbFeeCalc/1/Default.aspx> (last visited Jan. 12, 2017).

338 Compare FINMAC, *supra* note 333, with *Saikō Saibansho* [Sup. Ct.], *supra* note 334.

339 E-mail from JBA Secretariat to authors, *supra* note 93.

a place where they can resolve the claim themselves. As discussed above, the JBA reports almost all individual claimants appear pro se.³⁴⁰ Available statistics show only thirty percent of individual claimants represented by counsel in FINMAC, and only thirty-two percent of all claimants in all DDRO mediations represented by counsel.³⁴¹

Some of this is likely involuntary. Practitioners suggest that most investors that seek the advice of an attorney regarding investment losses come with losses in excess of several million yen.³⁴² Some attorneys find it difficult to take cases worth less than ¥1 million (\$10,000);³⁴³ others report that they look for damages in excess of ¥10 million (\$100,000) before pursuing civil litigation.³⁴⁴ At the same time, FINMAC statistics for 2015 indicate that 62.1% of the claims filed were for less than ¥5 million (\$50,000).³⁴⁵ In other words, the majority of claims currently filed with FINMAC are claims many attorneys would not pursue.

The Financial ADR System was intended to offer individual investors with small claims a means to timely resolve them.³⁴⁶ The numbers suggest that the types of claimants served by Japan's Financial ADR System and the size of the claims they bring has changed over time, but that the FINMAC and the JBA's DDROs routinely serve individual investors with small claims.

3. Adequacy of Relief

One might talk about how the Financial ADR System is typically faster at resolving a claim than the courts. One might talk about how resolving a dispute through the Financial ADR System is cheaper than going to court, and it may be the only option for some.

340 *Id.*

341 Fin. Servs. Agency, *supra* note 106.

342 E-mail from Tokyo attorney to authors, *supra* note 121.

343 *Id.*

344 E-mail from Osaka attorney to authors, *supra* note 8.

345 FINMAC, *Heisei 27 nendo no soudan, kujou, assen no tokuchou ni tsuite* [About Consultation, Complaints, Characteristics of Annuities in FY2015] 7 (2015), available at https://www.finmac.or.jp/tokei-siryo/pdf/highlight/highlight_2015.pdf (Japan). FINMAC categorizes its claims by those seeking damages under 10 million yen, those between 10 and 50 million yen, those between 50 and 100 million yen, and those between 100 million and 500 million yen. *Id.*

346 Kazuhiko Yamamoto, *Kin'yuu ADR no Hyouka to Kongo no Kadai*, *Zenkinkyou ADR5 Shūnen Symposium* [JBA Fifth Annual ADR Symposium] 7 (Nov. 13, 2015), available at http://www.zenginkyo.or.jp/fileadmin/res/news/news280125_1.pdf (Japan).

But there remains the basic concern: is it fair?

DDRO officials acknowledge public concerns that the mediators may be biased in favor of the financial service providers but suggest the practical reality is different.³⁴⁷ Consumer advocates, including the Chief Consultant for the Tokyo Metropolitan Consumer Center has described DDRO mediation services, at least as provided by the JBA, as unbiased and fair.³⁴⁸

a. Compensation & Comparative Fault

A more objective measure of fairness, and criticism, may be the amount awarded in mediation. Some plaintiffs' attorneys state flatly that they will explain FINMAC and the JBA's ADR systems to their clients, but they will not recommend them.³⁴⁹ Others indicate that they mediate only if their clients express a strong desire to avoid litigation because of time constraints or concerns about the stress of examination at trial.³⁵⁰ All express concerns about the inadequacy of the damages awarded.³⁵¹ They indicate that financial service providers rarely agree to settlement proposals exceeding fifty percent of the claimed damages.³⁵² As a result, mediators, whose job it is to settle cases, routinely propose damage awards of less than half the claimed damages, tending to propose settlements of approximately thirty percent.³⁵³

At first blush, these amounts seem patently unfair—and they may be—but they are not irrational. Japanese courts are aggressive in their comparative fault determinations.³⁵⁴ For litigants large or small, courts emphasize self-responsibility (*jiko sekinin*).³⁵⁵

347 E-mail from JBA Secretariat to authors, *supra* note 93.

348 Hiyama, *Kin'yuu ADR no Hyouka to Kongo no Kadai*, ZenkinkyouADR5Shuunen Symposium 7 (Nov. 13, 2015), http://www.zenginkyo.or.jp/fileadmin/res/news/news280125_1.pdf (Japan).

