OHIO STATE LAW JOURNAL

2019 – 2020

EDITOR IN CHIEF
Kaela King

EXECUTIVE EDITOR
S. Matthew Krzacak

CHIEF ONLINE EDITOR
Lindsey M. Woods

CHIEF ARTICLES EDITOR
Dustin M. Lorenzo

CHIEF NOTE EDITOR
Stacey A. Dettwiller

SYMMETRY EDITOR
Kathleen A. Evans

EXECUTIVE ARTICLES EDITORS
Samuel V. Lioi
Katherine M. Toth

MANAGING EDITORS
S. Matthew Krsacok

ARTICLES EDITORS
Timothy J. Beavers
Morgan J. Brant
Patrick J. Cleary
Matthew S. Hauer
Adam C. Kennett
Julia Lauritzen
Zach Leciejewski
Diane Long
Alison Roth
Brittany Stevenson

NOTE EDITORS
Seth J. Barany
Kelsie N. Hendren
Matthew Somogyi

SOURCE EDITORS
Amanda G. Jones
Brandon J. McNeill

BUSINESS EDITOR
John Osinski

MANAGING EDITORS
Allyson M. Fleming
Erin K. Hassett
Bethany Jones
Riley Kane
Rebecca E. Kanz
Spencer Meador
Todd Minton
Donald B. Naiman
Evan Poole
Benjamin J. Reeb
Micaela Taylor
Erica M. Van Heyde
Will White
Clairanne Wise

DARK EDITOR
Elizabeth A. Dengler

ASSOCIATE EDITORS
Meagan Dimond
Andria Dorsten Ebert
Philip Ewing
Silvia Francis-Bongue
H. Cole Hassay
Madison Hill
Samuel Huryn
CJ Kowalski
Austin R. Limes
Megan Luby
Nicole Mattingly
Richard McCutcheon
Michael McDonald
Sanjay Nevrekar
Jordan C. Patterson
Megan Porter
Nicholas O. Santos
Susanna Savage
Jacob Selsor
Dane Sowers

ASSOCIATE EDITORS
Caitlin Throne
Zachary Valdman
Alexander Walker
Jenson Wang
Sam Weir
Brittney M. Welch
Hannah Wirt
Christy Wolf
Jacob Zuberi

FACULTY ADVISORS
Deborah Jones Merritt
Anne E. Ralph
Christopher J. Walker

Published Six Times a Year by Students of The Ohio State University Moritz College of Law
OFFICERS OF ADMINISTRATION

Michael V. Drake, B.A., M.D., President of the University
Bruce A. McPherson, B.S., M.S., Ph.D., Executive Vice President and Provost of the University
Lincoln E. Davies, B.S., J.D., Dean of the College of Arts and Sciences
Daniel P. Tokaji, A.B., J.D., Assoc. Dean for Faculty, Charles W. Ebersold & Florence Whitcomb Ebersold Prof. of Constitutional Law, and Senior Fellow of Election Law at Moritz
Kathy S. Northern, B.A., J.D., Assoc. Dean for Diversity & Inclusion and Robert M. Duncan/Jones Day Designated Assoc. Prof. of Law
Paul Rose, B.A., J.D., Assoc. Dean for Academic Affairs, Robert J. Watkins/Procter & Gamble Prof. of Law, and Executive Director of Law, Finance and Governance Prog.
Laura Fernandez, B.S., J.D., Assist. Dean for International and Graduate Affairs and Adjunct Prof.
Sara A. Sampam, B.S., M.S., J.D., Assist. Dean for Information Services, Director of Law Library, and Senior Lecturer
Darren Nealy, B.A., J.D., Assist. Dean for Academic Affairs and Adjunct Prof.
Michael States, B.A., J.D., Assist. Dean for Admissions & Financial Aid

FACULTY EMERITI

Mary Beth Beazley, B.A., J.D.
Michael Braunstein, B.A., J.D.
Sanford N. Caust-Ellenberg, M.C.R.P., J.D.
Albert L. Clovis, B.A., M.A., L.L.B.
Sharon L. Davies, B.A., J.D.
Joshua Dressler, B.A., J.D.
Howard P. Fink, B.A., L.L.B.
Gordon E. Gee, B.A., Ed.D, J.D.
David A. Goldberger, B.A., J.D.
Lawrence R. Herman, A.B., L.L.B.
Louis A. Jacobs, A.B., J.D., L.L.M.
Bruce S. Johnson, B.A., J.D., M.L.S.
Michael Kindred, B.A., J.D., M.C.L., D.E.S.
Joan M. Krasskopf, A.B., J.D.
John B. Quigley, A.B., M.A., L.L.B.
Rhonda R. Rivera, B.A., M.P.A., J.D.
Nancy H. Rogers, B.A., J.D.
Michael D. Rose, B.A., J.D., L.L.M.
Allan J. Samansky, B.A., M.A., J.D.
Joseph B. Stulberg, B.A., J.D., M.A., Ph.D.
Gregory M. Travalo, B.A., J.D., L.L.M.
Vincenzo F. Verdun, B.A., J.D.
Douglas J. Whaley, B.A., J.D.
Charles E. Wilson, B.S., J.D.

FACULTY

Anna Akbar, B.A., J.D., Assoc. Prof. of Law
Douglas A. Berman, A.B., J.D., Newton D. Baker-Baker & Hostetler Chair in Law
Micah Berman, B.A., J.D., Assoc. Prof. of Public Health and Law
Gregory A. Caldeira, B.A. A.M., Ph.D., Distinguished Univ. Prof., Am and Dorrell Deuber Chair in Pol. Comm. and Policy Thinking and Prof. of Law
Cinnamon P. Carlarne, B.A., B.C.L., M.S., J.D., Alumni Society Designated Professor of Law
Martha Chamallas, B.A., J.D., Robert J. Lynn Chair in Law
Bryan H. Choi, A.B., J.D. Prof. of Law and Engineering
Daniel C.K. Chow, B.A., J.D., Frank E. & Virginia H. Brazier Chair in Business Law
Holly Coats, B.A., J.D., Visiting Assist. Prof. of Law
Amy J. Cohen, B.A., J.D., John C. Elam/Vorys Sater Chair in Law
Sarah Rudolph Cole, B.A., J.D., John W. Bricker Prof. of Law and Director of Prog. on Dispute Resolution
Ruth Colker, A.B., J.D., Distinguished Univ. Prof. and Grace Fern Heck Fueit Chair in Constitutional Law
Olwyn Conway, B.A., J.D., Prof. of Law
Ellen E. Deason, B.A., M.S., J.D., Joanne W. Murphy/Classes of 1965 and 1973 Prof. of Law
Terri L. Enns, B.A., Clinical Prof. of Law, Senior Fellow of Election Law at Moritz
Katherine Hunt Federle, B.A., J.D., LL.M., Joseph S. Platt-Porter Wright Morris & Arthur Prof. of Law and Director of the Center for Interdisciplinary Law and Policy Studies
Edward B. Foley, B.A., J.D., Charles W. Ebersold & Florence Whitcomb Ebersold Chair in Constitutional Law and Director of Election Law at Moritz
William Froehlich, B.A., J.D., Langdon Fellow in Dispute Resolution
Larry T. Garvin, B.A., B.S., M.S., J.D., Lawrence D. Stanley Prof. of Law
Art Gloyower, B.A., J.D., L.L.M. Assist. Prof. of Law
Arthur F. Greenbaum, B.A., J.D., James W. Shocknessy Prof. of Law
L. Camille Hebert, B.A., J.D., Carter C. Kissell Prof. of Law
Mohamed S. Helal, B.A., M.A., L.L.B., L.L.M, S.J.D., Assist. Prof. of Law
Steven D. Hirsch, B.A., J.D., Prof. of Law, Director of Prog. on Data and Governance
Stephanie R. Hoffer, B.S., J.D., LL.M., Prof. of Law
Steven F. Hurftner, A.B., J.D., C. William O'Neill Professor in Law and Judicial Administration, Senior Fellow of Election Law at Moritz and Director of Clinical Prog.
Croela Johnson, B.S., J.D., President's Club Prof. of Law
Kimberly Jordan, B.S., J.D., Clinical Prof. of Law and Director of Justice for Children Project
Katherine S. Kelly, B.A., M.A., J.D., Clinical Prof. of Law and Director of Academic Support Prog.
Katrina J. Lee, B.A., J.D., Clinical Prof. of Law
Deborah Jones Merritt, A.B., J.D., John Deaver Drinko-Baker and Hostetler Chair in Law
Alan C. Michael, A.B., J.D., Dean Emeritus and Edwin M. Cooperman Chair in Law
Dale A. Oesterle, B.A., M.P.P., J.D., Gilbert Reese Chair in Contract Law
Etzioniometrics Paradisid, B.A., MBE, J.D., Prof. of Law and Public Health
Anne Ralph, B.A., J.D., Clinical Prof. of Law
Courtlyn Roser-Jones, B.A., J.D., L.L.M., Assist. Prof. of Law
Gay A. Rub, B.A., L.L.B., M.A., L.L.M, S.J.D., Prof. of Law
Dakota Rudessil, B.A., J.D., Assist. Prof. of Law
Colleen Garry Settineri, B.S., B.A., J.D., Assist. Clinical Prof. of Law
Peter M. Shae, A.B., J.D., Jacob E. Davis & Jacob E. Davis II Chair in Law
Rick Simmons, B.A., M.A., J.D., Chief Justice Thomas J. Moyer Prof. for the Administration of Justice and the Rule of Law
Marc S. Spindelman, B.A., J.D., Isadore & Ida Topper Prof. of Law
Todd A. Starker, B.A., M.B.A., J.D., Clinical Prof. of Law
David Stebenne, B.A., M.A., J.D., Ph.D., Prof. of History and Law
Christopher J. Walker, B.A., M.P.P., J.D., Prof. of Law, Director of Washington D.C. Summer Prog.
Paige L. Wilson, B.A., J.D., Assist. Clinical Prof. of Law
Patricia J. Zettler, B.A., J.D., Assist. Prof. of Law

ADJUNCT FACULTY

Karim Ali
Elizabeth L. Anastaett
Susan Aryanar
David T. Ball
Christopher Ballard
Hubbard Barton
Jordan Berman
Allison Binckley
Amy Bittner
Ben Bodamer
Joseph Boeckman
Joseph M. Caliguri
Kwaine Christens
Hon. R. Guy Cole Jr.
Erica L. Cook
Matt Cooper
Lorenzo Corte
Jonathan E. Coughlan
Sanir Dahnman
Josh Dhyani
Scott E.ailor
Laura Fernandez-Alvarez
Tod Friedman
Gates Garry-Rokous
Paul Gatz
Franz Geiger
Peter Glenn-Applegate
David W. Grauer
Christopher Hammond
Michael Henderson
Robb Hnern
Hon. John E. Hoffman Jr.
Melissa Jackson
Caryn Kaufman
Marya C. Kolman
James K.L. Lawrence
George Limbert
Emma MacGudwin
Hon. AlgonIton L. Marbley
Susan Malloy

Greg Mathews
Patrick McCarthy
Hon. Stephen L. McIntosh
Richard M. Mescher
Robert J. Miller
Rebecca Monroe
Robert Oast
Ryan O'Rourke
Matt Palmer
Sarah Paxton
Matt Richardson
Michael Robertson
Dan D. Sandman
Hon. Edmund A. Sargus
Hon. Jennifer L. Sargus
Don Scheetz
Edward M. Segelken
Scott V. Simpson
Rob Solomon
Douglass Squares

Carter Stewart
Hon. Jeffrey S. Sutton
Justin Thompson
Katrina Thompson
Nikki Trautman-Basynski
Hon. Chelsey M. Vascara
Bradley Waller
Hon. Michael H. Watson
Robert Weiler
Terrence Wheeler
Suzanne Whisler
Stephen White
Megan Whited
Geoff Wilcox
Reid Wilson
Mary Beth Young
Stephanie Ziegler
Planning for Failure: Pipelines, Risk, and the Energy Revolution
Sara Gosman ..............................................................349

First Amendment (Un)Exceptionalism: A Comparative Taxonomy of Campaign Finance Reform Proposals in the United States and United Kingdom
Lori A. Ringhand ......................................................405

Betting On Education
Kathryn Kisska-Schulze & John T. Holden ..................465

Impoverished IP
Stephanie Plamondon Bair .........................................523

Note

Would You Like to Save Your Game?: Establishing a Legal Framework for Long-Term Digital Game Preservation
Will White ..............................................................567
Planning for Failure: Pipelines, Risk, and the Energy Revolution

SARA GOSMAN*

In 2014, as production soared in North Dakota’s oil fields, Energy Transfer Partners proposed a large pipeline to transport the oil to market. The very name of the project—Dakota Access—conveyed the company’s optimistic vision of a needed link between the prolific oil fields and the rest of the country. The vast scale of the pipeline project was matched only by the intensity of the opposition to the route. A bitter controversy erupted at Standing Rock over the risk of a catastrophic oil spill. Tribal members and environmentalists from across the nation united to protest the company’s decision to site the pipeline underneath a lake that serves as the sole source of drinking water for local tribes. The company defended the safety of its pipeline and ultimately prevailed. Dakota Access was completed in 2017. Since it began operation, it has leaked eight times.

One would expect risk governance to take a more preventative approach to risk, as the potential for catastrophic harm increases and the ability to predict an accident decreases. But projects such as Dakota Access raise troubling questions about the current system governing the risks of energy pipelines. Why are pipelines being sited in environmentally sensitive and densely populated areas? To what extent does the system address the long-term risks of spills and releases? These questions are more important than ever before, as the domestic revolution in oil and gas production fundamentally reshapes pipeline networks and the geographic and political landscape of risk.

This Article seeks answers by examining the laws governing energy pipelines through the lens of risk. The analysis reveals a critical flaw in risk governance: the risks associated with “siting” a pipeline are treated separately from the long-term “safety” of the pipeline. This formal legal distinction has a substantial practical effect on the risk landscape. By failing to consider the risks of an accident in the decision of where to locate a pipeline—that is, by failing to plan for failure—the system allows energy pipelines to be sited near people and sensitive ecosystems. This in turn leads to more accidents in vulnerable areas.

* Associate Professor of Law, University of Arkansas School of Law—Fayetteville. The author wishes to thank her colleagues and the many faculty who gave helpful feedback during presentations at the Colloquium on Environmental Scholarship at Vermont Law School, the Works in Progress Symposium at the University of Colorado School of Law, and the Southeastern Association of Law Schools Conference. The author also wishes to thank Tara Righetti for her response to the Article and for her insights on risk governance solutions.
and a greater risk burden on landowners and surrounding communities. The result is a system that sites first—and cleans up later.

TABLE OF CONTENTS

I. INTRODUCTION .................................................................350

II. RECONSTRUCTING THE ENERGY SYSTEM............................356
   A. Energy Pipeline Networks ..............................................357
   B. The Energy Revolution ................................................363

III. THE LEGAL FRAMEWORKS GOVERNING RISK .......................370
   A. The Risk of Energy Pipelines .......................................370
   B. The Safety Framework ...............................................373
   C. The Siting Framework ...............................................377

IV. EVALUATING THE RISK FRAMEWORKS ...............................389
   A. The Risk Policy Approaches .......................................389
   B. The Results ..........................................................394
   C. Three Potential Solutions .........................................399

V. CONCLUSION .....................................................................403

“The Dakota Access Pipeline is built to be one of the safest, most technologically advanced pipelines in the world.”

“Just because oil is flowing today doesn’t mean it won’t leak in the future . . . . There’s an uneasy feeling that any moment, this pipeline could pose a threat to our way of life. It’s something you have to carry and be wary of all the time, and be ready for.”

I. INTRODUCTION

In 2014, Energy Transfer Partners announced it would construct a new pipeline, the Dakota Access pipeline, to transport crude oil from the Bakken shale play in North Dakota to markets in the Midwest and on the Gulf Coast. The scale of the pipeline project reflected the bravado of the domestic revolution

---


in oil and gas production. The proposed pipeline was designed to transport up to 570,000 barrels of oil each day, approximately half of the daily production in the Bakken oil field.4 The planned route was over a thousand miles long and crossed fifty counties in four states.5 After receiving approval from each state, Energy Transfer Partners began to build the pipeline. But the company encountered significant resistance when it sought a federal easement to cross the Missouri River just above the Standing Rock reservation at Lake Oahe, the only source of drinking water for local tribes.6 The risk of an oil spill ignited a protest that would last months.7 Tribes and environmentalists challenged the company’s decision to site the pipeline in such a sensitive area and the legal framework that would allow a company to use such a route.8 After President Trump interceded and the U.S. Army Corps of Engineers (USACE) granted the easement, Energy Transfer Partners finally completed the project in 2017— at a cost of $7.5 billion, almost double the project estimate.9

The opposition to the Dakota Access pipeline is part of a larger trend: communities are increasingly organizing to fight new energy pipelines. Residents of the densely populated Northeast,10 farmers in Iowa,11 and rural

---


7 Meyer, supra note 2.


9 See Meyer, supra note 2.


homeowners in Georgia, Virginia, Kentucky, and Pennsylvania have sought to halt the construction of pipelines. Facing such public opposition, companies have cancelled or delayed several pipelines. Protesters are painted as discontented, “not-in-my-backyard” property owners complaining about unwanted land uses, or as fringe environmentalists who want to keep oil and natural gas in the ground. But there is an alternative explanation: communities are increasingly concerned about the long-term risks of energy pipelines, including the impacts of spills and releases on the environment, human health, and public safety.

The concern is justified. The legal frameworks governing energy pipelines impose unnecessary risks on communities and the environment by dividing approval of pipeline “siting” from standards for pipeline “safety.” When government agencies review a proposed pipeline project, they primarily focus on the need for the pipeline and the negative effects of pipeline construction, not on the need for energy pipelines. The legal frameworks governing energy pipelines impose unnecessary risks on communities and the environment by dividing approval of pipeline “siting” from standards for pipeline “safety.”