349 E-mail from Tokyo attorney to authors, *supra* note 121.

350 E-mail from Osaka attorney to authors, *supra* note 8.

351 E-mail from Tokyo attorney to authors, *supra* note 121. E-mail from Osaka attorney to authors, *supra* note 8.

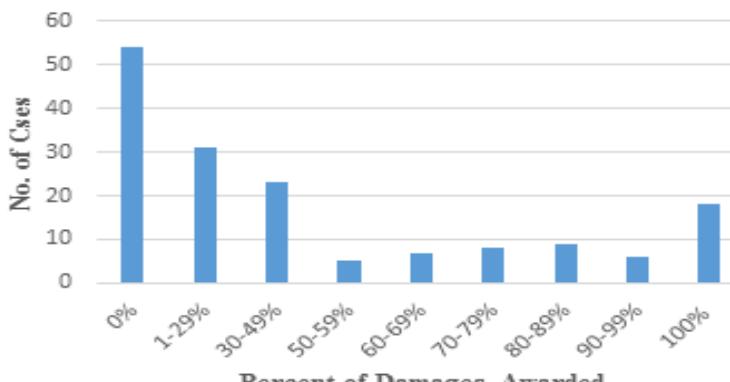
352 E-mail from Tokyo attorney to authors, *supra* note 121. E-mail from Osaka attorney to authors, *supra* note 8.

353 E-mail from Osaka attorney to authors, *supra* note 8.

354 See *supra* notes 336-345.

355 See, e.g., Katsunari Iwasaki, *Koi-Fuhokoi ni okeru Kashitsu-Sosai ni tsuite* [On Comparative Fault in Intentional Torts], 42-2 HOSEI-RIRON 52, 55-57 (2009) (Japan) (comparative fault is applied both to breach of contract and tort, and the latter includes negligence and intentional tort).

Broker-Dealer Cases 1998-2008



Analysis of a ten-year period spanning the collapse of the bubble economy through the beginning of the 2007–2008 financial crisis show courts awarding no damages to plaintiffs in 33.5% of all cases filed against broker-dealers; nineteen percent awarding damages of less than thirty percent; and fourteen percent awarding damages between thirty and forty-nine percent.³⁵⁶ In other words, in over half of the cases litigated, plaintiffs ended with damages of less than thirty percent; in over two-thirds, plaintiffs ended in a damage award of less than half the claimed damages.

Recent compilations of reported securities fraud cases concur that Japanese courts aggressively discount awards to investors based on comparative fault (*kashitsu sousai*).³⁵⁷ The Japanese Federation of Bar Associations' 2015 GUIDEBOOK ON REMEDIES FOR FINANCIAL TRANSACTION DAMAGES continues to cite to the judicial findings of seventy to eighty percent comparative fault found in the early securities cases.³⁵⁸ Other scholars identify comparative fault ranges

³⁵⁶ MAKOTO IKEYA ET AL., SHOKEN-SOSHNO KEIZAI-BUNSEKI: NICHIBEI NO JIREI DOKO TO SONGAI RISSHO APURUCHI [ECONOMIC ANALYSIS OF SECURITIES LITIGATIONS: ANALYSIS OF JAPANESE AND AMERICAN TREND AND APPROACH FROM PROVING DAMAGE] 80–84 (2008) (Japan).

³⁵⁷ NIHON-BENGOSHI-RENGO-KAI SHOHISHA-MONDAI TAISAKU IINKAI [CONSUMER AFFAIRS COMMITTEE OF JAPAN FEDERATION OF BAR ASS'NS], *supra* note 227, at 142–149.

³⁵⁸ *Id.*

from forty to seventy percent in securities and futures cases.³⁵⁹ Even victims of pyramid schemes routinely see awards discounted by forty to fifty percent for comparative fault.³⁶⁰

Courts are reported to be even more aggressive in their comparative fault determinations in litigation brought by businesses.³⁶¹ Practitioner interviews suggest that most cases where businesses seek damages for misrepresentation or unsuitability end in a court finding for the financial institution and no damages awarded.³⁶² The courts tend to see business actors as responsible for their own losses, and “outside of the scope of the intended protection of the law.”³⁶³

The reported judicial decisions and interviews with the JBA Secretariat regarding foreign exchange derivatives mediation support this assessment.³⁶⁴ The courts have focused on the business managers’ investment experience and “contract responsibility” in rejecting claims for damages.³⁶⁵ The JBA Secretariat, in contrast,

³⁵⁹ See Imagawa, *supra* note 166. Scholars have also found that judges have differed in their decisions in similar cases and judicial outcomes are highly unpredictable in currency swap cases. Hiroko Aoki, *Sonshitsu-Hoten-Kinshi-Gensoku no Haishi ni tsuite* [An Argument for Abolition of the Statutory Ban of Loss Compensation], *KIN'YU-SHOHIN-TORIHIKI-HOSEI NO CHORYU* [THE TRENDS OF FINANCIAL INSTRUMENTS LAW] 232 (Nihon-Shoken-Keizai-Kenkyujo, 2015) (Japan). Aoki, one such scholar, argues for settled outcomes of such disputes to increase predictability. *Id.* In a practical guidebook for attorneys representing investors, the authors highlight the unpredictability and ‘occasional abuse’ of comparative fault jurisprudence by Japanese courts in securities fraud cases as one of the biggest challenges for the investor’s bar. Although they cited some judicial decisions denying comparative faults for reference, they affirmed the unpredictable nature of comparative fault adjudication. *KAITEI Q&A TOSHI-TORIHIKI-HIGAI KYUSAI NO JITSUMU* [QUESTION AND ANSWER: PRACTICAL GUIDE TO REMEDIES FOR DAMAGES FROM INVESTMENT] 224-8 (Tetsuo Arai ed., 2015) (Japan).

³⁶⁰ Iwasaki, *supra* note 355, at 58–64.

³⁶¹ E-mail from Tokyo attorney to authors, *supra* note 121.

³⁶² *Id.*

³⁶³ E-mail from Osaka attorney to authors, *supra* note 8.

³⁶⁴ E-mail from JBA Secretariat to authors, *supra* note 93.

³⁶⁵ See, e.g., Tokyo Chisai [Tokyo Dist. Ct.] Feb. 22, 2013, Dai (wa) no. 18501, 1420 *KIN'YUU SHOUJI HANREI* 40 (Japan). In that case, apparel manufacturers entered into yen-dollar derivative contracts in order to hedge currency risk associated with supply orders made to Chinese manufacturers. *Id.* One manufacturer sought mediation after losses of ¥86m. *Id.* The mediation panel found the bank failed to examine the volume of China-bound orders to determine the appropriate size of dollar-yen derivative contracts and proposed an award of a percentage of the damages. *Id.* The investor rejected the proposed mediation award finding the percentage too low and sued. *Id.* The

suggests its mediators focus more broadly on suitability, including the amount of the hedge contract and the ability to bear the risk of loss, and, in doing so, it has settled approximately eighty percent of these cases.³⁶⁶ Viewed in the light of the courts' aggressive comparative fault analysis, the limited Financial ADR awards appear more reasonable.³⁶⁷

Both DDRO officials and practitioners acknowledge that the courts are better suited to resolve certain types of cases: cases requiring a finding of certain facts, e.g., embezzlement; cases where the investor demands an apology; and cases where an investor challenges the legality of the financial product itself.³⁶⁸