---

13 See id.
18 See Fifield, supra note 12 (quoting Melissa Ruiz, spokeswoman for Kinder Morgan, as saying “[t]here are always going to be people who say not in my backyard”).
19 See Jordan Blum, Protests, Arrests Pick Up as Environmentalists Target Pipelines, HOUS. CHRON. (Oct. 19, 2016), http://www.houstonchronicle.com/business/energy/article/Protests-arrests-pick-up-as-environmentalists-9983690.php [https://perma.cc/3XWV-U4TZ] (stating that “environmentalists are targeting pipelines as the new public enemy number one” and “[w]ith the ‘keep it in the ground’ movement gaining support, the focus of efforts to slow the extraction of oil and gas has increasingly become pipelines”).
20 There are even longer-term risks of energy pipelines, notably the climate change enabled by fossil fuel infrastructure. Climate activists are increasingly focused on halting new pipelines to mitigate future greenhouse gas emissions. This Article considers the risks of spills and releases, leaving the issue of climate change risk to future work.
rather than on the risks of operation. Managing the risks of spills and releases to the surrounding area is left to the pipeline safety standards—which go into effect once the route is already established. Combined with a compelled transfer of property rights through eminent domain, the result is a governance system that defers largely to the pipeline company and its chosen route.

The Standing Rock controversy is the exception that proves the rule. In addition to siting approval from each state, Energy Transfer Partners was required to obtain an easement from the USACE because the Dakota Access pipeline crossed a small area of a civil works project. Under the National Environmental Policy Act (NEPA), the agency could not make a decision on the easement until it conducted an environmental assessment of the project to determine whether there were significant environmental impacts. Relying on federal safety standards, the USACE assessed the risk of an oil spill as low and concluded that a leak during pipeline operation was an “unlikely event.” This assessment would prove incorrect for the pipeline as a whole. From 2017 to 2018, the Dakota Access pipeline has leaked eight times. The leaks were relatively small, however, and none have affected Lake Oahe.

The risk governance of oil and natural gas pipelines has never been more important. A revolution in domestic oil and gas production is occurring, as development companies extract the resources from shale and other “tight” rock

---

23 See discussion infra Part III.C.
24 See discussion infra Part III.B.
27 JOHN W. HENDERSON, U.S. ARMY CORPS OF ENG’RS, ENVIRONMENTAL ASSESSMENT: DAKOTA ACCESS PIPELINE PROJECT 88–94 (July 2016). Tribes and environmental organizations challenged the assessment, but the court deferred to the USACE and upheld the risk analysis. See Standing Rock Sioux Tribe v. U.S. Army Corps of Eng’rs, 255 F. Supp. 3d 101, 127 (D.D.C. 2017). The court remanded the assessment to the agency, however, to consider whether the project would be highly controversial in light of scientific critiques of the risk analysis and to analyze the effects of an oil spill on tribal hunting and fishing rights. Id. at 145–48. On remand, the USACE determined that the effects were not significant. JOHN L. HUDSON, U.S. ARMY CORPS OF ENG’RS, ANALYSIS OF THE ISSUES REMANDED BY THE U.S. DISTRICT COURT FOR THE DISTRICT OF COLUMBIA RELATED TO THE DAKOTA ACCESS PIPELINE CROSSING AT LAKE OAHE 41 (Aug. 2018). As this Article went to press, the court has once again remanded the assessment to the agency because the “[u]nrefuted expert critiques regarding [the risk of oil spills] mean that the easement approval remains ‘highly controversial’” and the USACE must prepare an Environmental Impact Statement. Standing Rock Sioux Tribe v. U.S. Army Corps of Eng’rs, No. 16-1534 JEB, 2020 WL 1441923 (D.D.C. Mar. 25, 2020).
29 Id. One affected a drinking water source area in Patoka, Illinois.
formations in plays across the country.\textsuperscript{30} The plays supply immense amounts of oil and gas, and production is growing.\textsuperscript{31} Because of this revolution, the United States became a net exporter of natural gas in 2017\textsuperscript{32} and the largest producer of crude oil in the world in 2018.\textsuperscript{33} The rapid increase in oil and gas supply and changing demand are altering pipeline networks in profound ways. From 2004 to 2017, the pipeline industry constructed over 125,000 miles of pipelines to transport oil and gas from production areas in the United States and the border of Canada to areas where the commodities will be used or exported.\textsuperscript{34} The industry expects to continue construction, with an additional 150,000 miles of new pipelines by 2035.\textsuperscript{35}

The growing scholarly literature on energy pipelines has focused on the siting frameworks that govern natural gas pipelines, oil pipelines, and other energy transportation infrastructure.\textsuperscript{36} Scholars have explored questions of

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{30}Michael Ratner & Mary Tiemann, Cong. Research Serv., R43148, An Overview of Unconventional Oil and Natural Gas: Resources and Federal Actions 1–7 (2015).
\item \textsuperscript{34}See discussion infra Part II.B.
\item \textsuperscript{35}Keven Petak et al., ICF, North America Midstream Infrastructure through 2035, at 88 (June 2018), https://www.ingaa.org/File.aspx?id=34703 [https://perma.cc/EKW6-WPY6] [hereinafter ICF, North America] (reporting a total estimate of 149,128 miles of new pipelines by 2035 across the following three categories: 110,018 miles of new oil and gas gathering pipelines; 36,129 miles of oil, gas, and natural gas liquid pipelines; and 2,981 miles of oil product pipelines). Energy Transfer Partners has joined the fray again; it is proposing three new pump stations to increase the amount of oil the Dakota Access pipeline can transport. See Growth Opportunity, Dakota Access Pipeline, https://daplpipelinefacts.com/Updates.html [https://perma.cc/35BT-XRJN].
\item \textsuperscript{36}See generally, e.g., Alexandra B. Klass, Future-Proofing Energy Transport Law, 94 Wash. U. L. Rev. 827 (2017) [hereinafter Klass, Future-Proofing] (considering how siting laws can encourage transportation infrastructure for clean energy); Alexandra B. Klass & Danielle Meinhardt, Transporting Oil and Gas: U.S. Infrastructure Challenges, 100 Iowa L. Rev. 947 (2015) (assessing whether siting laws governing oil and gas pipeline networks are sufficient to facilitate new oil and gas pipelines); Tara K. Righetti, Siting Carbon Dioxide Pipelines, 3 Oil & Gas Nat. Resources & Energy J. 907 (2017) (analyzing the current siting framework for carbon dioxide pipelines).
\end{itemize}
\end{footnotesize}
federalism, the use of eminent domain, the role of climate change, and the criteria governing the siting of transboundary pipelines. But the academy has paid less attention to the risks posed by pipeline accidents. This Article builds on the author’s previous work on pipeline safety and risk to consider how pipeline siting laws and safety laws address the long-term risk of pipeline accidents. It presents the first comprehensive analysis of siting laws through the lens of risk. As described in the Article, the siting framework is certainly fragmented. Jurisdiction is scattered among the federal government and the states, and various agencies apply different criteria to pipeline proposals. These differences are largely cosmetic, however, when viewed through the lens of long-term risk. The overarching legal framework produces a surprisingly uniform result: This is a framework that sites first—and cleans up later.

The Article is divided into four parts. In Part II, I provide a detailed analysis of current data to describe the pipeline networks and explain the effect of the domestic energy revolution on those networks. In Part III, I explain the risks of energy pipelines and the ways in which the “safety” and “siting” legal frameworks address risk. This Part includes a fifty-state analysis of pipeline siting laws. In Part IV, I present a typology of risk policy approaches—a preventative approach, a management approach, and a remedial approach. While a preventative approach appears best suited to pipeline risk, I argue that the division between the siting and safety frameworks leads to less prevention. I contend that this creates several negative results, such as an increase in the

37 See, e.g., Alexandra B. Klass & Jim Rossi, Reconstituting the Federalism Battle in Energy Transportation, 41 HARV. ENVTL. L. REV. 423, 458–63 (2017) (exploring federalism tensions in energy transportation infrastructure, including in the siting of interstate gas pipelines); Amy L. Stein, The Tipping Point of Federalism, 45 CONN. L. REV. 217, 217, 237–38 (2012) (comparing state authority over siting of electricity generation facilities to federal control over the siting of other energy infrastructure, such as natural gas pipelines).


39 See generally, e.g., James W. Coleman, Beyond the Pipeline Wars: Reforming Environmental Assessment of Energy Transport Infrastructure, 2018 UTAH L. REV. 119.

40 See, e.g., Sam Kalen, Thirst for Oil and the Keystone XL Pipeline, 46 CREIGHTON L. REV. 1, 10–25 (2012) (examining the presidential permit criteria for the transboundary Keystone XL pipeline).

41 Cf. Klass, Future-Proofing, supra note 36, at 887–96 (arguing that energy law should encourage transportation of oil by rail rather than by pipeline, in part because of the greater potential for rail safety improvements); Righetti, supra note 36, at 925–27 (describing the safety framework governing carbon dioxide pipelines); see also David B. Spence, Regulation and the New Politics of (Energy) Market Entry, 95 NOTRE DAME L. REV. 327, 357–75, 378–81 (2019) (presenting the results of an empirical study on nonprofit organizations that oppose energy infrastructure and arguing that opponents are misrepresenting the risk of projects).

number and catastrophic harm of pipeline incidents. Finally, I identify a policy solution that would combine the siting and safety frameworks into one risk governance system.

II. RECONSTRUCTING THE ENERGY SYSTEM

Today’s energy system depends on vast networks of underground pipelines that transport petroleum to users, in order to heat buildings, generate electricity, fuel cars, and manufacture thousands of products.43 Approximately three million miles of pipelines44 lie underneath cities and rural areas, coastal waters and inland streams, forests and grasslands, in almost every part of the United States.45 The pipelines transport 28 trillion cubic feet of natural gas,46 11.38 billion barrels of crude oil,47 and 10.19 billion barrels of petroleum products48 per year. No other transportation mode for commodities—including railways,49 inland waterways,50 and paved roadways51—stretches as far to connect


45 Id. (describing pipelines as “geographically widespread, running alternately through remote and densely populated regions—from Arctic Alaska to the Gulf of Mexico and nearly everywhere in between”); Where Are Pipelines Located?, PIPELINE 101, https://pipeline101.org/Where-Are-Pipelines-Located [https://perma.cc/SF3Z-V57B].


48 Id.


50 See id. (reporting an estimated 25,000 miles of navigable channels, “which include rivers, bays, channels, and the inner route of the Southeast Alaskan Islands, but does not include the Great Lakes or deep ocean traffic”).

production and use. The sheer size of these pipeline networks is rivaled only by the transmission and distribution system for electricity. Unlike most electric wires, pipelines are hidden from public view until an operator constructs a new pipeline or an accident occurs, bringing the networks to the surface. A dramatic increase in domestic oil and natural gas production has spurred pipeline construction across the country—and reshaped the energy system.

A. Energy Pipeline Networks

There are two primary energy pipeline networks in the United States: a network that transports natural gas, and a network that transports liquid petroleum. In general, both networks carry domestically produced and imported fossil fuels to users and export facilities. But it is important to understand the specific features of the networks, both to appreciate the effects of the energy revolution and to evaluate the laws governing pipeline risk.

The natural gas pipeline network consists of three interconnected types of systems: gathering systems, transmission systems, and distribution systems. Operators of gathering pipeline systems collect raw natural gas from production fields. Depending on the constituents in the raw gas, the pipelines may carry the gas to a central processing facility or directly to the transmission pipeline system. Operators of transmission pipeline systems transport methane from gathering systems and import facilities to areas where the gas will be consumed or to terminals for export. Some large users—like power plants or manufacturing facilities—take delivery of the gas from transmission systems.

---


53 See How Do Pipelines Work?, PIPELINE 101, https://pipeline101.org/How-Do-Pipelines-Work [https://perma.cc/3L4Y-N9UZ]. Liquid pipelines also transport liquefied gases, such as carbon dioxide, and liquefied hydrocarbons associated with natural gas production, such as ethane, propane, and butane. Id.


55 In many cases, the gas extracted from the underground rock formation must be separated from water, solids, and crude oil. Id. at 89. The oil and heavier hydrocarbons, or lease condensate, are sent to a crude oil pipeline system. Id. The raw gas enters the gathering system at the separator outlet. Id.

56 Id. at 5. At a processing facility, methane is separated from impurities and other hydrocarbons. Id.

57 Id. at 6. The United States imports and exports natural gas by cross-border transmission pipelines and by tanker vessel. Id. at 92. To efficiently transport natural gas across the ocean, operators of export facilities must cool the gas until it liquefies. Id. Import terminals then re-gasify the methane before piping it to the transmission pipeline system. Id.

58 Id. at 6.
Transmission operators also deliver to local distribution companies, which transport the gas to customers through distribution pipelines.\textsuperscript{59}

There is very little information available on gas gathering systems because the vast majority are unregulated.\textsuperscript{60} An estimated 3500 private companies operate 356,000 to 400,000 miles of pipelines.\textsuperscript{61} In production areas where there are many wells and where operators have built processing and transmission infrastructure, the gathering systems are compact; short branch lines funnel gas from wells into larger pipelines that carry the gas to central facilities.\textsuperscript{62} In areas with fewer wells and less infrastructure development, the systems consist of longer pipelines that connect distant production fields.\textsuperscript{63} Traditionally, gathering systems are composed of small-diameter steel pipes that operate at low pressure.\textsuperscript{64} Operators utilize plastic and composite piping as cheaper alternatives to steel, but to what extent is unclear.\textsuperscript{65}

\textsuperscript{59} Id. at 99.

\textsuperscript{60} See PIPELINE & HAZARDOUS MATERIALS SAFETY ADMIN., PRELIMINARY REGULATORY IMPACT ASSESSMENT: NOTICE OF PROPOSED RULEMAKING—PIPELINE SAFETY: SAFETY OF GAS TRANSMISSION AND GATHERING PIPELINES 144 (Mar. 2016) [hereinafter PHMSA, PRELIMINARY REGULATORY IMPACT ASSESSMENT] (estimating that PHMSA regulates 3% of onshore gathering pipelines and contending that data should be collected on all gathering systems); U.S. GOV’T ACCOUNTABILITY OFFICE, GAO-12-388, COLLECTING DATA AND SHARING INFORMATION ON FEDERALLY UNREGULATED GATHERING PIPELINES COULD HELP ENHANCE SAFETY 7 (Mar. 2012) [hereinafter GAO, COLLECTING DATA AND SHARING INFORMATION].


\textsuperscript{65} See U.S. GOV’T ACCOUNTABILITY OFFICE, GAO-17-639, PIPELINE SAFETY: ADDITIONAL ACTIONS COULD IMPROVE FEDERAL USE OF DATA ON PIPELINE MATERIALS AND CORROSION 13–14, 16 (Aug. 2017) [hereinafter GAO, PIPELINE MATERIALS AND CORROSION].
In contrast to localized gathering systems tied to production fields, gas transmission pipeline systems crisscross the nation. More than 210 systems form an integrated transmission grid of 300,000 miles of line pipe. The industry is very concentrated: just twenty-seven private companies operate almost two-thirds of these miles. Interstate “trunkline” systems, which transport gas over long distances between a few collection and distribution points, make up most of the grid. The rest of the systems generally operate within a major market and serve that region; unlike trunkline systems, the main pipelines have many interconnections and branch lines where gas can enter and exit the system. Because transmission systems transport large volumes of gas, the pipes are generally wide—trunklines can be greater than three feet in diameter—and are constructed of steel that can withstand the high operating pressures. Mechanical devices compress gas from gathering systems to boost its pressure before the gas enters the transmission system, and compress gas once it is in the system to maintain pressure and counteract friction. There are more than 1400 compressor stations along the transmission grid.

Of the three types of natural gas pipeline systems, distribution systems are by far the most ubiquitous. Approximately 1400 local gas utilities distribute gas through 2.2 million miles of pipelines to 69.3 million customers in communities across the nation. The distribution systems begin at city gates, where operators

---

66 See Natural Gas Explained: Natural Gas Pipelines, supra note 46 (“The U.S. natural gas pipeline network is a highly integrated network that moves natural gas throughout the continental United States.”).


70 Transportation Process and Flow, supra note 69.


72 GAO, PIPELINE MATERIALS AND CORROSION, supra note 65, at 12–13.

73 See MIESNER & LEFFLER, supra note 54, at 5–6, 51, 105.

74 About U.S. Natural Gas Pipelines—Transporting Natural Gas, supra note 67.

75 Pipeline Mileage and Facilities, supra note 67 (follow “2010+ Pipeline Miles and Facilities” hyperlink) (reporting 1372 operators, 1,307,826.6 miles of main pipelines, and
receive gas from transmission pipelines and add an odorant to give methane its distinctive smell.76 Gas is delivered to residences and businesses through main pipelines and small service lines that branch off the “mains.”77 These pipes—particularly service lines—are generally smaller in diameter and operate at lower pressure.78 Beyond the basic function, the systems are remarkably heterogenous: they vary from small municipal systems of less than 1000 customers to large investor-owned systems that service over 100,000 customers.79 They also vary in materials. Systems may contain piping constructed of steel, plastic, cast iron, or copper.80 Most mains and service lines are now plastic because the material is flexible and resistant to corrosion.81

The liquid petroleum pipeline network also consists of three types of systems: crude oil gathering systems, crude oil transmission systems, and product transmission systems.82 The gathering pipeline systems, like their counterparts in the natural gas network, connect production fields to the transmission systems that provide long-distance transport.83 But crude oil is processed at refineries that can be hundreds of miles away from production areas.84 The crude oil transmission systems transport raw crude from domestic gathering systems to refineries,85 while separate transmission systems carry


77 Natural Gas Pipeline Systems, supra note 71.

78 See Fact Sheet: Distribution Pipelines, supra note 76.


80 Pipeline Mileage and Facilities, supra note 67; Pipeline Replacement Background, PIPELINE & HAZARDOUS MATERIALS SAFETY ADMIN., https://opsweb.phmsa.dot.gov/pipeline_replacement/ [https://perma.cc/LS58-35FU] (last updated Sept. 20, 2019) (noting that although cast and wrought iron pipelines are some of the oldest pipeline materials in the United States, operators have made progress in replacing these high-risk pipelines).