JBA officials have also suggested that if the facts support a clear finding of illegality, such a finding by the court may result in a more favorable award to the investor.³⁶⁹ For example, a Kyoto District

court found the investor had experience investing; found no violation of the duty to explain or the suitability principle; and awarded no damages. *Id.* See also, e.g., Osaka Chisai [Osaka Dist. Ct.] Nov. 6, 2015, 1484 KIN'YUU SHOUJI HANREI [KINHAN] 37 (Japan). In that case, the Defendant sold currency options to a company with a foreign representative director with limited Japanese language ability. *Id.* The mediators recommended a partial award of damages, which the bank rejected. *Id.* The mediators followed with a special conciliation proposal, which the defendant financial institution rejected. *Id.* The financial institution filed suit, and the court found no violation of the duty to explain or suitability principles and awarded plaintiff no damages. *Id.* See also E-mail from JBA Secretariat to authors, *supra* note 93.

- 366 E-mail from JBA Secretariat to authors, *supra* note 93. Some practitioners have suggested that this largesse was temporary. E-mail from Tokyo attorney to authors, *supra* note 121. For a period of time during which the Diet and Financial Services Agency were focusing on losses incurred by investors purchasing foreign exchange derivatives, the JBA offered faster and more likely compensation than the courts, but after public attention waned, the benefits to mediating before the JBA disappeared. *Id.*
- 367 Direct involvement by a financial regulator in negotiating investor compensation presents a different but interesting point of comparison: Following the collapse of Lehman Brothers and the 2007-2008 financial crisis, regulators in both Hong Kong and Singapore directly negotiated with banks to coordinate settlement agreements related to investments in Lehman Brother's mini-bonds. Ali, *supra* note 41, at 346. This resulted in Hong Kong with settlements restoring approximately 85 to 95 percent of claimants' initial investment. *Id.* The Monetary Authority of Singapore's involvement resulted in investors recovering an average of 64.5 percent of their initial investment. *Id.*
- 368 Yamamoto, *supra* note 346, at 8–9. E-mail from JBA Secretariat to authors, *supra* note 93. E-mail from FINMAC Secretariats to authors, *supra* note 127.
- 369 E-mail from JBA Secretariat to authors, *supra* note 93.

Court awarded 100 percent of the damages sought by an investor with limited Japanese language skills for losses arising from the sale of mutual funds.³⁷⁰ Mediators have suggested it would be difficult to craft a mediation proposal requiring the bank to pay 100 percent of the damages.³⁷¹

Conversely, as discussed above, where courts have awarded no damages to investors finding no violation of the duty to explain or the suitability principle, mediators may propose a partial award based on a violation of an industry standard. And they have proposed partial awards where the financial institution failed to consider the investor's documented need to maintain funds to provide for an elderly parent or a lack of prior investor experience.³⁷²

The judicial decisions and practitioner interviews suggest that there are some cases where claimants are clearly better off filing suit.³⁷³ Those interviews combined with an analysis of the comparative fault numbers also suggest that the opposite may be true. Business claimants may be better off mediating, and most individuals will see settlements at least comparable to litigated outcomes.

b. Voluntary Participation & Settlement Rates

Given investors' voluntary participation in the Financial ADR System and the ease with which they may terminate the process and litigate, another measure of fairness may be settlement rates.

Reports on settlement rates vary. One scholar suggests that, for the period from 2010 to 2015, the JBA mediated 2,598 cases, settling 59.35% of them.³⁷⁴ Other reports suggest a recent decline in JBA settlement rates to approximately forty percent following a drop-off in the number of derivative cases filed.³⁷⁵ Our research shows JBA

370 Kyoto Chisai [Kyoto Dist. Ct.] Sept. 25, 2014, Heisei 25 (wa) no. 904, 1455 KIN'YUU SHOUJI HANREI [KINHAN] 28 (2014) (Japan).