81 See GAO, PIPELINE MATERIALS AND CORROSION, supra note 65, at 14–16.


83 See MIESNER & LEFFLER, supra note 54, at 2–3.


85 MIESNER & LEFFLER, supra note 54, at 3–4.
petroleum products and hydrocarbon gas liquids to direct users and distribution terminals.\textsuperscript{86} Both types of transmission systems also serve import and export markets.\textsuperscript{87} There are no distribution pipeline systems similar to the ones for natural gas; instead, trucks generally transport products to end users.\textsuperscript{88}

Like gas gathering systems, crude oil gathering systems operate largely without regulation and so data on the systems is sparse.\textsuperscript{89} There are an estimated 30,000 to 40,000 miles of gathering pipelines;\textsuperscript{90} the number of operators is unknown. Gathering pipelines collect crude oil from storage tanks on well sites and carry it to oil terminals, where oil is aggregated in large tanks to await further transportation by transmission pipeline, rail, or truck.\textsuperscript{91} A system may also transport oil from several well sites to an interim central tank battery, before carrying the oil to an oil terminal.\textsuperscript{92} Depending on the location of the oil terminal and tank batteries, small pipelines may feed oil from production sites into main pipelines, or pipelines may connect one site to another in a more haphazard fashion.\textsuperscript{93} The pipes are traditionally smaller in diameter, constructed of steel, and operated at lower pressures.\textsuperscript{94}

Crude oil transmission pipeline systems connect domestic and foreign producers with U.S. refineries while also bringing domestically produced crude to the coast for export.\textsuperscript{95} Over 250 companies operate a loose grid of over 80,000 miles of crude oil transmission pipelines, which extends across much of the nation.\textsuperscript{96} The systems share many features with gas transmission systems. The largest crude oil systems use long-distance trunklines with only a few receipt

\begin{itemize}
  \item \textsuperscript{86} Id. at 4.
  \item \textsuperscript{87} Id.; TRENCH \& MIESNER, supra note 84, at 17–18.
  \item \textsuperscript{88} How Do Pipelines Work?, supra note 53.
  \item \textsuperscript{89} Pipeline Safety: Safety of Hazardous Liquid Pipelines, 80 Fed. Reg. 61,610, 61,612 (proposed Oct. 13, 2015) (to be codified at 49 C.F.R. pt. 195) (“Recent data indicates . . . that PHMSA regulates less than 4,000 miles of the approximately 30,000 to 40,000 miles of onshore hazardous liquid gathering lines in the United States.”).
  \item \textsuperscript{90} GAO, COLLECTING DATA AND SHARING INFORMATION, supra note 60, at 3; TRENCH \& MIESNER, supra note 84, at 13; Pipeline Safety: Safety of Hazardous Liquid Pipelines, 80 Fed. Reg. at 61,612.
  \item \textsuperscript{91} MIESNER \& LEFFLER, supra note 54, at 3; ENERGY \& ENVTL. RESEARCH CTR., UNIV. OF N.D., LIQUIDS GATHERING PIPELINES: A COMPREHENSIVE ANALYSIS 11 (Dec. 2015), https://undeerc.org/bakken/pdfs/EERC%20Gathering%20Pipeline%20Study%20Final%20Dec15.pdf [https://perma.cc/DJ7G-5J6D]. The fluid mixture extracted through the production well is treated on site to separate crude oil from water, gases, and impurities. Id.
  \item \textsuperscript{92} ENERGY \& ENVTL. RESEARCH CTR., supra note 91, at 11.
  \item \textsuperscript{93} See id. at 15–16 (showing maps of gathering systems).
  \item \textsuperscript{94} Id. at 79, 122, 128.
  \item \textsuperscript{96} Pipeline Mileage and Facilities, supra note 67 (follow “2010+ Pipeline Miles and Facilities” hyperlink) (reporting 80,748 miles of crude oil transmission pipelines and 267 operators in 2018, not including those who operate only tanks).
\end{itemize}
362  

**OHIO STATE LAW JOURNAL**  

[Vol. 81:3

and delivery points. The piping is constructed of steel, is wider in diameter, and is operated at higher pressures. Operators also use mechanical devices—pumps—to increase the pressure of oil before it enters the transmission system and to ensure continued flow. But there is one significant difference between the two: while gas systems carry one commodity, crude oil transmission systems generally transport multiple grades of oil. Based on the demands of refineries, operators “batch” the grades by sequentially injecting the fluids into the transmission pipeline and removing the batches once they reach their destinations.

Other liquid transmission systems transport two categories of commodities—refined products and hydrocarbon gas liquids—for both the domestic and import-export markets. Refined product pipeline transmission systems deliver the petroleum products made from crude oil at refineries to large users and fuel terminals. These products include jet fuel, diesel fuel, gasoline, and heating oil. Liquefied gas transmission systems transport hydrocarbon gas liquids—the natural gas liquids separated from methane at natural gas processing plants and the liquefied gases produced by refineries—to manufacturers and distributors. The grids are similar in size: 188 companies operate approximately 62,000 miles of refined product pipelines, while 192 companies operate approximately 70,000 miles of liquefied gas pipelines. Like crude oil transmission systems, these liquid systems generally transport or

---

97 See Trench & Miesner, supra note 84, at 17–18. Shorter transmission pipelines transport crude oil to nearby refineries or receiving points for trunklines. Id. at 17.


99 How Do Pipelines Work?, supra note 53 (stating that most crude oil transmission pipelines are 8 to 24 inches and a few are 48 inches in diameter).

100 See Miesner & Leffler, supra note 54, at 3, 70, 247.

101 See Trench & Miesner, supra note 84, at 12–15.

102 Id. at 15.

103 Petroleum Pipeline Systems, supra note 82; How Do Pipelines Work?, supra note 53.

104 How Do Pipelines Work?, supra note 53.

105 See Hydrocarbon Liquids Explained, U.S. ENERGY INFO. ADMIN., https://www.eia.gov/energyexplained/index.php?page=hgls_home [https://perma.cc/85TG-RQ2H] (last updated Oct. 31, 2019). For example, ethane is a natural gas liquid used as feedstock for plastics and chemicals, while propane is a liquefied petroleum gas used as fuel for heating and cooking. Id.

106 Pipeline Mileage and Facilities, supra note 67 (follow “2010+ Pipeline Miles and Facilities” hyperlink) (reporting 188 operators and 62,714.4 miles of refined products pipelines, and 192 operators and 70,268.6 miles of pipelines transporting “Highly Volatile Liquids (HVL), flammable, and toxic liquids,” which are generally liquefied hydrocarbon gases).
“batch” multiple commodities. But unlike the trunkline systems that carry oil over large distances, they are typically shorter in length and serve a region. The steel piping is also smaller in diameter, though operated at high pressure.

B. The Energy Revolution

Over the last fifteen years, a revolution in oil and gas production has reshaped the energy system of the United States. Development companies have combined two techniques—high-volume hydraulic fracturing and horizontal drilling—to extract large volumes of oil and natural gas from shale and other “tight” rock formations. These unconventional reservoirs underlie vast swaths of the country, including areas that have not traditionally produced oil and gas. The results are dramatic. From 2004 to 2018, the nation’s annual domestic gas production has increased by 64%, from 18.6 to 30.6 trillion cubic feet. In 2017, the United States produced so much natural gas that it became a net exporter, a feat that would have seemed impossible at the beginning of this century. Over the same period, annual domestic crude oil production rose from 2.0 to 4.0 billion barrels, an increase of 100%. In 2018, the nation became the top crude oil producer in the world and a net exporter of petroleum liquids—reclaiming energy independence for the first time in seventy-five years.

107 TRENCH & MIESNER, supra note 84, at 15.
108 Id. at 18–19.
109 Id. at 19.
110 RATNER & TIEMANN, supra note 30, at 1–7. In this type of development, a company drills vertically down to the target rock formation and then horizontally within the formation. GROUND WATER PROT. COUNCIL & ALL CONSULTING, MODERN SHALE GAS DEVELOPMENT IN THE UNITED STATES: A PRIMER ES-3 (Apr. 2009), https://www.energy.gov/sites/prod/files/2013/03/f0/ShaleGasPrimer_Online_4-2009.pdf [https://perma.cc/T8GY-H3G8]. The horizontal leg can extend two or more miles, which increases the surface area in contact with the well. Id. To stimulate production, the company then hydraulically fractures the rock by injecting large volumes of fluid and sand into the well under high pressure. Id. at ES-4. The fractures allow the oil or gas to flow out of the formation and up to the surface. Id. Several of these wells can be drilled on one well site, the horizontal legs spreading out in different directions to efficiently drain the reservoir. Id. at ES-3, ES-5.
111 See IHS GLOBAL INC., supra note 32, at 15–16.
113 IHS GLOBAL INC., supra note 32, at 14; EIA, ANNUAL ENERGY OUTLOOK 2019, supra note 32, at 14.
As production increased, so did the need for transport. The nation’s energy pipeline networks were originally built to carry oil and gas from conventional production areas to nearby users.\textsuperscript{116} When the reservoirs began to empty and it seemed that the United States would be ever more dependent on oil and gas from other countries, energy companies responded by investing in new infrastructure that would bring imports to processing facilities and consumers.\textsuperscript{117} Starting in the late 1970s, companies built large crude oil transmission pipelines to transport imported oil from Gulf Coast ports to the north and expanded the Canadian oil pipeline system to bring oil from Alberta to the south.\textsuperscript{118} U.S. oil refineries invested in technologies to process these imported heavy crude oils.\textsuperscript{119} Even at the beginning of this century, companies were constructing or restarting large import terminals on the East and Gulf coasts to accept liquefied natural gas by tanker vessel from countries as distant as Australia.\textsuperscript{120} The terminals were connected to interstate transmission pipelines to bring the gas to local distribution companies and large users.\textsuperscript{121}

But once the energy revolution began, pipeline companies that had pivoted away from domestic production towards imports suddenly faced new demands for transportation.\textsuperscript{122} Producers needed gathering pipelines to transport the oil and gas from thousands of wells in prolific shale plays to processing facilities and interstate transmission pipeline systems.\textsuperscript{123} This need was particularly acute in nontraditional production areas—such as the Marcellus and Utica shale gas plays in the Appalachian region—which contained little to no existing pipeline infrastructure.\textsuperscript{124} The gathering pipeline industry struggled to keep up with the number of wells drilled in these regions, frustrating development companies and

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{116} See MIESNER & LEFFLER, supra note 54, at 9–15 (chronicing the early history of oil and gas pipelining).
\item \textsuperscript{117} See IHS GLOBAL INC., supra note 32, at 12.
\item \textsuperscript{119} ANTHONY ANDREWS ET AL., CONG. RESEARCH SERV., R41478, THE U.S. OIL REFINING INDUSTRY: BACKGROUND IN CHANGING MARKETS AND FUEL POLICIES 16 (2010).
\item \textsuperscript{120} IHS GLOBAL INC., supra note 32, at 14.
\item \textsuperscript{121} See id. See generally MIESNER & LEFFLER, supra note 54, at 5–6 (describing natural gas pipeline value chain).
\item \textsuperscript{122} See IHS GLOBAL INC., supra note 32, at 15–19 (discussing the U.S. shale oil and gas revolution and the corresponding increase of capital investment in pipeline infrastructure).
\item \textsuperscript{123} See id.
\end{itemize}
\end{footnotesize}
slowing production. Yet even plays in more traditional production areas, such as the Bakken shale play in North Dakota, required significant expansions in gathering systems: one to transport the produced oil to oil terminals, and another to transport the gas that flowed to the surface with the oil to the rest of the gas pipeline network. By one estimate, approximately 28,000 miles of gas gathering pipelines and 20,000 miles of oil gathering pipelines were built between 2013 and 2017 alone.

Meanwhile, producers needed transmission pipelines to transport oil, gas, and hydrocarbon liquids from shale play regions to shifting markets. When production increased and prices dropped, domestic demand rose as well. The electric power sector turned more and more towards natural gas to generate electricity, encouraged by federal and state environmental policies and the retirement of aging coal-fired power plants. Residential, commercial, and industrial users consumed more natural gas—and distribution systems grew. Demand for hydrocarbon liquids—notably chemical feedstocks—increased. Producers also responded to declining prices by seeking to export oil and gas to markets abroad. Owners of liquefied natural gas import terminals proposed export facilities. Because domestic oil refineries had invested in processing equipment for imported heavy crude oil before the energy revolution, they had limited capacity to refine the light crude oil extracted from shale plays. After intense lobbying by producers, Congress cemented the shift in energy transport

---

125 See id. (stating that, at the time, over 1000 wells in northern Pennsylvania had been drilled but were unable to produce because of a lack of pipeline infrastructure).
128 ICF, NORTH AMERICA, supra note 35, at 88.
129 See id. at 3; see also Natural Gas Consumption by End Use, U.S. ENERGY INFO. ADMIN., https://www.eia.gov/naturalgas/data.php [https://perma.cc/GGK8-AYTX] (select “Consumption,” then “Total consumption”) (reporting that U.S. natural gas deliveries to the electric power sector roughly doubled between 2004 and 2018).
130 See Natural Gas Consumption by End Use, supra note 129; see also Both Natural Gas Supply and Demand Have Increased from Year-Ago Levels, U.S. ENERGY INFO. ADMIN. (Oct. 4, 2018), https://www.eia.gov/todayinenergy/detail.php?id=37193 [https://perma.cc/22MU-ABQH].
133 RATNER ET AL., supra note 132, at 3.
134 BROWN ET AL., supra note 132, at 5, 20; IHS GLOBAL INC., supra note 32, at 30.
from imports to exports in 2015 by lifting the ban on exports of crude oil that had been enacted forty years earlier in the wake of the Arab oil embargo.\textsuperscript{135}

Transmission pipeline companies responded to these significant changes in supply and demand in three ways. First, they sought to expand the footprint of transmission systems to connect new delivery and receipt points and to increase transportation capacity.\textsuperscript{136} Some projects were designed to lengthen existing pipeline systems,\textsuperscript{137} and others to construct new interstate pipelines.\textsuperscript{138} Second, they sought to transport the large volumes of oil and gas produced in shale plays by increasing the capacity of existing routes.\textsuperscript{139} Projects ranged from adding compression or pump stations to boost operating pressures,\textsuperscript{140} to constructing new pipelines parallel to older pipelines,\textsuperscript{141} to replacing older pipelines with new, higher capacity pipelines.\textsuperscript{142} Third, they sought to alter the purpose of existing transmission systems so the systems could carry products in the right


\textsuperscript{137}See id. (describing two projects, the Leach Xpress and Rayne Xpress, to expand the Columbia Pipeline natural gas transmission system).

\textsuperscript{138}See BROWN ET AL., supra note 132, at 17–18; see also id. at 17 n.80 (noting Enbridge’s proposed Sandpiper project, a new crude oil pipeline, which was not built after public opposition).


\textsuperscript{141}See Overview, ENERGY TRANSFER, https://marinerpipelinefacts.com/overview/ [https://perma.cc/V4R5-KRKA] (describing project to build two hydrocarbon liquids pipelines, primarily along the route of an existing pipeline).

direction. These projects included reversing the flow of commodities and converting pipelines to transport new products.

The result is a pipeline transportation system that is remaking itself, one project at a time. Oil and hydrocarbon liquids transmission companies have steadily expanded the systems’ footprint by building new, large pipelines that carry the commodities to export terminals or to be refined or processed. From 2004 to 2018, the total length of oil systems increased by almost 30,000 miles or 64%. One of these pipelines was the Dakota Access pipeline. Hydrocarbon liquids systems grew by approximately 18,000 miles or 36% during the same time period. Natural gas transmission systems added approximately 34,000 miles of new pipeline during this time, but the size of the gas transmission footprint declined slightly because the industry decommissioned older pipelines. In the last few years, the pace of new gas transmission pipeline construction has skyrocketed. In 2017 alone,

143 See Mike Kirkwood, Pipeline Reversals and Conversions: Case Studies and Best Practices, 242 PIPELINE & GAS J. 26, 26 (2015) (discussing the increase in pipeline flow-reversals caused by the boom in U.S. shale plays).
144 See BROWN ET AL., supra note 132, at 17 n.79 (noting the Seaway oil pipeline reversal project, which was completed in 2012).
148 Annual Report Mileage for Hazardous Liquid or Carbon Dioxide Systems, supra note 146 (showing that annual reported mileage grew from 51,794 to 70,267 miles for highly volatile liquid systems).

companies built 773 miles of new pipeline, most of which were designed to carry natural gas from the Marcellus and Utica shale plays to the Midwest.151 That same year, the federal government approved forty-nine more projects to construct 2739 miles of pipeline.152 The projects included largest-capacity gas pipeline ever approved: the 713-mile Rover pipeline, which transports gas from the Marcellus and Utica plays to Michigan and to market hubs in Canada.153

All predictions are that the domestic energy revolution is not over. Vast potential drilling areas and improving technology will likely spur more development.154 Annual natural gas production is expected to rise from 29.5 trillion cubic feet in 2018 to 43.4 trillion cubic feet in 2050, driven by production from the Marcellus and Utica shale plays in the northeast and the Eagle Ford and Haynesville shale plays in the Gulf Coast region.155 Development of unconventional resources is also expected to drive an increase in annual oil production, from 4.0 billion barrels in 2018 to a peak of 5.3 billion barrels in 2031, before production falls to 4.3 billion barrels in 2050.156 Production in the Permian Basin play in west Texas and eastern New Mexico is predicted to rapidly rise, while the Bakken shale play and plays in the Gulf Coast region will contribute smaller amounts.157

In this energy future, the pipeline industry will reshape itself even more to carry oil, gas, and petroleum products to consumers and export markets. From 2018 to 2035, the industry expects to invest $266 to $336 billion and build approximately 150,000 miles of new pipelines in the United States.158 This predicted buildout would be greater than the one that has already taken place.

The majority of the pipeline construction is expected to occur in the natural gas pipeline network.159 The industry predicts that gathering pipeline companies will add 73,500 miles of new gas gathering pipelines to the current network

152 Id.
153 Id.; Rover Pipeline LLC, 158 FERC ¶ 61,109 (2017); Rover Pipeline Project, ROVER PIPELINE, https://www.roverpipelinefacts.com/ [https://perma.cc/FU7N-FAFJ].
154 EIA, ANNUAL ENERGY OUTLOOK 2019, supra note 32, at 76 (stating that there are roughly 500,000 square miles of oil and gas resources in the United States).
155 Id. at 78; Annual Energy Outlook 2019: Oil and Gas Supply, U.S. ENERGY INFO. ADMIN., https://www.eia.gov/outlooks/aeo/data/browser/#/?id=14-AEO2019&cases=ref2019&sourcekey=0 [https://perma.cc/75MC-DGM7].
156 EIA, ANNUAL ENERGY OUTLOOK 2019, supra note 32, at 57–58.
158 See id. at 80, 88.
between 2018 and 2035—an increase of 20%.\textsuperscript{160} About half of this new construction is expected to occur in Gulf Coast plays, in conjunction with a rapid rise in production.\textsuperscript{161} During the same period, the industry predicts that the transmission grid will transport an additional 50 billion cubic feet per day of natural gas, half of which will be from the Marcellus and Utica shale plays.\textsuperscript{162} To carry this amount of gas, transmission companies will build 24,000 miles of new pipeline, an increase of 8%.\textsuperscript{163} There are different predictions as to where the pipelines will be constructed. The industry expects to build pipelines in the Northeast, South, and Gulf Coast regions, where the added capacity will primarily serve gas-fired power plants and export facilities.\textsuperscript{164} The federal government also predicts that the natural gas transmission grid will expand to transport increasing supplies from the Marcellus and Utica shale plays.\textsuperscript{165} But it expects that the buildout will occur in the Midwest, which will serve as a throughput for gas from the shale plays to the South and Gulf Coast.\textsuperscript{166}

The oil and hydrocarbon liquids pipeline network is expected to grow less overall because the industry has already invested in several large pipelines.\textsuperscript{167} But the Permian Basin is experiencing dramatic growth and will need many more gathering pipelines.\textsuperscript{168} The industry predicts that gathering pipeline systems will double between 2018 and 2035, as 36,000 miles of new pipelines are built in the Permian Basin and Gulf Coast plays.\textsuperscript{169} Oil transmission pipeline systems will expand by 6400 miles, or 8% of the current grid, during the same time period.\textsuperscript{170} These new pipelines will transport an additional 2.1 billion barrels of oil, primarily from the Permian Basin to refineries along the Gulf Coast.\textsuperscript{171} Pipeline capacity is also needed to transport imports of Canadian heavy crude oil and domestic oil from the Bakken shale through a central corridor to U.S. refineries.\textsuperscript{172} Construction on most of these projects is expected to occur between 2023 and 2028.\textsuperscript{173} Finally, the industry predicts that 8600 miles of hydrocarbon liquids pipeline will be built, primarily to transport liquids from production areas to manufacturers and the coast for export.\textsuperscript{174}

\textsuperscript{160}Id. at 88.
\textsuperscript{161}See id. at 146.
\textsuperscript{162}Id. at 37–38.
\textsuperscript{163}Id. at 88.
\textsuperscript{164}ICF, NORTH AMERICA, supra note 35, at 4–5, 125–26.
\textsuperscript{165}EIA, ANNUAL ENERGY OUTLOOK 2019, supra note 32, at 80.
\textsuperscript{166}Id.
\textsuperscript{167}ICF, NORTH AMERICA, supra note 35, at 156, 159.
\textsuperscript{168}Id. at 6.
\textsuperscript{169}Id. at 80, 85, 88.
\textsuperscript{170}Id. at 88.
\textsuperscript{171}Id. at 36–37.
\textsuperscript{172}Id. at 34, 36–37; IHS GLOBAL INC., supra note 32, at 28.
\textsuperscript{173}ICF, NORTH AMERICA, supra note 35, at 6.
\textsuperscript{174}Id. at 88.
III. THE LEGAL FRAMEWORKS GOVERNING RISK

Energy pipelines pose risks to the environment, health, and public safety. When a pipeline system releases gases or liquids, it can damage natural resources, harm human health, and injure or kill members of the public. Regulation of these risks is divided between decisions about the “safety” of a pipeline and decisions about the “siting” of a pipeline. The safety framework focuses on preventing and managing accidents from the sited pipeline. The framework is concentrated in federal law and consists of minimum standards for most types of pipelines. In contrast, the siting framework focuses on whether the pipeline should be built, the location of the pipeline, and the acquisition of property rights. The framework is fragmented by type of pipeline and decision-maker, resulting in a patchwork of policy approaches. The two frameworks are formally separate, leaving the relationship between the location of a pipeline and the long-term risk of a spill or release unaddressed.

A. The Risk of Energy Pipelines

The risk posed by energy pipelines is a function of two elements: (1) the probability—or likelihood—of a failure of the integrity of a pipeline system, and (2) the magnitude of the consequences of a release or spill of product to public safety, human health, and the environment. Energy pipelines pose a low-frequency, high-consequence risk. Accidents are relatively infrequent given how vast the networks are. Each year, energy pipelines cause an average of 295 significant accidents. Onshore accidents are distributed roughly evenly among the pipeline systems; the annual average ranges from sixty-eight accidents on gas transmission pipelines to eighty-four accidents on hydrocarbon

---

175 “Risk” is defined in different ways. At its most general, risk “refers to uncertainty about and severity of the events and consequences (or outcomes) of an activity with respect to something that humans value.” Terje Aven & Ortwin Renn, On Risk Defined as an Event Where the Outcome Is Uncertain, 12 J. RISK RES. 1, 6 (2009).


177 National Pipeline Performance Measures, Pipeline & Hazardous Materials Safety Admin., https://www.phmsa.dot.gov/data-and-statistics/pipeline/national-pipeline-performance-measures [https://perma.cc/2BF8-EE66] (follow “Significant Incidents” hyperlink) (reporting the annual average over the past ten years) (last updated Mar. 4, 2020). A significant incident is an accident that causes a “[f]atality or injury requiring in-patient hospitalization; [...] $50,000 or more in total costs, measured in 1984 dollars; [...] [h]ighly volatile liquid releases of 5 barrels or more or other liquid releases of 50 barrels or more; [or] [l]iquid releases resulting in an unintentional fire or explosion.” Id.
liquid transmission pipelines. By mile of pipeline, gas distribution pipelines are by far the safest. Energy pipeline accidents cause twelve deaths, seriously injure sixty-three people, and result in $671.6 million of property damage on average per year. But when a pipeline accident happens, catastrophic consequences are more likely to occur than would be predicted by a normal distribution of harm. In 2010, for example, a PG&E gas transmission pipeline exploded in a residential neighborhood of San Bruno, California, killing eight people, injuring fifty-eight, and causing $558 million in property damage. That same year, an oil transmission pipeline operated by Enbridge spilled over one million gallons of heavy crude oil into a creek near Marshall, Michigan. The oil flowed almost forty miles down the Kalamazoo River, causing more than $840 million in damage; cleanup took over four years.

There are many threats that could cause a pipeline to fail, including defects in the manufacturing of system components, incorrect construction or installation of the system, degradation of materials over time, mistakes in operation, or external human or natural forces. Over the last twenty years, the leading cause of accidents in the few regulated natural gas gathering pipelines

---

178 Id. (on average over the last ten years, onshore hydrocarbon liquids pipelines caused 84 accidents, gas distribution pipelines caused 65 accidents, crude oil transmission pipelines caused 65 accidents, gas transmission pipelines caused 58 accidents, and regulated gas gathering pipelines caused 2 accidents). But note that only a small percentage of gas gathering pipelines are regulated and thus required to report accidents.

179 Id.; Pipeline Mileage and Facilities, supra note 67.

180 National Pipeline Performance Measures, supra note 177 (reporting the average over the last ten years).


185 PHMSA, PIPELINE RISK MODELING, supra note 176, at 32–33.
is internal corrosion of the metal piping.\textsuperscript{186} These pipelines carry raw gas, which is more likely to contain contaminants that react with metal.\textsuperscript{187} In contrast, the primary cause of accidents in natural gas and liquid transmission pipelines is a failure in the physical components of the pipeline system—flaws in the construction or installation of a pipeline, cracks in the pipe, or malfunction of control equipment.\textsuperscript{188} These failures can occur in any pipeline system, but they are more likely to lead to accidents in transmission pipelines because the systems are complex and transport large volumes of commodities at high pressures. Finally, the leading cause of accidents in gas distribution pipelines is excavation damage, usually by third parties engaged in construction activities.\textsuperscript{189} This is because gas distribution pipelines run underneath streets and commercial and residential property in developed areas.\textsuperscript{190} The consequences of an accident depend on the commodity the pipeline is carrying, the volume of the product released, the rate at which the product disperses, and the people and environment in the affected area.\textsuperscript{191} Natural gas and some hydrocarbon liquids are flammable, so a rupture in a pipeline transporting one of these substances can result in an immediate explosion and fire, killing or injuring people and destroying property.\textsuperscript{192} If the hydrocarbon liquids are volatile but dense, the heavy gas can move along the surface of the ground and suffocate people before exploding.\textsuperscript{193} A liquid spill—whether of crude oil or other hydrocarbon liquids that remain liquefied—can flow over the ground, polluting surface waters, harming natural resources, and damaging property.\textsuperscript{194}

\textsuperscript{186}National Pipeline Performance Measures, supra note 177 (select “Gas Gathering” from the drop-down menu) (reporting that 47.5% of forty total incidents were caused by internal corrosion).


\textsuperscript{188}National Pipeline Performance Measures, supra note 177 (select “Gas Transmission” from the drop-down menu) (reporting that 36% of 1,098 total onshore incidents were caused by such failures); id. (select “Hazardous Liquid” from the drop-down menu) (36.9% of 2683 onshore incidents were caused by such failures).

\textsuperscript{189}Id. (select “Gas Distribution” from the drop-down menu) (reporting that 35% of 1439 total incidents were caused by excavation damage).


\textsuperscript{191}PHMSA, PIPELINE RISK MODELING, supra note 176, at 46–48.


\textsuperscript{193}Id. at 2.

\textsuperscript{194}Id.
the ground and contaminate groundwater.\footnote{Id.} Exposure to compounds in crude oil can also harm human health.\footnote{See, e.g., Mark D’Andrea & G. Kesava Reddy, \textit{Crude Oil Spill Exposure and Human Health Risks}, 56 \textit{J. OCCUPATIONAL \& ENVTL. MED.} 1029, 1029 (2014).}

Just as the domestic energy revolution has fundamentally changed oil and gas pipeline networks, so has it changed the landscape of risk. Each time a new pipeline is installed, more communities and natural resources are put at risk. In a commonly recognized phenomenon known as the “bathtub curve,” newer pipelines generally experience a higher accident rate than older pipelines because material and construction defects are likely to surface when pipelines first begin operation.\footnote{See, e.g., \textit{BUREAU OF SAFETY \& ENVTL. ENFORCEMENT, OVERVIEW OF RISK ASSESSMENT AND MANAGEMENT FOR OFFSHORE PIPELINES 6–7, https://www.bsee.gov/sites/bsee.gov/files/tap-technical-assessment-program//223aj.pdf} [https://perma.cc/TQ8W-LKEY]; \textit{PIPELINE SAFETY TR., HOW OLD IS TOO OLD? A LOOK AT AGING TRANSMISSION PIPELINE INFRASTRUCTURE ISSUES} (2015), http://pstrust.org/wp-content/uploads/2015/12/Weimer-Old-Pipes.pdf [https://perma.cc/PA9J-LUAQ].} But the risk profile of the pipelines themselves is also changing. To accommodate the increased flow of oil, gas, and hydrocarbon liquids, companies are building pipelines with wider diameters and operating them at higher pressures. Some natural gas gathering pipelines in shale plays are so large that they function similarly to gas transmission lines.\footnote{Pipeline Safety: Safety of Gas Transmission and Gathering Pipelines, 81 Fed. Reg. 20,722, 20,728 (proposed Apr. 8, 2016) (to be codified at 49 C.F.R. pts. 191, 192) (stating that “[p]roducers are employing gathering lines with diameters as large as 36 inches and maximum operating pressures up to 1480 psig”).} In the future, the industry expects to build larger gathering and transmission pipelines than it is building today—across every type of system.\footnote{ICF, NORTH AMERICA, \textit{supra} note 35, at 8.} The size of transmission pipelines, which are already wider than other types of pipelines, will increase the most.\footnote{\textit{Id.} at 8, 48 (predicting that the average size of oil, gas, and hydrocarbon liquids transmission pipelines will increase from 20.4 to 26.9 inches). “[M]ost pipeline capacity added in each of the scenarios is large pipe . . . . Some of the projects, particularly the oil projects transporting heavy crude oil from Western Canada into the U.S. and toward the Gulf Coast require very large pipe, each of which is upwards of 32-inches in diameter.” \textit{Id.}}

B. The Safety Framework

The federal Pipeline Safety Act is the primary legal framework governing the risk of accidents from gathering, transmission, and distribution energy pipelines.\footnote{\textit{Id.}} The Act grants the Pipeline and Hazardous Materials Safety Administration (PHMSA) broad authority to prescribe minimum safety

\begin{footnotes}
\footnote{Id.} \footnote{See, e.g., Mark D’Andrea & G. Kesava Reddy, \textit{Crude Oil Spill Exposure and Human Health Risks}, 56 \textit{J. OCCUPATIONAL \& ENVTL. MED.} 1029, 1029 (2014).} \footnote{See, e.g., \textit{BUREAU OF SAFETY \& ENVTL. ENFORCEMENT, OVERVIEW OF RISK ASSESSMENT AND MANAGEMENT FOR OFFSHORE PIPELINES 6–7, https://www.bsee.gov/sites/bsee.gov/files/tap-technical-assessment-program//223aj.pdf} [https://perma.cc/TQ8W-LKEY]; \textit{PIPELINE SAFETY TR., HOW OLD IS TOO OLD? A LOOK AT AGING TRANSMISSION PIPELINE INFRASTRUCTURE ISSUES} (2015), http://pstrust.org/wp-content/uploads/2015/12/Weimer-Old-Pipes.pdf [https://perma.cc/PA9J-LUAQ].} \footnote{Pipeline Safety: Safety of Gas Transmission and Gathering Pipelines, 81 Fed. Reg. 20,722, 20,728 (proposed Apr. 8, 2016) (to be codified at 49 C.F.R. pts. 191, 192) (stating that “[p]roducers are employing gathering lines with diameters as large as 36 inches and maximum operating pressures up to 1480 psig”).} \footnote{ICF, NORTH AMERICA, \textit{supra} note 35, at 8.} \footnote{\textit{Id.} at 8, 48 (predicting that the average size of oil, gas, and hydrocarbon liquids transmission pipelines will increase from 20.4 to 26.9 inches). “[M]ost pipeline capacity added in each of the scenarios is large pipe . . . . Some of the projects, particularly the oil projects transporting heavy crude oil from Western Canada into the U.S. and toward the Gulf Coast require very large pipe, each of which is upwards of 32-inches in diameter.” \textit{Id.}} \footnote{\textit{49 U.S.C. §§ 60101–41} (2012). As described in \textit{supra} Part II.A, only a small fraction of oil and gas gathering pipelines are regulated. See \textit{49 C.F.R. §§ 192.8–9, 195.11} (2019).}
standards for the transportation of gases and hazardous liquids by pipeline.\textsuperscript{202} These standards, which are adopted by rule, must be “practicable” and “designed to meet the need for . . . pipeline safety . . . and protecting the environment.”\textsuperscript{203} But the statute’s concept of “safety” is limited to the pipeline system: the design, installation, and construction of the new pipeline; the operation and maintenance of the pipeline once it is in the ground; the plans and procedures for a pipeline emergency; and the eventual replacement or abandonment of the pipeline.\textsuperscript{204}

The Act draws a bright line between “safety” and “siting” by specifically providing that it “does not authorize [PHMSA] to prescribe the location or routing of a pipeline facility.”\textsuperscript{205} The standards contain only one modest restriction on siting—an admonition that a right-of-way for a hazardous liquid pipeline “must be selected to avoid, as far as practicable, areas containing private dwellings, industrial buildings, and places of public assembly.”\textsuperscript{206} In recognition that this restriction will not prevent operators from continuing to locate pipelines near buildings, the standard also provides that the operator must bury the liquid pipeline more deeply if it is located within fifty feet of a private dwelling or an industrial building or place of public assembly “in which persons work, congregate, or assemble.”\textsuperscript{207}

As the domestic energy revolution has radically reshaped the pipeline networks during the last fifteen years, the approach to risk has remained the same: to “keep the product in the pipe.” Rather than limit where pipelines can be sited, the safety framework focuses on protecting the area around the pipeline after the fact. PHMSA does not approve new pipeline projects or expansions in a pipeline system.\textsuperscript{208} Instead, the materials, design, and construction of a new pipeline and its components are governed by technical engineering standards.\textsuperscript{209} The standards include specific prescriptive requirements, such as a design

\textsuperscript{202} 49 U.S.C. §§ 60101(a)(18)–(19), 60102 (2012). A “hazardous liquid” includes all petroleum products, such as crude oil and hydrocarbon liquids, as well as “nonpetroleum fuel” and liquid substances that “may pose an unreasonable risk to life or property” when transported in pipelines. Id. § 60101(a)(4). While the safety standards for gas and hazardous liquid pipelines are separate, they have many similarities.

\textsuperscript{203} Id. § 60102(b)(1).

\textsuperscript{204} Id. § 60102(a)(2)(B).