371 E-mail from JBA Secretariat to authors, *supra* note 93. JBA officials have suggested the mediation is not about finding one party 100% at fault. *Id.* It is a matter of understanding fault on the part of the bank, acknowledging the customer's responsibility as a contracting party and seeking to craft a settlement proposal based on concession. *Id.*

372 Hiroshima Kousai [Hiroshima App. Ct.] June 14, 2012, 1387 HANREI TAIMUZU [HANTA] 230 (2012) (Japan).

373 See *supra* notes 330-351.

374 Tanaka, *supra* note 9. At the same time, JBA surveys vary, but show user satisfaction rates ranging from a simple majority to eighty percent. *Id.* at 3.

375 JAPANESE BANKERS ASS'N, ZENKOKU GINKO KYOKAI FUNSO-KAIKETU-

complaint resolution rates vary, ranging from twenty-nine to fifty-seven percent, and mediated settlement rates range from forty-six to fifty-nine percent.³⁷⁶ FINMAC, in turn, reports complaint resolution rates ranging from seventy-three percent to eighty-eight percent and mediation settlement rates ranging from forty to fifty-six percent for the years 2011 to 2015.³⁷⁷

In comparison, special conciliation proposals—those proposals by mediators forcing a financial service provider to accept or litigate—are used rarely. A 2015 Financial Supervisory Agency report shows FINMAC mediators having issued only two since inception, with one resulting in settlement, and the JBA issuing only six, with four resulting in settlement.³⁷⁸

FINMAC and JBA mediators rarely attempt to force a settlement, and when they do it rarely works.³⁷⁹ Yet their mediated settlement rates compare favorably with the courts. Approximately forty percent of all cases filed in the district courts result in judicial opinions; thirty-five percent of all cases filed are settled.³⁸⁰

Statistics for all DDROs show similarly high resolution rates. In 2014, the DDROs collectively resolved seventy-one percent of all complaints filed, and, in 2015, seventy-eight percent.³⁸¹ Mediated settlements rates for all DDROs are lower, but still similar to what

TO GYOMU NO JISSHI-JOKYO(HEISEI 28 NENDO DAI 2 SHIHANKI) [STATE OF IMPLEMENTATION OF DISPUTE RESOLUTION BY JBA (JUL. – SEP., 2016)], http://www.zenginkyo.or.jp/fileadmin/res/abstract/adr/conditions/conditions01_2802_1.pdf (Japan).

376 Statistics compiled from Fin. Servs. Agency, *Kin'yū toraburu renraku chōsei kyōgikai* [Financial Trouble Contact Adjustment] (Heisei 27 Nen 4 Gatsu 1 Nichi – Heisei 28 Nen 3 Gatsu 31 Nichi) and (Heisei 27 Nen 4 Gatsu 1 Nichi – Heisei 28 Nen 3 Gatsu 31 Nichi), available at http://www.fsa.go.jp/singi/singi_trouble/siryou (Japan).

377 *Id.*; Fin. Servs. Agency, *Dai 49 Kai Kin'yuu Toraburu Renraku Chousei Kyougikai: Shitei Funso Kaiketsu Kikan Muke no Sougou Teki na Kantoku Shishin nado he no Taiyou Joukyou* [The 49th Financial Trouble Contact Adjustment Council: Status of Compliance with Comprehensive Supervision Guidelines for Designated Dispute Resolution Organizations], 149–150 (June 15, 2015), available at http://www.fsa.go.jp/singi/singi_trouble/siryou/20150615/05.pdf (Japan).

378 Fin. Servs. Agency, Zaimu ADR Renraku Kaigi [Financial ADR Liaison Meeting], (June 15, 2015), http://www.fsa.go.jp/singi/singi_trouble/siryou/20150615/05.pdf (Japan).