\textsuperscript{205} Id. § 60104(e).

\textsuperscript{206} 49 C.F.R. § 195.210(a) (2019).

\textsuperscript{207} Id. § 195.210(b).

\textsuperscript{208} The operator may apply for a special permit if it would like an exception or modification to the rules. 49 U.S.C. § 60118(c) (2012); 49 C.F.R. § 190.341 (2019). In 2017, PHMSA responded to the energy revolution by requiring companies to notify it of pipeline construction projects or other operational changes such as a change in product or reversal of the flow of a product. Pipeline Safety: Operator Qualification, Cost Recovery, Accident and Incident Notification, and Other Pipeline Safety Changes, 82 Fed. Reg. 7972, 7973 (Jan. 23, 2017) (to be codified at 49 C.F.R. pts. 190–92, 195, 199).

\textsuperscript{209} See, e.g., 49 C.F.R. §§ 192.51–.329, 195.100–.264 (2019).
formula for steel pipe that determines the maximum safe operating pressure. They also include general standards of performance—such as requiring the pipe to be sufficiently thick to withstand the internal and external pressures on the pipeline system.

Once the pipeline is in the ground, PHMSA is prohibited from applying design, installation, and construction standards to existing pipelines—thus ensuring that a company need not reconstruct pipelines in the ground to current standards. The operation and maintenance of the pipeline is primarily governed by management directives. The standards require operators to create and follow a written manual of procedures for each task, as well as to develop programs to address particular issues, such as training of employees and monitoring of pipelines in control centers. To address the risk of third-party excavation damage, the standards require pipeline operators to mark the location of their pipelines as part of “call-before-you-dig” programs. If the technical and management standards fail to keep the product in the pipe, the safety framework relies on operators’ emergency response planning to mitigate the damage. Operators must have emergency procedures that include a “prompt and effective response” to notice of the accident and the necessary actions to minimize the spill or release and public exposure to injury.

The safety framework responds to the risks associated with siting a pipeline in a certain location by focusing more attention on “high-consequence” areas. These areas are limited to ones where an accident could cause catastrophic harm. For gas pipelines, the standards focus on densely settled areas near the pipeline, where an explosion would cause the most fatalities and injuries. The definition is so narrow that it only captures 6.9% of the gas transmission pipeline grid. For hazardous liquid pipelines, the standards focus on municipal census tracts, commercially navigable waterways, and “unusually sensitive areas”—drinking water sources and certain habitats of imperiled, threatened, or endangered species.

---

211 Id. §§ 192.103, 195.112(a).
214 Id. §§ 192.605, .615, 195.402–.403.
215 Id. §§ 192.801–.809, 195.501–.509.
216 Id. §§ 192.614, 195.442.
217 Id. §§ 192.605(a), (e), .615, 195.402(a), (e).
218 Id. § 192.903 (defining “high-consequence area[s]” to include areas near the pipeline where multi-story buildings are prevalent or there are forty-six or more buildings; or where the potential impact circle contains twenty or more buildings or a regularly occupied structure such as a playground, church, or hospital). The standards use different terms for protected areas: class locations and high-consequence areas. See id. §§ 192.5, .903. The approach to risk, however, is the same.
endangered species. This definition is more expansive than the one for gas pipelines; approximately 42% of the total hazardous liquid transmission pipeline grid lies within these areas or is near enough that a spill could affect the areas. The difference can largely be explained by the use of census tracts, not by the inclusion of natural resource areas. The standards focus on a remarkably narrow list of environmental features.

The framework seeks to protect high-consequence areas in three different ways. First, it regulates more pipelines in these areas. The safety standards apply to gathering pipelines only if the pipelines are in or could affect the protected areas. In practice, this threshold is quite high. The vast majority of gathering pipelines built during the domestic energy revolution are unregulated because they are located in rural areas where the oil and gas wells are located. Second, when companies site new gas transmission pipelines in protected areas, they must comply with additional construction and operation requirements. As the number of buildings in the area around a gas transmission pipeline increases, for example, the pipeline must be able to withstand more pressure and use more closely spaced safety valves. Pipelines in navigable rivers, streams, and harbors must be buried at least forty-eight inches in soil or twenty-four inches in consolidated rock. And third, the framework requires operators of pipelines in protected areas to improve their risk management through special integrity management programs. In these programs, operators must assess the condition of the pipelines, identify threats to the lines, analyze the risks, and take “prompt action” to address defects and other anomalous conditions.

The states may choose to impose additional “compatible” safety requirements on intrastate pipelines, if they have a pipeline safety program...

---

222 See Pipeline Safety: Areas Unusually Sensitive to Environmental Damage, 65 Fed. Reg. 80530, 80534 (Dec. 21, 2000) (explaining that the agency chose not to include many natural resources in its definition of “unusually sensitive areas” and that the definition would concentrate on “areas that are most susceptible to permanent or long-term damage”).
224 Id. §§ 192.9, 195.1–2, .11.
227 Id. § 192.327(e).
228 Id. §§ 192.901–.951, 195.452. Operators of gas distribution pipeline systems are also required to create integrity management programs; the programs apply to the entire system because all of the pipelines are generally in populated areas. Id. §§ 192.1001–.1015.
229 Id. §§ 192.911, 195.452.
certified by PHMSA.\textsuperscript{230} All states except Alaska and Hawaii regulate the safety of intrastate gas pipelines.\textsuperscript{231} Just one-third of the states regulate the safety of intrastate hazardous liquid pipelines; this group includes some states with productive oil plays—but not, it must be noted, North Dakota.\textsuperscript{232} States generally retain the bright-line distinction between safety and siting in their delegated programs; indeed, several state legislatures have enacted an identically worded provision that prevents the delegated agency from prescribing the route of a pipeline.\textsuperscript{233} While the states have adopted numerous requirements that exceed the federal minimum standards—by one report there are over 1300 such requirements—the scope of the standards is largely limited to enhanced reporting, recordkeeping, and testing.\textsuperscript{234} Many states, for example, require operators to prepare more detailed maps of their systems than is mandated by PHMSA.\textsuperscript{235} Texas requires pipeline operators to obtain a permit from the state and to submit maps of systems in digital shape files.\textsuperscript{236} The states that do regulate the location of pipelines through their safety programs impose relatively minor restrictions.\textsuperscript{237}

C. The Siting Framework

The legal framework governing the siting of new energy pipelines or the expansion of existing pipeline systems is a scattered collection of policy

\textsuperscript{230} 49 U.S.C. §§ 60104(c), 60105 (2012).
\textsuperscript{232} The states that regulate intrastate hazardous liquid pipelines are Alabama, Arizona, California, Indiana, Louisiana, Maryland, Minnesota, New York, New Mexico, Oklahoma, Pennsylvania, Texas, Virginia, Washington, and West Virginia. See id. at 2.
decisions, divided across jurisdictions and fragmented by the type of pipeline, the product being transported, and whether the pipeline crosses state borders. States have authority under their police powers to regulate the route of gathering pipelines and gas distribution pipelines, but few states have chosen to do so. In effect, the decisions about the location of these pipelines—and the risks that a pipeline accident poses to communities and the environment—are left to the companies who will operate the lines. In contrast, the Federal Energy Regulatory Commission (FERC) approves the siting of interstate natural gas transmission pipelines, and some states have chosen to regulate the siting of other types of transmission pipelines.

Given PHMSA’s authority over the safety of pipelines, the issue of whether other federal or state entities have the authority to consider the risks of an accident in their siting decisions is contested. At the federal level, FERC acknowledges PHMSA’s exclusive authority to set safety standards but claims the authority to impose conditions on interstate natural gas pipelines that “mitigate the impact of construction or operation on the environment.” As to states, the Pipeline Safety Act provides that states “may not adopt or continue in force safety standards” for interstate pipelines and certified states may only adopt additional or more stringent safety standards if they are “compatible” with federal minimum standards. There are no court decisions directly on point, but a state likely has the authority to regulate the siting of an interstate or intrastate pipeline based on the risk of an accident. Federal “safety standards” only apply to aspects such as design, construction, and operation—and PHMSA is not authorized to prescribe the location or route of a pipeline. But a state is likely preempted from imposing “safety standards” on the construction or operation of an interstate pipeline as a condition of siting approval, and may

241 See Wash. Gas Light Co. v. Prince George’s Cty. Council, 711 F.3d 412, 422 (4th Cir. 2013) (“Even if we were to find that the [Pipeline Safety Act] has preemptive effect beyond the express preemption provision . . . we would not conclude that Congress intended the [law] to occupy the field of natural gas facility siting.”); ANR Pipeline Co. v. Iowa State Commerce Comm’n, 828 F.2d 465, 473 (8th Cir. 1987) (noting that the state “may be able to enact legislation to protect its valuable topsoil and other aspects of the environment” but that the issue was not before the court); Portland Pipe Line Corp. v. South Portland, 288 F. Supp. 3d 321, 430–31 (D. Me. 2017) (stating that “Congress did not intend the [Pipeline Safety Act] to preempt state and local authority ‘to prescribe the location or routing of a pipeline facility,’” that “[u]nder their police power, states and localities retain their ability to prohibit pipelines altogether in certain locations,” and that “it is unlikely Congress intended to remove [a lesser restriction than a ban]”).
242 See Kinley Corp. v. Iowa Util. Bd., 999 F.2d 354, 359 (8th Cir. 1993) (stating that “[t]his Congressional grant of exclusive federal regulatory authority [over interstate pipelines] precludes state decision-making in this area altogether and leaves no regulatory
be preempted from doing so as to intrastate pipelines if it is not certified to regulate the safety of those lines.  

The siting of gathering pipelines, like the safety of the lines, is almost entirely unregulated. In oil and gas producing states, state commissions generally oversee gathering pipelines as part of their well permitting programs. This oversight does not extend to siting, however. Twenty-eight of the thirty states that assert jurisdiction over gathering pipelines do not require companies to obtain approval from the state commission to site gathering pipeline systems. These states accept the risks of the pipelines without regard to their location—presumably because they deem the risks of gathering pipelines to be so low that the location does not matter. Even the two states that require gathering pipeline companies to seek approval before constructing projects do not include a standard on risk in their siting criteria.

As companies have installed larger gathering pipelines and operated them at higher pressures, the states’ approach to risk has remained the same. The location of gathering systems is not regulated in states where much of the rapid build out has occurred during the domestic energy revolution, such as North Dakota, Ohio, Pennsylvania, and West Virginia. Neither is the siting of systems in states where companies are expected to build many more gathering pipelines in the future: Louisiana, New Mexico, and Texas. The states’ only concession to providing room for the state to either establish its own safety standards or supplement the federal safety standards”.

---

243 See Olympic Pipe Line Co. v. Seattle, 437 F.3d 872, 880 (9th Cir. 2006) (holding that only certified state or local government agencies may regulate the safety of intrastate pipelines); United Steelworkers of Am. v. Skinner, 768 F. Supp. 30, 35 (D.R.I. 1991) (same).


245 The states are Alaska, Arkansas, California, Colorado, Delaware, Illinois, Indiana, Kansas, Louisiana, Maryland, Michigan, Mississippi, Montana, Nebraska, Nevada, New Mexico, New York, North Dakota, Ohio, Oklahoma, Pennsylvania, South Dakota, Tennessee, Texas (unregulated pipelines), Utah, Virginia, West Virginia, and Wyoming. See id. app. A at A-3 to -227 (providing the oil and gas gathering line rules and regulations for each state).

246 See ALA. ADMIN. CODE r. 400-1-8-.03 (2000) (requiring operators to obtain approval for “[a]ll intrastate gathering lines, located in a rural location” and specifying that information on the “[l]ocation, route and length of line” must be provided); 805 KY. ADMIN. REGS. 1:190 (2019) (requiring operators to obtain approval for all oil and gas gathering lines that are not federally regulated and specifying that the “approximate locations of property lines, dwellings, environmentally sensitive features and road and stream crossings along the path of the gathering line” must be provided).

247 As part of the state’s pipeline safety program, Texas requires regulated gathering pipelines to obtain a permit. 16 TEX. ADMIN. CODE §§ 3.70(a), 8.1(a)(1)(B), 8.1(b)(4) (2017). The state does not, however, approve the route of the pipelines.
to risk is to require operators to report the location of the systems. The siting of natural gas distribution pipelines is also generally not regulated. Investor-owned local distribution companies are governed by state public utility laws. As public utilities, the companies are expected to provide safe and adequate service at reasonable rates. Several states specifically require utilities to obtain a “certificate of convenience and public necessity” before construction of any facilities. Even so, the focus of certification is on the need for the service, the adequacy of the service, and the reasonableness of the cost of infrastructure—not on the location of pipelines. And states do not require utilities to obtain a certificate for expansions of existing systems within a service area; at most, a company might be required to obtain a city’s consent to construct pipelines under its public roads. Municipally owned utilities, meanwhile, oversee their own pipeline systems and routes. Leaving decisions about the risk of a route to distribution utilities presumes that the utility is best able to design its system, the possible routes are limited by the need to service customers, and distribution pipelines are operated at low pressures. Once again, the domestic energy revolution has not altered this perception of the risk—even as distribution utilities build more pipelines and transport more gas.

---

248 See, e.g., N.D. CENT. CODE § 38-08-26 (2014) (requiring all operators of underground gas and liquid gathering lines to submit geographic information system (GIS) data to the state within 180 days of putting the pipeline into service); OHIO ADMIN. CODE 4901:1-16-15 (2019) (requiring operators of gas gathering pipelines to notify the state prior to construction and after construction of the pipeline route); see also N.Y. COMP. CODES R. & REGS. tit. 16, § 255.9(c) (2019) (requiring operators of gas gathering pipelines to notify the state prior to construction).


251 Id.

252 See, e.g., ARIZ. REV. STAT. ANN. § 40-281(A) (2011) (“A public service corporation . . . shall not begin construction of . . . plant, service or system, or any extension thereof, without first having obtained from the commission a certificate of public convenience and necessity.”).

253 See, e.g., GA. CODE ANN. § 46-4-25(a) (2004) (listing criteria for a certificate that include the demand for the natural gas, the economic feasibility of the system, the propriety of the costs, and the effects on other distribution systems).

254 See, e.g., ARIZ. REV. STAT. ANN. § 40-281(B) (2011) (“This section shall not require such corporation to secure a certificate for an extension within a city, county or town within which it has lawfully commenced operations, or for an extension into territory either within or without a city, county or town, contiguous to its . . . plant or system, and not served by a public service corporation of like character, or for an extension within or to territory already served by it, necessary in the ordinary course of its business.”).

255 See, e.g., MD. CODE ANN., PUB. UTIL. § 7-102(b) (LexisNexis 2010).

256 See, e.g., ARIZ. CONST. art. 15, § 2 (defining “public service corporations” as “[a]ll corporations other than municipal engaged in furnishing gas . . . for . . . fuel”).
In comparison to gathering and distribution pipelines, the location of transmission pipelines is more likely to be regulated. Authority over the siting of transmission pipelines is divided between the federal government and the states. At the federal level, the location of all interstate natural gas transmission pipeline projects must be approved by FERC. Jurisdiction over the location of other types of transmission lines—intrastate gas transmission lines and both interstate and intrastate crude oil and hydrocarbon liquid lines—is left to the states. Most of the siting laws were enacted well before the domestic energy revolution. FERC has exercised authority over interstate natural gas transmission projects since 1938. In the following decades, some states expanded their traditional authority over public utilities or common carriers to include the location of transmission pipelines. Others began to regulate transmission pipelines as “energy facilities” under general siting acts.

The states have split evenly on whether to approve transmission pipeline projects. Half of the states do not require pipeline companies to obtain approval for any type of transmission project; half require approval of at least one type of transmission pipeline system. More states regulate intrastate gas transmission projects than liquid transmission projects: twenty-three states require companies to obtain permission for intrastate gas transmission

---

258 FERC’s authority over interstate crude oil and hydrocarbon liquid pipelines is limited to economic regulation, which allows states to regulate the siting of these pipelines. See Interstate Commerce Act, 49 U.S.C. app. § 6 (1988).
259 Natural Gas Act, § 1(b), 52 Stat. 821, 821 (1938).
260 See, e.g., ALASKA STAT. § 42.06.240(a) (2018) (requiring oil and gas pipeline carriers to obtain approval for pipeline construction as of January 1, 1974).
262 Twenty-five states do not specifically require approval for transmission pipeline projects, although public utility commissions may decide to review such projects if they are constructed by local distribution companies. The states are Alabama, Arizona, Colorado, Delaware, Hawaii, Idaho, Indiana, Kansas, Kentucky, Louisiana, Maine, Maryland, Mississippi, Missouri, New Jersey, New Mexico, North Carolina, Oklahoma, Pennsylvania, South Carolina, Tennessee, Texas, Utah, West Virginia, and Wisconsin. The District of Columbia also does not require approval. In Colorado, county governments may choose to designate certain gas transmission pipeline projects as an “activity of state interest” and require approval before construction. COLO. REV. STAT. §§ 24-65.1-104, -201 (2019). In Texas, transmission pipelines are required to obtain a permit from the state under the safety program but there is no siting review. 16 TEX. ADMIN. CODE §§ 3.70, 8.1 (2017).
projects,\textsuperscript{264} while sixteen states require the same for interstate and intrastate liquid transmission projects.\textsuperscript{265} Only two states—Georgia and Nebraska—have passed new siting laws to respond to the expansion in pipeline networks caused by the domestic energy revolution; both laws govern liquid transmission pipelines.\textsuperscript{266}

\textsuperscript{264} ALASKA STAT. §§ 42.06.240(a), .630 (2018); ARK. CODE ANN. §§ 23-18-503(6)(C), -510(a) (2015) (greater than one mile and at least 125 psi); CAL. PUB. UTIL. CODE §§ 1001, 1002.5 (West 2010) (public utilities that add new natural gas capacity to the state); CONN. GEN. STAT. §§ 16-50i, -50k (2013) (at least 200 psi and design capacity of 20% of specified minimum yield strength); FLA. STAT. § 403.9405 (2015) (at least 15 miles in length and crosses a county line); 220 ILL. COMP. STAT. 5/15-401 (2013); IOWA CODE ANN. § 479.5 (West 2019) (at least 5 miles in length and more than 150 psi); MASS. GEN. LAWS ch. 164, §§ 69G, 69J (2015) (more than 1 mile in length and 100 psi); MICH. COMP. LAWS ANN. § 483.109 (West 2008); MINN. STAT. ANN. § 216G.02 (West 2010) (more than 275 psi); MONT. CODE ANN. §§ 75-20-104(9)(b), -201 (2019) (greater than 50 miles in length and 25” in diameter); NEV. REV. STAT. §§ 704.860, .865 (2014) (outside incorporated city); N.H. REV. STAT. ANN. §§ 162-H:2(VII)(a), .5(I), .10-b (2014); N.Y. PUB. SERV. LAW §§ 120(2), 121 (McKinney 2019) (more than 1000 feet in length and more than 125 psi); N.D. CENT. CODE § 49-22.1-017(a), .04 (2014) (at least one mile and more than 4.5” in diameter); OHIO REV. CODE ANN. §§ 4906.01(B)(1)(c), .04 (West Supp. 2019) (more than 500 feet in length, 9” in diameter, and 125 psi); OR. REV. STAT. §§ 469.300(11)(a)(E)(ii), 320 (2017) (at least five miles in length and 16” in diameter); 42 R.I. GEN. LAWS §§ 42-98-3, -4 (2006); S.D. CODIFIED LAWS §§ 49-41B-2.1(2), -4 (2004) (design capacity of at least 20% of specified minimum yield strength, exception for pipes less than 4” diameter, and one mile or less is constructed outside of public right-of-way); VT. STAT. ANN. tit. 30, § 248(a)(3) (2017); VA. CODE ANN. § 56-265.2-2:1 (2012) (owned by public utility); WASH. REV. CODE ANN. §§ 80.50.020(21)(b), .060 (West 2001) (at least 15 miles in length and greater than 14” in diameter that delivers to distribution facility); WYO. STAT. ANN. §§ 37-1-101(a)(vi)(G), -2-205 (2019).

\textsuperscript{265} ALASKA STAT. § 42.06.240(a) (2018); GA. CODE ANN. §§ 12-17-2, 22-3-83 (2019); 220 ILL. COMP. STAT. 5/15-401 (2013); IOWA CODE ANN. §§ 479.5, 479B.4 (West 2009) (at least 5 miles in length and more than 150 psi); MASS. GEN. LAWS ch. 164, §§ 69G, 69J (2015) (more than 1 mile in length); MICH. ADMIN. CODE r. 792.10447(1)(c) (2010); MINN. STAT. ANN. § 216G.02 (West 2010) (at least 6” diameter); MONT. CODE ANN. §§ 75-20-104(9)(b)(i), -201 (2019) (greater than 50 miles in length and 25” in diameter); NEB. REV. STAT. ANN. § 57-1404(2), -1405(1) (LexisNexis 2014) (greater than 6” diameter); N.H. REV. STAT. ANN. §§ 162-H:2(VII)(a), .5(I) (2014); N.D. CENT. CODE § 49-22.1-017(a), .04 (2014) (at least one mile and more than 4.5” in diameter); OR. REV. STAT. §§ 469.300(11)(a)(E)(i), 320 (2017) (at least five miles in length and 6” in diameter); 42 R.I. GEN. LAWS §§ 42-98-3(d), -4 (2006) (“facilities associated with the transfer of oil . . . via pipeline”); S.D. CODIFIED LAWS §§ 49-41B-2.1(2), -4 (2004) (design capacity of at least 20% of specified minimum yield strength, exception for pipes less than 4” diameter, and one mile or less is constructed outside of public right-of-way); WASH. REV. CODE ANN. §§ 80.50.020(21)(a), .060 (West 2001) (at least 15 miles in length and greater than 6” diameter); WYO. STAT. ANN. §§ 37-1-101(a)(vi)(G), -2-205 (2019). In Oklahoma, companies may choose to apply to the state commission for an order authorizing “the siting, construction, expansion, or operation of a crude oil or refined petroleum product pipeline facility.” OKLA. STAT. ANN. tit. 52, § 67(D) (2011).

\textsuperscript{266} Phil McKenna, \textit{Property Rights Outcry Stops Billion-Dollar Pipeline Project in Georgia}, INSIDE CLIMATE NEWS (Apr. 1, 2016), https://insideclimatene.org/news/
This comprehensive analysis of federal and state siting laws reveals that pipeline siting laws take two different approaches to risk: (1) a public utility approach that views risk as one factor in a broad inquiry to determine whether the pipeline is in the public interest, or (2) a facility approach that views risk as an issue of the acceptability of the pipeline, separate from the question of public need. These approaches in turn reveal two different conceptions of energy infrastructure: in the first, energy infrastructure is a conduit for the beneficial provision of energy with inherent risks; in the second, it is a large land use that provides benefits but also has significant side effects that can be minimized.

The first type of siting law draws on public utility concepts to holistically evaluate whether the project is in the public interest. The risk of pipeline accidents to communities and the environment is treated as a factor in a larger determination of public good. FERC uses this approach to determine whether to approve an interstate natural gas transmission pipeline project. In deciding whether the project “is or will be required by the present or future public convenience and necessity,” FERC must weigh the public benefits against the adverse effects of the project. The balancing test includes the effects of the pipeline on the surrounding community and potential disruption of the environment—both factors that could include the risk of accidents. FERC may also set “reasonable terms and conditions” on its approval.

Of the twenty-five states that approve pipeline projects, eight use a similar public utility approach. Like FERC, the states apply a balancing test to determine whether a pipeline is required by the “public convenience and necessity” or the “public interest.” Some states explicitly include negative

---

2069 Id.
2071 The states are Alaska, California, Illinois, Iowa, Michigan, Nebraska, Virginia, and Wyoming. See infra note 272.
2072 ALASKA STAT. § 42.06.270(a) (2018); CAL. PUB. UTIL. CODE §§ 1001, 1002.5 (West 2010); 220 ILL. COMP. STAT. 5/15-401 (2013); IOWA CODE ANN. § 479.12 (West 2009); MICH. COMP. LAWS ANN. § 483.109 (West 2015); NEB. REV. STAT. ANN. § 57-1407(4)
effects on the environment and “public safety” among the balancing factors.\textsuperscript{273} In Illinois, for example, the law directs the commission to consider the environmental impact, the impact on natural resources, and the impact on public safety of the project.\textsuperscript{274} Other states grant decision-makers the authority to impose conditions to address negative effects, thereby implicitly including these effects in the balancing test.\textsuperscript{275} In Iowa, the board has authority to impose conditions on intrastate pipelines as to “safety requirements” that are “determined by it to be just and proper.”\textsuperscript{276} But the only state that directly addresses the risk of an accident in its siting law—Nebraska—prohibits its commission from considering the “risk or impact of spills or leaks” in its determination that the route is in “the public interest.”\textsuperscript{277}

While the public utility approach considers environmental effects and public safety, it presumes that the risk of an accident is a generic factor inherent to the project—that is, an issue of whether the product will remain in the pipe—rather than an issue that can be proactively solved through a land use decision. FERC and the states that use this approach view the risk of accidents across the entire pipeline as acceptable and rely on the safety framework, rather than the location of the pipeline, to protect vulnerable areas. For example, FERC responded to a landowner’s concern about the safety of the Rover Pipeline, the largest-capacity natural gas pipeline ever considered by the Commission, by stating that the “minimal number of incidents distributed over more than 300,000 miles of natural gas transmission pipelines indicates a low risk for an incident at any given location.”\textsuperscript{278} FERC requires pipeline operators to certify that the project complies with PHMSA’s safety standards but does not otherwise seek to reduce the risk of accidents.\textsuperscript{279} Similarly, states that weigh public safety

\begin{footnotesize}(LexisNexis 2014); VA. CODE ANN. § 56-265.2(A)(1) (2012); WYO. STAT. ANN. § 37-2-205 (2019).\end{footnotesize}

\begin{footnotesize}\textsuperscript{273} CAL. PUB. UTIL. CODE § 1002(a)(4) (West 2010); 220 ILL. COMP. STAT. 5/15-401(c) (2013); NEB. REV. STAT. ANN. § 57-1407(4) (LexisNexis 2014); VA. CODE ANN. § 56-265.2:1(A) (2012). In Wyoming, the statute is unclear, but the rules require applicants to provide information on the facility and site. 2-3 WYO. CODE R. § 21 (LexisNexis 2019).\end{footnotesize}

\begin{footnotesize}\textsuperscript{274} 220 ILL. COMP. STAT. 5/15-401(b)(1), (3), (4) (2013).\end{footnotesize}

\begin{footnotesize}\textsuperscript{275} IOWA CODE ANN. § 479.12 (West 2009); VA. CODE ANN. § 56-265.2:1(A) (2012); see also ALASKA STAT. § 42.06.240(d) (2018) (the state commission may impose conditions “necessary for the protection of the environment” and for the “best interest” of the oil or gas pipeline and the general public).\end{footnotesize}

\begin{footnotesize}\textsuperscript{276} IOWA CODE ANN. § 479.12 (West 2009); cf. id. § 479B.9 (authorizing conditions on interstate hazardous liquid transmission pipelines only “as to location and route”).\end{footnotesize}

\begin{footnotesize}\textsuperscript{277} NEB. REV. STAT. ANN. § 57-1407(4) (LexisNexis 2014). According to the state’s law, the purpose of this provision is to “acknowledge[] and respect[] the exclusive federal authority over safety issues established by the federal . . . Pipeline Safety Act . . . and the express preemption provision stated in that act.” Id. § 57-1402(2) (2014). As discussed supra notes 239–43 and accompanying text, this is an overly restrictive interpretation of the state’s authority.\end{footnotesize}

\begin{footnotesize}\textsuperscript{278} Rover Pipeline LLC, 158 FERC ¶ 61,109, at 85 (2017).\end{footnotesize}

\begin{footnotesize}\textsuperscript{279} 18 C.F.R. § 157.14(a)(10)(vi) (2019) (requiring an applicant to “certify that it will design, install, inspect, test, construct, operate, replace, and maintain the facilities for which
in a siting decision view the risk as negligible and defer to federal safety standards rather than address the potential harm through the route. Illinois\textsuperscript{280} and Iowa,\textsuperscript{281} for example, compared the generic risks of pipeline transportation of crude oil against rail transportation and determined that the Dakota Access Pipeline was much safer. The states ultimately relied on PHMSA’s standards and the pipeline operator’s private commitments to protect the public and the environment.\textsuperscript{282}

The second type of siting law draws on land use planning concepts to recognize the significant impacts of locating and operating energy facilities on communities and the environment. In this approach, the risk of a pipeline is treated as a distinct issue that must be addressed before siting, separate from the benefits of the pipeline. The decision-maker—which may be a public utility commission or a designated siting board or council\textsuperscript{283}—must consider the public need for and the effects of the project. But rather than weigh the benefits against the risks, the decision-maker must determine that the pipeline fulfills a public need \textit{and} that the negative effects of the pipeline are acceptable. Seventeen states utilize the facility approach for the pipelines they regulate, making this approach the more popular of the two approaches to risk. Most of the states adopted general energy facility siting acts in the 1970s and included
the “hazard” of a potential pipeline presume that the line either poses an undue hazard or it does not. Thus, the facility approach does not encourage decision-makers to consider the siting of linear infrastructure as a series of risk decisions that shift based on the potential consequences of an accident. Nor does the approach necessarily create a proactive response to risk by empowering the decision-maker to choose the best route. South Dakota’s law, for example, prohibits its commission from establishing a route.290

A few states that utilize the approach seek to protect vulnerable areas from the consequences of a pipeline accident. In Florida, the state board certifies a corridor for a gas transmission pipeline after considering whether the pipeline “[a]void[s] densely populated areas to the maximum extent feasible.”291 Similarly, in Minnesota, the commission considers criteria such as the “existence of populated areas” in designating a pipeline route.292 North Dakota is the only state to explicitly protect certain areas. The route of a pipeline may not traverse any “exclusion” areas such as national and state parks, historic sites, and wilderness areas.293 The company must also avoid other sensitive areas, such as drinking water sources, unless there is no reasonable alternative.294 In determining whether there is a reasonable alternative, the commission can weigh the management of adverse impacts, orderly siting, system reliability and integrity, and efficient use of resources.295

South Dakota’s and North Dakota’s reviews of the Dakota Access Pipeline demonstrate the limitations of the facility approach to risk. South Dakota’s siting law requires the operator to demonstrate that the pipeline will not “pose a threat of serious injury to the environment” or “substantially impair the health, safety or welfare of the inhabitants.”296 But the commission concluded that the operator met these requirements by complying with PHMSA’s safety standards.297 The commission also concluded that it could not “compel the [a]pplicant to select an alternative route” or base its decision “on whether . . . the selected route is the route [the state] might itself select.”298

North Dakota’s siting law protects vulnerable areas, yet the state commission’s application of the law is much less demanding than the text would seem to require. The law prohibits a pipeline from being routed through an “exclusion” area, which includes critical habitat for endangered and threatened

292 Minn. Stat. Ann. § 216G.02, subdiv. 3(b)(4) (West 2010). Recognizing the issue of preemption, Minnesota prohibits the commission from “set[ting] safety standards for the construction of pipelines.” Id. § 216G.02, subdiv. 3(a).
295 N.D. Admin. Code 69-06-08-02(2).
298 Id. at 24–25.
species. The commission, however, interpreted this as a prohibition on surface impacts and allowed the Dakota Access Pipeline to be bored underneath critical habitat in the Missouri River. When Energy Transfer Partners proposed a tank terminal on the other side of a city’s groundwater supply protection area, the commission concluded that there was no reasonable alternative to routing the pipeline through the area even though the operator had created the dilemma.

The figure below shows which states have adopted transmission pipeline siting laws and the approaches they use to assess risk. Most of the states that regulate pipeline siting are in the North and along the West and East Coasts. The author’s future research will explore relationships between the states’ policy choices and pipeline risk characteristics, but it is interesting to note that states with extensive transmission pipeline systems, such as Texas and Louisiana, have not adopted siting laws. There is also no obvious correlation between the type of pipeline and the type of policy approach. The two states that adopted pipeline siting laws in response to the domestic energy revolution—Georgia and Nebraska—chose different policy approaches to address the same type of pipeline. Instead, it appears that certain geographic regions prefer one approach, perhaps because the states in that region look to each other for policy models. The Northeast and the northern states in the Mountain West use the facility approach, while states in the Midwest and the Plains are more likely to use the public utility approach.

---

301 Id. at 8.
IV. EVALUATING THE RISK FRAMEWORKS

These regulatory frameworks raise descriptive and normative questions of risk governance—how society makes collective decisions about the acceptability of risk, and how it should make such decisions. A common governance approach to environmental, health, and safety risks is to separate the decision on siting from the decision on the ongoing risk of the industrial activity. This approach presumes that the decision about the location of a source of risk can be divorced from a decision about the acceptable level of risk—or that the distinction is not important because the risk can be managed regardless of its geographic location. Both presumptions can be questioned. But even if the distinction is valid, one would expect risk governance to adopt a more cautious approach when the potential for catastrophic harm increases and the ability to predict the consequences of the next accident decreases.

A. The Risk Policy Approaches

In risk governance, there are three primary policy approaches: a preventative approach, a management approach, and a remedial approach.

302 See ORTWIN RENN, RISK GOVERNANCE: COPIING WITH UNCERTAINTY IN A COMPLEX WORLD 8–9 (2008) (describing risk governance as “structures and processes for collective decision-making involving governmental and non-governmental actors” as applied to the “context of risk and risk-related decision-making,” which includes the “complex web of actors, rules, conventions, processes and mechanisms concerned with how relevant risk information is collected, analysed and communicated, and how management decisions are taken”); Timothy F. Malloy, Disrupting Conventional Policy: The Three Faces of Nanotechnology, 28 UCLA J. ENVTL. L. & POL‘Y 1, 3 (2010) (defining risk governance as “the social, legal and institutional decision-making processes used in identifying and responding to risks facing society”).

303 Major environmental laws such as the Clean Water Act and the Clean Air Act do not directly address siting by specifying the best location for a facility. See Rodger C. Field, Siting, Justice, and the Environmental Laws, 16 N. ILL. U. L. REV. 639, 645–46 (1996). Instead, the laws accept the proposed site as a given and set pollution standards based on the environmental characteristics of the area. See id. at 645.


Each policy approach chooses to intercede at a different point in the risk timeline, from before the risk-creating activity begins to after harm occurs. Underlying this choice are distinct risk philosophies—from policies that focus primarily on protecting the public and the environment to those that consciously allow some harm. Each approach is thus built on a particular characterization of risk, a judgment about the acceptability of risk, and an assumption about the efficacy of policy measures to respond to risk. While the approaches are conceptually independent, they are not mutually exclusive. A policymaker can apply multiple approaches to the same risk activity.