379 *Id.*

380 Tanaka, *supra* note 9, at 4.

381 Fin. Servs. Agency, *Shitei Funsou Kaiketsu Kikan no Funsou Kaiketsu tou Gyōmu Jisshi Joukyou* [Designated Dispute Settlement Organization Status of Implementation] 4 (June 9, 2016), available at http://www.fsa.go.jp/singi/singi_trouble/siryou/20160609/03.pdf (Japan).

one finds in the courts. For all DDROs operating in Japan in 2015, thirty-nine percent of all cases filed settled during mediation.³⁸²

c. Voluntary Participation & Procedural Fairness

Fairness may also be defined in procedural terms, and, in this respect, the evidence is mixed. Concerns have been voiced about the fairness of the process.³⁸³ Practitioners point to truncated fact finding and “rough justice.”³⁸⁴ FINMAC permits multiple hearings, but, in practice, most mediations have only one, and they may be cursory.³⁸⁵ In 2014, out of 108 cases, eighty-six ended following one session; ten cases were terminated after the second.³⁸⁶ JBA mediations are reported to focus on documents and transactional suitability to the exclusion of other relevant factors.³⁸⁷ Practitioners suggest that if they are going to mediate a claim using a private ADR organization, they much prefer the mediation services provided by the Japanese Federation of Bar Associations, where subject matter experts are also available and the fact-finding and mediation process is conducted with greater care and deliberation.³⁸⁸

At the same time, there is evidence suggesting that some mediators have applied the rules in such a manner to facilitate fact-finding that is unlikely to occur in other forms of private ADR, or in the courts.³⁸⁹ The Financial ADR manuals used by mediators suggest that investors may obtain discovery of documents unavailable in civil court proceedings.³⁹⁰ While the Japanese Code of Civil Procedure

382 *Id.* at 9. NOBUKORE, *supra* note 319, at 6.

383 Summary of Arguments at Council of Advisers for Follow-Up of Financial ADR System, *supra* note 106.

384 E-mail from Osaka attorney to authors, *supra* note 8.

385 *Id.*

386 FINMAC, Heisei 26 Nendo no Sōdan, Kujō, Assen Saru-tate no Tokuchō Ni Tsuite [On the Features of Consultation, Complaints, and Complaints Filed in FY 2006] 9 (May 29, 2017), http://www.finmac.or.jp/tokei-siryo/pdf/highlight/highlight_2014.pdf (Japan).

387 Aoki, *supra* note 212; Ikenaga et al., *supra* note 126 at 15–16.

388 E-mail from Osaka attorney to authors, *supra* note 8.

389 See *infra* notes 369–372.

390 The manual authors suggested that internal memos reflecting hierarchical decision-making (*ringisho*) could be the subject of a request for documents by mediators. See Tokyo Bengoshikai Bengoshi-Gyomu-Bu Kin’yuu Funso Kenkyukai, Q&A Kin’yuu ADR Katsuyo Gaidobukku: Kaikeitsudekiru Shoken Ginko Hoken no Toraburu! [Q&A Guidebook for Utilizing Financial ADR: You can Resolve Troubles about Securities, Banking, and Insurance!] 85 (Nihon Kajoshiki Shuppan, 2012) (Japan); ISHIZUKA, *supra* note 198 at 165–166; Hironori Nishikino,

codifies an obligation to submit documents to the court, a court order to submit documents is usually entered only after voluntary submissions from parties during the first several hearings, and the parties may protect as privileged a wide range of corporate internal documents.³⁹¹ In contrast, the JBA Secretariat reports that mediators may request, e.g., production of internal memos (*ringi-sho*) and internal bank policies for interacting with elderly bank customers.³⁹² In other cases, mediators are reported to have requested and reviewed voice recordings of the telephone interactions between the broker and investor.³⁹³ Members of the plaintiffs' bar suggest this happens rarely,³⁹⁴ but there is evidence that it does happen.³⁹⁵

There may be other procedural advantages to the Financial ADR System. In some instances, they preserve a necessary relationship. Practitioners report that small-to-medium businesses are so dependent on financing from their banks that they hesitate to bring disputes to court.³⁹⁶ Doing so constitutes an open denunciation of the bank and could result in the denial of loans necessary to the