In the preventative approach, policymakers respond to risk through measures that avert harm. At the extreme, a preventative policy embraces the strong version of the precautionary principle and prohibits an activity unless the risk creator can demonstrate that the activity would be safe. More pragmatically, a preventative policy protects communities and the environment from the worst harms. By seeking to prevent harm, the policy approach treats negative consequences as unacceptable—either because from a technocratic perspective the benefits of the activity do not justify the risks, or because the public perceives the risks to be intolerable regardless of the benefits. From a democratic governance perspective, the policy approach is particularly suited to “dread” risks that evoke significant public concern because they cause fatalities, create the potential for catastrophic outcomes, are not controllable, and do not equally distribute the risks and benefits. The policy approach is also suited to

---


308 See Malloy, supra note 305, at 149 (stating that “one can adopt a prevention-based approach without embracing the precautionary principle”); Robert V. Percival, Who’s Afraid of the Precautionary Principle?, 23 PACE ENVTL. L. REV. 21, 36–75 (2005) (describing range of preventative or precautionary actions in U.S. environmental law).

risks that are uncertain, in the sense that the probability cannot be calculated or the consequences are unknown.\textsuperscript{310}

In the management approach, policymakers respond to risk through measures that limit the risk of harm. A management policy permits an activity to occur while imposing controls that reduce—but do not prevent—risk.\textsuperscript{311} While the activity still has the potential to create harm, the risk is acceptable because the benefits of the activity outweigh the risks or the public perceives the risks to be tolerable. In a technocratic paradigm in which risks are just another policy problem to be competently managed, policymakers can adopt this approach to address a range of risks with different characteristics.\textsuperscript{312} But because management depends on the presumption that policy measures can reliably reduce risk to an acceptable level,\textsuperscript{313} the approach is best suited to regulating risk-creating activities with more familiar characteristics. From a democratic governance perspective, management is also appropriate for risks that do not evoke the same level of public concern—that is, they are controllable; voluntary; known to the public; have immediate, observable effects; or are more equitable in the distribution of risks and benefits.\textsuperscript{314}

In the remedial approach, policymakers respond to risk through measures that reduce the harm after an accident. A remedial policy allows an activity but requires the risk creator to plan for emergencies and to take immediate action when harm occurs.\textsuperscript{315} The approach accepts not only the presence of risk, but also the inevitability of some damage—because it presumes that remedial measures can lessen the consequences to an acceptable level.\textsuperscript{316} It is best suited to respond to risks that create large benefits and cause the type of harm that can

\textsuperscript{310}The traditional formulation of the precautionary principle provides that policymakers should take action to prevent threats of serious or irreversible harm even in the face of scientific uncertainty. See Percival, supra note 308, at 22–36; see also Slovic, supra note 309, at 282–83 (describing the public’s concern about “unknown” risks that are new, latent, or not observable).

\textsuperscript{311}See Malloy, supra note 305, at 112 (describing conventional risk management as “setting ‘acceptable’ exposure levels and relying on engineering controls to achieve such levels”).


\textsuperscript{313}Cf. Malloy, supra note 305, at 134–35 (discussing potential ineffectiveness of control measures).

\textsuperscript{314}See Slovic, supra note 309, at 282–83.


be easily controlled. This approach can be combined with a management approach, since it addresses the remaining potential for harm. To the extent the approach is used by itself, it is suited to risks that create less significant consequences, such as activities that do not cause catastrophic harm, many fatalities, or widespread property damage.

Energy pipelines benefit consumers and the economy, but they also pose catastrophic risks to populated and environmentally sensitive areas. Because a pipeline accident is more likely to cause extreme consequences than would be expected, the scale of a future accident is difficult to predict based on past events. As companies build ever more pipelines across the United States, a preventative policy approach would seem best suited to addressing the risks to the areas where the consequences of an accident would be most severe.

But when the safety and siting frameworks are combined, they take a markedly constrained approach to prevention. The safety framework seeks to prevent risk using engineering standards that regulate the design, installation, and construction of pipelines. Once the pipelines are constructed, the safety framework works backward to reduce the risks to a limited number of vulnerable areas through special requirements such as risk management programs. The siting framework largely defers to the engineering and management standards in safety regulation, rather than take its own preventative approach. To address the inevitable accidents that follow management policies, the safety framework takes a remedial approach and requires pipeline operators to plan for emergencies and respond promptly.

The frameworks treat risk as a technical problem in the physical pipeline systems and as a management problem in the organizations that operate those systems, not as a problem of incompatible land uses. This vision of risk reifies pipelines and organizations, extracting the problem from the particulars of the surrounding communities and environment. It frames the risk as controllable rather than catastrophic, and places decision-making in the hands of experts instead of land use planners. Finally, it presumes that the policy measures can effectively reduce risk to an acceptable level, putting faith in the engineering standards and technologies that prevent and reduce risk—as well as in remedial measures to minimize the effects of a spill or release.

---

317 See Marchant & Stevens, supra note 315, at 246 (arguing that ex post governance is appropriate for technologies that have major net safety benefits but also cause accidents).
319 Cf. W. Kip Viscusi & Richard J. Zeckhauser, Deterring and Compensating Oil-Spill Catastrophes: The Need for Strict and Two-Tier Liability, 64 Vand. L. Rev. 1717, 1734 (2011) (“There are two implications of catastrophes being characterized by fat-tailed distributions. First, where disasters are concerned, the past may not be prologue. A future disaster could easily be many times worse. Second, a single extreme outcome may readily account for most of the losses from a particular type of catastrophe.”).
321 Id. §§ 192.901–951, 1001–1015, 195.452.
322 Id. §§ 192.605(a), (e), .615, 195.402(a), (e).
In viewing the risk problem as a pipeline problem or an operator management problem, the frameworks follow an approach that emerged from the pipeline industry. The division between “safety” and “siting” began in the industry’s safety codes. Originally, the pipeline industry regulated safety itself, using standards developed by private standard-setting organizations. In the 1950s, the states began adopting the codes as safety regulations for public utilities. Congress accepted this narrow vision of safety and gave the new Department of Transportation authority to set standards for natural gas pipelines, then adopted the same approach for liquid pipelines.

Siting laws, where they exist, primarily treat pipelines as inert construction projects. They seek to prevent and mitigate short-term “effects” more than long-term risk. Indeed, states that regulate the siting of both gas and liquid transmission pipelines use the same criteria, indicating that the focus of the laws is on the pipeline more than on the product it transports. Only one state—North Dakota—has adopted a completely preventative approach to ensure that transmission pipelines will not harm protected areas. But even this siting law takes a relatively narrow approach to prevention that appears to be focused on construction effects. The excluded areas are lands that have already been protected for their recreational or historical value by the federal government or state: national and state parks, historic sites, and wilderness areas. The only private lands that are excluded are those that serve as habitat for threatened, endangered, unique, or rare species.

In theory, laws that require agencies to analyze alternatives to the route of a pipeline should lead to pipelines that are sited away from vulnerable areas. Under NEPA and similar “stop and think” state laws, agencies must compare the environmental effects of the proposal against the alternatives. Several of the pipeline siting laws also require the decision-maker to consider alternate routes. Even if the laws do not mandate approval of the route with the least

324 See id. at 5 (describing the industry code for natural gas pipelines as a “consensus of informed engineering judgment as to minimum construction requirements for safety”).
325 Id. at 126–27 tbls.3 & 4.
328 N.D. ADMIN. CODE 69-06-08-02(1) (2019).
329 Id.
331 See, e.g., CONN. GEN. STAT. § 16-50(a)(1)(D) (Supp. 2019); FLA. STAT. ANN. § 403.9412 (West 2015); IOWA CODE ANN. § 479.6(8) (West 2009); N.H. REV. STAT. ANN. § 162-H:7(V)(b) (2014).
impacts, the analysis should create external and internal pressure on the agency to make the best decision. In practice, however, the evaluation of alternatives is limited by the same divide between siting and safety. In FERC’s environmental analysis of the Rover pipeline, for example, the commission compared the surface environmental impacts of the alternative routes but not the consequences of an accident. When FERC evaluated the safety of the pipeline, it took the location of the pipeline as a given and relied on PHMSA’s more stringent design and management standards to control the risk of accidents in more densely populated areas. FERC ultimately concluded “that [the pipeline] would represent a slight increase in risk to the nearby public.” This conclusion points to a deeper concern about using laws such as NEPA to compare risks. It is particularly difficult for agencies to predict low-probability, high-consequence events in a changing risk landscape because the frequency and consequences of accidents in the past do not necessarily predict the future.

B. The Results

The failure of risk governance to take a robust preventative policy approach to the catastrophic risks posed by energy pipelines to vulnerable areas creates four negative effects: (1) it leads to more pipeline accidents that cause the worst harms; (2) it places a significant burden on the public and local governments to manage risk; (3) it relies disproportionately on emergency response measures to mitigate harm; and (4) it encourages pipeline operators to build more pipelines than is efficient.

332 See Robertson v. Methow Valley Citizens Council, 490 U.S. 332, 349 (1989) (explaining that environmental analysis forces action by “ensur[ing] that the agency...will...carefully consider...detailed information concerning significant environmental impacts...[and] guarantee[ing] that the relevant information will be made available to the larger audience that may also play a role in both the decisionmaking process and the implementation of that decision”).

333 Fed. Energy Regulatory Comm’n, FERC/FEIS-0267F, Rover Pipeline, Panhandle Backhaul, and Trunkline Backhaul Projects: Final Environmental Impact Statement 3-10 to -35 (July 2016), https://www.ferc.gov/industries/gas/enviro/eis/2016/07-29-16-rover-pipeline/impact-statement.pdf [https://perma.cc/YV7V-87SV] [hereinafter FERC, Rover Pipeline Final EIS]. The criteria included residences within 50 feet, but this criterion must have measured nuisance impacts; the potential impact radius of an explosion is 628 to 1100 feet. Id. at 4-270.

334 Id. at 4-258 to -267.

335 Id. at 4-264 to -266, -271. The Commission also presented a risk ladder that showed that the annual number of deaths from tractor turnovers was higher than the deaths from natural gas pipelines, a statistic unlikely to sway anyone concerned about the involuntary risks of pipelines. Id. at 4-271.

336 See Jamison E. Colburn, Necessarily Unpredictable? Oil Spill Risks Beyond the Horizon, 30 Miss. C. L. Rev. 307, 328 (2011) (explaining the challenges of assessing low-probability, high-consequence risks under NEPA and contending that frequentist data from past accidents fails to predict unprecedented events like Deepwater Horizon).
First, the lack of a preventative approach leads to more pipeline accidents and to more catastrophic harm to the environment and the public from pipeline accidents. Because the regulatory frameworks do not require the risk of an accident to be taken into account when a pipeline is sited, it is not surprising that operators continue to build energy pipelines in or near highly populated and environmentally sensitive areas when it makes economic sense to do so.\textsuperscript{337} Approximately 7\% of natural gas transmission pipelines and 42\% of hazardous liquid transmission pipelines are located in areas where PHMSA has determined that an accident could have significant consequences.\textsuperscript{338} And during the energy revolution, companies have continued to build transmission pipelines—particularly the interstate pipelines that are more likely to pose significant risks—in high-consequence areas at the same pace as pipelines in other areas.\textsuperscript{339} Of course, pipelines must sometimes be built in a specific location.\textsuperscript{340} But operators appear to have significant latitude to choose the path from one point to another, particularly when they construct new “greenfield” projects.\textsuperscript{341} For large projects such as the Keystone XL pipeline, operators have proposed...
substantial changes to the pipeline route when required to submit route alternatives to a federal or state agency.\footnote{342 See, e.g., Application, Transcanada Keystone Pipeline, L.P., Neb. Pub. Serv. Comm’n, No. OP-0003, at 2–8 (2017) (describing alternative routes).}

According to PHMSA’s accident data, pipelines located in high-consequence areas are more likely to have spills or releases than pipelines located in other, “low-consequence” areas.\footnote{343 See HL IM Performance Measures, supra note 221 (follow “HL IM reporting data” hyperlink); National Pipeline Performance Measures, supra note 177 (follow “Onshore Significant Incident HCA” hyperlink).} This is in part because the external threats to pipelines are greater in such areas.\footnote{344 See National Pipeline Performance Measures, supra note 177 (follow “Onshore Significant Incident HCA” hyperlink) (noting that “[s]ince [high-consequence areas] are typically developed areas, [natural gas transmission] pipelines have increased integrity risks from excavation and outside force damage”).} But even if the likelihood of an accident were the same in both types of areas, the siting of a pipeline in a high-consequence area instead of a low-consequence area would result in greater harm over the long term. A spill or release from a pipeline in a high-consequence area is, by definition, more likely to cause severe effects than the same accident in a low-consequence area.\footnote{345 See Pipeline Safety: Pipeline Integrity Management in High Consequence Areas, 65 Fed. Reg. 21,695, 21,699 (proposed Apr. 24, 2000) (to be codified at 49 C.F.R. pt. 195) (describing “high-consequence areas” as areas where “a pipeline failure could pose the greatest threat to public safety, the environment, and water commerce”).}

PHMSA has tried to reduce the risk of accidents in high-consequence areas by requiring operators to develop integrity management programs to identify, assess, and manage threats to the integrity of their pipelines.\footnote{346 49 C.F.R. §§ 192.901–951 (natural gas transmission pipelines), .1001–1015 (natural gas distribution pipelines), 195.450, .452 (hazardous liquid pipelines) (2019).} Even putting aside the limited number of communities and natural resources protected by these programs, the effectiveness of the approach to risk is debatable. If the programs were designed to deter operators from siting new pipelines in high-consequence areas, the additional cost of the requirements does not appear to have altered the operators’ decisions. Nor have these risk management programs been successful in lowering the number of accidents or the total amount of damage; the rate of transmission pipeline accidents in high-consequence areas has stayed the same or increased since the programs went into effect.\footnote{347 RICK KOWALEWSKI, PIPELINE INTEGRITY MANAGEMENT: A REPORT TO THE SECRETARY OF TRANSPORTATION 24–25 (Oct. 2013), http://pstrust.org/wp-content/uploads/2015/10/Kowalewski-IM-PE_Report.pdf [https://perma.cc/5W8J-3SDW]; NAT’L TRANSP. SAFETY BD., NTSB/SS-15/01, SAFETY STUDY: INTEGRITY MANAGEMENT OF GAS TRANSMISSION PIPELINES IN HIGH CONSEQUENCE AREAS 65 (Jan. 2015), https://www.ntsb.gov/safety/safety-studies/Documents/SS1501.pdf [https://perma.cc/7GH9-5LYE].} Indeed, the incidents that have caused the greatest harm have occurred recently and in high-consequence areas.\footnote{348 KOWALEWSKI, supra note 347, at 25–27.}

\footnote{343 See HL IM Performance Measures, supra note 221 (follow “HL IM reporting data” hyperlink); National Pipeline Performance Measures, supra note 177 (follow “Onshore Significant Incident HCA” hyperlink).}
\footnote{344 See National Pipeline Performance Measures, supra note 177 (follow “Onshore Significant Incident HCA” hyperlink) (noting that “[s]ince [high-consequence areas] are typically developed areas, [natural gas transmission] pipelines have increased integrity risks from excavation and outside force damage”).}
\footnote{345 See Pipeline Safety: Pipeline Integrity Management in High Consequence Areas, 65 Fed. Reg. 21,695, 21,699 (proposed Apr. 24, 2000) (to be codified at 49 C.F.R. pt. 195) (describing “high-consequence areas” as areas where “a pipeline failure could pose the greatest threat to public safety, the environment, and water commerce”).}
\footnote{348 KOWALEWSKI, supra note 347, at 25–27.}
Second, the lack of a preventative approach to siting shifts the responsibility for preventing pipeline accidents from pipeline operators to landowners, nearby residents, and local governments. Rather than requiring operators to site pipelines away from people, the regulatory system relies heavily on the people exposed to risk to protect pipelines. This includes landowners who involuntarily encounter the risk because their property has been taken through eminent domain. PHMSA supports this shift by requiring pipeline operators to develop “public awareness” programs. The purpose of these programs is to educate residents about the hazards associated with pipelines and to enlist them as partners in safety so they will refrain from activities that damage the pipeline and monitor the right-of-way for any threats. The effectiveness of the operators’ programs is questionable. Local governments are also expected to manage the risks of existing pipelines through zoning and planning requirements. PHMSA recommends that local governments require developers to consult with pipeline operators and restrict land uses that could interfere with pipelines. The burden of reducing risk is thus on the community surrounding the pipeline. In practice, localities often do not curb incompatible land uses, because officials are not aware of the risk or because such restrictions are politically unpopular.

Third, the lack of a preventative approach puts extraordinary pressure on mitigation strategies to address the effects of accidents. PHMSA requires

---

349 As other scholars have explored, pipeline companies generally have the authority to exercise eminent domain to site pipelines under either federal or state law. See Klass & Meinhardt, supra note 36, at 951.