Kin'yuu ADR Taio no Chakuganten: Seido Tetuzuki no Tokucho to Taioji no 10 no Gensoku [How to Deal with Financial ADR as an Attorney for a Financial Service Provider: Its Distinctive Features and Ten Principles to Remember], 1946 KIN'YUU HOMU JIJO 34 (Japan). Both Ishizuka and Nishikino also suggest that banks can and should ask for in camera proceedings for such trade secret materials as internal memos or manuals. See ISHIZUKA, *supra* note 198 at 165–166; Hironori Nishikino, *Kin'yuu ADR Taio no Chakuganten: Seido Tetuzuki no Tokucho to Taioji no 10 no Gensoku* [How to Deal with Financial ADR as an Attorney for Financial Service Provider: Its Distinctive Features and Ten Principles to Remember], 1946 KIN'YUU HOMU JIJO 34 (Japan). In contrast, Japanese courts, in general, have privileged internal memos and manuals of banks as “documents only for internal use,” arguing that their disclosure poses a danger to the bank’s business, which normally surpasses the documents’ evidential value. See Saikō Saibansho [Sup. Ct.] Nov. 12, 2009, no. 1787, SAIKŌ SAIBANSHO MINJI HANREISHŪ [MINSHŪ] 8, 53 (Japan).

391 C. CIV. PRO., Law No. 109 of 1996, art. 220 (Japan).

392 E-mail from JBA Secretariat to authors, *supra* note 93. See also Tanaka, *supra* note 9 (discussing bank manuals for dealing with elderly clients).

393 See, e.g., FINMAC, FINMAC FUNSO-KAIKETSU TETSUZUKI JIREI (HEISEI 28 NEN 10-12 GATSU: SHITEI-FUNSO-KAIKETSU-KIKAN GYOMU) [FINMAC DISPUTE SETTLEMENT PROCEDURE (OCTOBER-DECEMBER HEISEI 20): DESIGNATED DISPUTE SETTLEMENT BODY OPERATION] 2, available at https://www.finmac.or.jp/tokei-siryo/pdf/sitei-assen/2014_10-12_conflict.pdf (Japan).

394 E-mail from Tokyo attorney to authors, *supra* note 121. E-mail from Osaka attorney to authors, *supra* note 8.

395 E-mail from JBA Secretariat to authors, *supra* note 93.

396 E-mail from Tokyo attorney to authors, *supra* note 121.

operation of their business.³⁹⁷ Yet, they are willing to mediate a claim in a confidential proceeding. And, banks are reportedly more ready to negotiate compensation to small-to-medium business customers for fear of default.³⁹⁸

Business entities turned early on to the Financial ADR System to mediate substantial claims. A Financial Supervisory Agency report citing numbers from 2011 stated that, with the increase in foreign exchange derivative claims, business entities brought thirty-eight percent of the total claims heard by all eight DDROs and forty-two percent of those claims exceeded ¥10 million (\$100,000) in claimed damages.³⁹⁹

Businesses turned to Financial ADR with such frequency that observers voiced concerns about having adequate resources to address the small investors' claims.⁴⁰⁰ The Advisory Council reviewing the Financial ADR System three years after its implementation, however, endorsed the use of DDRO mediation by business entities and by attorneys.⁴⁰¹ They pointed out that many small-to-medium

397 *Id.* See also Ogata, *supra* note 217; ASAHI SHINBUN, *supra* note 227; Tetsuhiro Sato, *Tagaku Sonshitsu no Kawase Deribatibu: ADR de Amai ni Naru Ginko no Sekinin* [Foreign Exchange Derivatives Caused Huge Loss: Now Banks' Responsibility Blurred by Disposition through ADR], 32 EKONOMISUTO 4232 (Japan).

398 Ogata, *supra* note 217. Ogata reported anonymous cases where bank representatives asked customers not to file cases with the JBA and argued that banks could negotiate de-facto settlement in forms of additional loans to cover early termination charges. *Id.* With such arrangements, banks could avoid loss by settlement payment, turning their due early termination charge into debt with decent interest. Other news coverages also reported arrangements for additional loans came with high interest rates and additional mortgages. See *id.*; ASAHI SHINBUN, *supra* note 227; Sato, *supra* note 397.

399 Fin. Servs. Agency, *supra* note 106.

400 Yamamoto, *supra* note 346, at 3. Fin. Servs. Agency, *supra* note 106, at 4. Some proposed limiting use to individuals or small claims while drafting the Financial ADR legislation. See Yanase, *supra* note 2, at 79. A number of jurisdictions have implemented such limitations. The Hong Kong Financial Dispute Resolution Center restricts eligibility to individual consumers and sole proprietors. Ali & Da Roza, *supra* note 80, at 504. The UK and Australia restrict use to retail clients including defined small businesses. *Id.* Some jurisdictions also restrict eligibility by capping maximum awards: The maximum that the Ombudsman in the United Kingdom may award is £150,000; Australia's maximum is AU\$280,000, absent special circumstances. See Ali, *supra* note 41, at 341–43.

401 See Fin. Servs. Agency, *Kin'yū shōhin torihiki-hō-tō no ichibu o kaisei suru hōritsu [fusoko]* (Heisei 21-nen 6 tsuki 24-nichi kaisei) [An Act to Revise Part of the Financial Instruments and Exchange Act (Supplementary Provisions) (Revised on June 24, 2009) art. 21] 2 (Nov. 7, 2012), available at <http://www.fsa.go.jp/singi/adr-followup/secchi.pdf> (Japan); Fin. Servs. Agency, *supra* note

enterprises in Japan are no better equipped with information and bargaining power than individual customers.⁴⁰² And the Financial ADR System may serve the needs of the small business community in a way that the courts cannot—they offer a means to resolve disputes and obtain a measure of compensation without jeopardizing financing.

In short, codification of the Financial ADR System has resulted in a forum for investors who may otherwise have no legal remedy or hesitate to pursue one.

VI. Conclusion

Many practitioners who specialize in representing investors avoid the Financial ADR System.⁴⁰³ They mediate only when the client does not want to file suit or circumstances preclude it.⁴⁰⁴ They suggest that mediation does not work most of the time and awards are almost always lower than what investors will receive in court.⁴⁰⁵ They call for increased transparency, independence, and use of third party experts to evaluate claims.⁴⁰⁶

Yet, there are indications that, as an institution, the Financial ADR System serves a useful purpose: it supplements Japan's civil justice system. The Financial ADR System serves an underserved segment of the Japanese population: the elderly, those bringing small claims, and parties who are unable or unwilling to litigate.

For routine cases, it offers an expedited, inexpensive forum driven by legal professionals applying legal norms. In exceptional cases, it may offer expanded discovery through reference to records that are often not discoverable in civil litigation. In exceptional cases, it may offer expanded remedies, e.g., offering "sophisticated" business investors a partial remedy where none may be had in the courts.

The fact that participation is voluntary and the investor stays

106, at 4.

402 *Id.*

403 E-mail from Tokyo attorney to authors, *supra* note 121. E-mail from Osaka attorney to authors, *supra* note 8.

404 E-mail from Tokyo attorney to authors, *supra* note 121. E-mail from Osaka attorney to authors, *supra* note 8.

405 E-mail from Tokyo attorney to authors, *supra* note 121. E-mail from Osaka attorney to authors, *supra* note 8.

406 E-mail from Tokyo attorney to authors, *supra* note 121. E-mail from Osaka attorney to authors, *supra* note 8. E-mail from Kyoto attorney to authors, *supra* note 145.

in the driver's seat—free to terminate mediation and litigate at any time—makes the participation and settlement rates meaningful. They suggest the Financial ADR System offers reasonable relief, when compared to published comparative fault averages. And it does so in a timely, inexpensive manner.

Participants in the process readily admit that it is not for the big cases, and, given its emphasis on compromise, it is not for cases involving cold liability or novel questions of law or fact. But the courts remain open to such cases, permitting continued application and evolution of the law.

Japan's voluntary Financial ADR System offers a modern paradigm for institutionalized mediation that meaningfully supplements the civil justice system.

* * *