352 See NAT’L TRANSP. SAFETY BD., ENBRIDGE REPORT, supra note 184, at 43 tbl.1 (noting the results of a survey measuring the effectiveness of the public awareness program, in which only 23% of the affected public, 39% of public officials, and 47% of emergency officials said that they were “very well informed” about pipelines in their community); NAT’L TRANSP. SAFETY BD., SAN BRUNO REPORT, supra note 182, at 59 (noting the results of a survey measuring the effectiveness of the public awareness program, in which “the affected public was [the operator’s] least informed audience,” with 89% reporting that “they did not recall receiving information” from the operator and 34% reporting that they considered themselves “somewhat or very well informed”).


pipeline operators to develop emergency plans with written procedures that detail how the operators will respond to an emergency.\[355\] The operator must train its employees to follow the procedures.\[356\] It must also establish a liaison with local emergency responders so that they can coordinate efforts when there is an incident.\[357\] In addition, operators of oil pipelines must prepare a plan that identifies the resources that are available to clean up a “worst case discharge.”\[358\]

In practice, operators have found it difficult to put the procedures into action and respond quickly to incidents.\[359\] Overburdened fire departments and other emergency officials often do not understand the hazards of pipelines and how to respond to emergencies, either because the operator has failed to adequately train them or because the officials did not take the time to attend trainings.\[360\] And even if the system worked perfectly, there is a limit to mitigation as policy approach. Natural gas explosions, for example, will immediately destroy the buildings within the zone of impact.\[361\]

Fourth, the lack of a preventative approach results in a greater number of pipelines than is economically efficient. When the operator is deemed a common carrier and there is no review of the need for the pipeline or alternative routes, the absence of a governance mechanism to consider long-term risks will lead to overbuilding. An efficient system of risk governance would require a pipeline operator to fully weigh the cost of the risk of a spill or release against the benefits in deciding whether to build a pipeline in a particular location.\[362\]

Pipeline operators may have at least some incentive to site their pipelines in a way that minimizes their liability for future damages.\[363\] But given the difficulty

\[356\] Id. § 192.615(b)(2).
\[357\] Id. §§ 192.615(c), 195.402(c)(12).
\[358\] Id. §§ 194.101, .105.
\[359\] See NAT’L TRANSPI. SAFETY BD., ENBRIDGE REPORT, supra note 184, at 8–18; NAT’L TRANSPI. SAFETY BD., SAN BRUNO REPORT, supra note 182, at 98–99.
\[360\] See NAT’L TRANSPI. SAFETY BD., ENBRIDGE REPORT, supra note 184, at 105; NAT’L TRANSPI. SAFETY BD., SAN BRUNO REPORT, supra note 182, at 77 (noting that the fire chief of the city did not know there was a pipeline in that location).
\[362\] Cf. Russell S. Jutlah, Economic Theory and the Environment, 12 VILL. ENVTL. L.J. 1, 14 (2001) (stating that “the objective [of environmental economics] is to achieve a socially optimal allocation of resources by ensuring that polluters, and others whose activities may adversely affect environmental quality, bear the full costs that their activities may impose”).
\[363\] Cf. Viscusi & Zeckhauser, supra note 319, at 1737–38 (describing the “conventional retrospective liability approach,” in which “payment for all damages leads the injurer to internalize costs and to take appropriate levels of care”).
in predicting the consequences of an accident, it seems unlikely that liability results in full internalization of the cost of risk.\footnote{Cf. id. at 1738 (arguing that it is difficult to determine the magnitude of harm when there is a catastrophic accident).}

Even when the operator is required to obtain approval for a route from FERC or state governments, there will still be too many pipelines. This is because decision-makers do not consider the risks of an accident to be a part of economic regulation. For example, FERC frames the economic inquiry as whether existing customers will subsidize the pipeline project and the extent to which customers and property owners will bear additional financial costs.\footnote{Certification of New Interstate Natural Gas Pipeline Facilities, 88 FERC § 61,227, at 19 (1999) (describing the initial assessment of adverse effects as “essentially an economic test” and explaining that “[o]nly when the benefits outweigh the adverse effects on economic interests will the Commission then proceed to complete the environmental analysis where other interests are considered”).}

The risk of long-term spills and releases should affect the economic calculus and thus the public need for the pipeline. FERC is already criticized for approving too many projects,\footnote{Since 1999, FERC has denied only two applications under Section 7 of the Natural Gas Act. SUSAN TIERNEY, ANALYSIS GRP., NATURAL GAS PIPELINE CERTIFICATION: POLICY CONSIDERATIONS FOR A CHANGING INDUSTRY 12–13 (Nov. 2017), https://www.analysisgroup.com/uploadedfiles/content/insights/publishing/ag_ferc_natural_gas_pipeline_certification.pdf [https://perma.cc/H6T5-PCMY].
} but critics have not focused on the treatment of risk as a cause.\footnote{See, e.g., CATHY KUNKEL & TOM SANZILLO, INST. FOR ENERGY ECONS. & FIN. ANALYSIS, RISKS ASSOCIATED WITH NATURAL GAS PIPELINE EXPANSION IN APPALACHIA 1 (Apr. 2016), http://ieefa.org/wp-content/uploads/2016/04/Risks-Associated-With-Natural-Gas-Pipeline-Expansion-in-Appalachia_-April-2016.pdf [https://perma.cc/H2K6-93LM] (arguing that FERC’s allowed rate of return on equity and a lack of comprehensive planning encourage overbuilding, and concluding that so many new pipelines create a significant safety issue for landowners).}

C. Three Potential Solutions

To avoid these results, the siting and safety frameworks should adopt a more preventative approach to the risks of energy pipelines. This Article sketches three potential policy solutions: a policy that sets aside protected areas, a policy that returns decision-making to landowners and municipal governments, and a policy that combines safety and siting decisions. It concludes that the third solution provides the greatest benefits to communities and the environment. At the very least, this policy should apply to gas and hazardous liquid transmission pipelines as well as larger gas and oil gathering pipelines that effectively create the same risks as transmission lines. This would capture most of the projects that will be built in the next few decades, as well as the types of pipelines that are most likely to have a flexible route.
One policy solution is to follow the North Dakota model and create exclusion and avoidance areas. Under this policy, FERC and the states that choose to regulate siting would prohibit companies from siting pipelines in certain exclusion areas that are vulnerable to significant harm in an accident. Companies would also be required to avoid areas that were vulnerable to moderate harm unless there were no reasonable alternative. Because the policy would specifically focus on the risk of spills and releases, it would encompass pipelines sited outside of vulnerable areas if a worst-case release could impact the areas. This expansion would prevent companies from manipulating the system by building pipelines that skirt sensitive features yet still pose risks.

Each jurisdiction could choose the areas it would protect. Alternatively, federal and state environmental, natural resource, and public health agencies could work together to establish risk-based criteria for designating vulnerable areas, taking into account the type of pipeline and the risks posed by different products. Each decision-maker would then apply the criteria to the areas within its jurisdiction. The policy could also draw on local knowledge by allowing local governments to nominate areas within their jurisdictions for consideration. If the governance process limits exclusions to locations where an accident would cause the most significant consequences, any exception to the prohibition should be narrow. For example, it could allow only those projects that have a demonstrated need to be in the area or that would pose greater risks to communities and the environment if they were required to use alternative routes.

The benefits of this policy solution are that it would fit easily into the current regulatory frameworks, and it would take a completely preventative approach to risk by protecting sensitive areas from potential harm. This would reduce the number of accidents with significant consequences and free landowners and communities from the burdens of managing the risk and responding to accidents. But the policy would treat the safety of pipelines as a dichotomous choice between no harm and the risk of an accident. And the larger the protected area, the greater the pressure would be to make exceptions to the policy. Based on the North Dakota Public Service Commission’s implementation of the law, the decision-maker may find it challenging to protect areas when companies contend that it is necessary to site the pipeline in that location. This is particularly true because the decision-maker is likely to be hampered by information asymmetry; the pipeline company would have much more information about the need for the route and the characteristics of alternative routes than would the agency.

Another solution would be to transfer the governance of siting to landowners and local governments. This policy would remove the power of condemnation from pipeline companies and give local governments the authority to regulate pipelines like any other land use under their zoning and planning laws. Federal and state agencies would retain jurisdiction over pipeline safety and public lands, but landowners and local governments would exert primary decision-making power over the pipeline route. The result would be a siting framework in which private landowners would decide the level of
acceptable risk, subject to local land use requirements. A version of this policy solution is already being implemented by Indian tribes, who have sovereign authority to decide whether to grant rights-of-way to pipeline companies and who may set conditions or restrictions on the approval.\footnote{See 25 C.F.R. § 169.107(a) (2019) (acknowledging that applicants for rights-of-way must obtain consent from tribes if the project is on tribal land).}

The benefits of this policy solution are that it would internalize the social costs of a pipeline project by requiring companies to pay for the use of property and for the imposition of risk. Agreements could be structured to include yearly risk payments, which would more accurately reflect the cost of the pipeline operation.\footnote{In Wyoming, courts may award yearly payments to landowners for the taking of an easement under the state’s eminent domain statute. See Barlow Ranch, Ltd. P’ship v. Greencore Pipeline Co., 301 P.3d 75, 103 (Wyo. 2013); see also Kelianne Chamberlain, Unjust Compensation: Allowing a Revenue-Based Approach to Pipeline Takings, 14 Wyo. L. Rev. 77, 95–98 (2014) (arguing that landowners should receive annual payments tied to revenue as compensation).}

This would result in an efficient number of pipelines. But the effectiveness of the policy solution depends on whether individuals and local governments have enough information to make a decision on the acceptability of the risk. It is also possible that important pipeline projects that serve a public need would not be built because companies would be unable to obtain easements from holdouts.\footnote{See Coleman & Klass, supra note 38, at 717.} This is a potential problem, but it is not clear that it would be an insuperable hurdle for projects. Pipeline companies are currently siting projects in a few states without the power of eminent domain, presumably by paying more for the privilege.\footnote{For example, in 2012, the Colorado Supreme Court held that companies may not exercise eminent domain authority to site oil pipelines in the state. Larson v. Sinclair Transp. Co., 284 P.3d 42, 46 (Colo. 2012) (en banc); see also Klass & Meinhardt, supra note 36, at 987. Yet operators have built 1,157 miles of oil pipeline in Colorado since 2012. Pipeline Mileage and Facilities, supra note 67. In Pennsylvania, companies have sited thousands of miles of gathering pipelines without eminent domain. See 66 Pa. CONS. STAT. § 1104 (2000) (requiring a public utility to obtain a certificate of public convenience before exercising the power of eminent domain). But see Press Release, Penn. Pub. Util. Comm’n, PUC Continues Consideration of Laser Northeast Gathering Co. Application (May 19, 2011), http://www.puc.state.pa.us/about_puc/press_releases.aspx?ShowPR=2759 [https://perma.cc/MGD7-6459] (describing 3-2 decision to grant a certificate to a gathering pipeline company).}

This policy solution would not necessarily reduce the number of accidents or the burden of accidents on communities. Some individual property owners may choose not to allow pipelines in vulnerable areas or may be more likely to detect a problem because they appreciated the risk. In the absence of these private actions, it would be up to local governments to protect the community by adopting planning and zoning requirements. For example, local governments could adopt a minimum setback requirement for new pipelines and require that any developments in the area around the pipeline obtain prior approval from the zoning authority. But these actions would be dependent on the capacity of local...
governments to understand the risk and to make informed decisions that protect the public and environment.

A final policy solution would be to combine the safety and siting frameworks into one risk governance system. Unlike the other solutions, this policy would grant a single agency the authority to review all of the risks of a pipeline project—from the risks of harm to the environment caused by construction, to the risks of operation, to the risks of an accident to individuals and property. The agency would then approve the pipeline if the location and the safety measures together made the risk acceptable. PHMSA would be the best agency to regulate the route and safety of gas and liquid interstate transmission pipelines because of its expertise on pipeline risk, while the states could regulate intrastate transmission lines.

The primary benefit of this policy is that the agency would make more knowledgeable and comprehensive decisions about pipeline risk. In contrast to the current system and the other policy solutions, the decision-maker would be able to consider the entire life cycle of risk at the beginning of the project, from the location to the operation to the consequences of accidents. This comprehensive review would result in pipelines that have fewer accidents with less significant consequences, create less of a burden on communities and landowners to manage risk, and required fewer emergency response measures. The policy would also improve economic efficiency by ensuring that only those pipelines that create an acceptable risk would be built.

One potential criticism is that the comprehensiveness of the review may delay the approval of worthy projects. FERC addresses this issue by offering the option of pre-filing procedures. PHMSA and the states could also look to the Nuclear Regulatory Commission’s (NRC’s) regulation of nuclear power plants as a model. The NRC regulates all risks associated with the plants, from siting, to design, to operation, to the final stage of decommissioning. To streamline the process, however, the NRC allows operators to apply for an early site permit. The commission weighs safety issues related to the location, effects on the environment, and emergency planning in the permitting decision. The risks posed solely by the design and operation of the plant are then evaluated in separate proceedings.

This solution would no doubt face political hurdles. The states would likely oppose any attempt to give jurisdiction over the siting of interstate oil and hydrocarbon liquid pipelines to the federal government, even if the states were given more authority over the risks of siting intrastate pipelines in return. FERC may not be any more amenable to relinquishing its authority over the siting of natural gas pipelines, though it could retain economic regulation of the lines. But the second option would transfer even more authority from state and federal

374 Id. §§ 2133, 2137.
376 Id. §§ 52.41–.63 (design certification), .71–.110 (combined licenses).
agencies to local governments and individuals. And setting aside their parochial interests, all of the government agencies must recognize that the current system is not successful.

V. CONCLUSION

The United States is in the midst of a radical reshaping of energy pipeline networks that will create a new landscape of risk for decades to come. The recent controversies surrounding pipelines demonstrate that the public is deeply skeptical of risk decisions made by government agencies and pipeline companies. Yet the legal frameworks governing the siting and safety of these pipelines fail to use the one tool that would prevent the worst harms to communities and the environment—preventative land-use planning. This failure is based on a shuttered vision of pipelines as self-contained systems that can be placed in almost any location with the right safety devices. It is time for risk governance to expand its remit, thus matching the geographic scope and optimism of the domestic energy revolution.
First Amendment (Un)Exceptionalism: A Comparative Taxonomy of Campaign Finance Reform Proposals in the United States and United Kingdom

LORI A. RINGHAND

ABSTRACT

There is an urgent conversation happening among the world’s democracies about how to respond to the combined threat of online electioneering and foreign interference in domestic elections. Despite the shadow such activities cast over the 2016 presidential election in the United States, the United States has been largely absent from comparative discussions about how to tackle the problem. This is not just because of a recalcitrant president. The assumption that America’s “First Amendment Exceptionalism”—the idea that American freedom of expression law is simply too much of an outlier to warrant useful comparative consideration—is strong on both sides of the Atlantic. This is especially true in regard to the regulation of political campaigns.

This Article challenges that assumption, and argues that America’s more libertarian approach to the legal regulation of political speech does not pose a barrier to fruitful comparative work in this area. It does so by comparing the law of the United States to that of the United Kingdom. Specifically, it organizes reform proposals being considered in the United States and United Kingdom into a common taxonomy, and sets out the legal standard governing each type of proposal in each country. Considering each country’s law through this organizational structure allows us to see that the legal differences between the United States and United Kingdom, while significant, rarely bar the types of changes being considered in either nation. Indeed, the two countries have much to learn from each other’s efforts in this area, and

* J. Alton Hosch Professor of Law, University of Georgia College of Law; and Spring 2019 Fulbright Scotland Visiting Professor of Law, University of Aberdeen School of Law. I am indebted to the U.S.-U.K. Fulbright Commission, the University of Aberdeen, and the University of Georgia for their generous support of this project. I also have benefited from thoughtful feedback provided by colloquium and conference participants at the Gresham College Fulbright Lecture, the Centre for Law, Policy and Society at the University of Southampton, the Centre for American Legal Studies at Birmingham City University School of Law, and the Southeastern Association of Law Schools 2019 Annual Conference. Special thanks are owed to Jacob Eisler, Heather Green, Richard L. Hasen, Daniel R. Lorentz, Joseph S. Miller, Morgan Pollard, Robert Taylor, Daniel P. Tokaji, Sonja West, Paul Wragg, and Stephen Wolfson. Errors are my own.
lawmakers, regulators, and scholars should not hesitate to engage with the experiences of their transatlantic peers.

In reaching this conclusion, the Article makes three distinct contributions. First, by clustering reform proposals into a taxonomy, it provides a structure for comparative work that will be useful not just in the United States and United Kingdom, but in all countries working to bring their election laws fully into the internet era. Second, by providing an in-depth yet accessible guide to the legal structures undergirding election law in the United States and United Kingdom, it provides a useful tool for scholars attempting to understand these systems. The U.S. system in particular is often quickly dismissed by other nations, but without a deeper understanding of how and why U.S. law has ended up as it has, those nations risk inadvertently following in its footsteps. Finally, it identifies several concrete areas where the United States and United Kingdom can benefit from each other’s expertise, thereby providing a roadmap for regulators, lawmakers, and reform advocates in both countries.

TABLE OF CONTENTS

I. INTRODUCTION ..........................................................407
II. THE 2016 ELECTIONS .....................................................408
III. FOUNDATIONS ......................................................................................................410
    A. Campaign Finance Law in the United States ..................410
    B. Political Party Funding and Campaign Finance Law in the United Kingdom ...........................................415
IV. TAXONOMY OF REFORM PROPOSALS ..........................................418
    A. Public Education ......................................................................................................421
    B. Transparency ..............................................................422
        1. Transparency in the United States ........................................423
        2. Transparency in the United Kingdom ..........................431
        3. Transparency Recap ..................................................436
    C. Source Exclusions ..............................................................437
        1. Source Exclusions in the United States .................438
        2. Source Exclusions in the United Kingdom .............442
        3. Foreign Source Exclusion Recap .........................451
    D. Content Exclusions ..............................................................452
        1. Content Exclusions in the United States ..................452
        2. Content Exclusions in the United Kingdom .............457
        3. Content Exclusions Recap ........................................462
V. CONCLUSION ........................................................................................................462
I. INTRODUCTION

The 2016 presidential election in the United States and referendum on European Union membership (Brexit) that same year in the United Kingdom were a wake-up call to those nations about the extent to which disinformation, propaganda, and “fake news” spread online can amplify extremism and undermine democratic elections. Regulators, lawmakers, and reform advocates have responded by recommending a wide array of legal reforms, including changes to the rules governing political campaigns. In both the United States and the United Kingdom, this process has generated dozens of proposals to more effectively counter online and foreign efforts to influence voters and destabilize democratic institutions.

Despite this, there has been little cross-country analysis comparing those proposals. This is not surprising. The assumption that America’s “First Amendment Exceptionalism”—the idea that American freedom of expression law is simply too much of an outlier to warrant comparative consideration—is widely held on both sides of the Atlantic.1 This is especially true in regard to the regulation of political campaigns, where the United States is most commonly held up in the United Kingdom (and elsewhere) as a negative example to be avoided at all costs.2 In this Article, I challenge that assumption by demonstrating that the legal differences between the United States and the United Kingdom rarely bar the types of reform proposals being considered in either country, and that the two countries can in fact gain considerable insight from each other’s efforts to bring their election laws fully into the internet era.

The Article has four parts. Part II illustrates the challenges faced in both nations by briefly recapping what we know now about online and foreign interference in the 2016 elections.3 Part III contextualizes those challenges by

---

1 See Ronald J. Krotoszynski, Jr., Free Speech Paternalism and Free Speech Exceptionalism: Pervasive Distrust of Government and the Contemporary First Amendment, 76 OHIO ST. L.J. 659, 659 (2015) (“[W]hen viewed from a global perspective, the American position of affording near-absolute protection to speech is strongly exceptionalist.”).


3 Referendum campaigns like the Brexit vote in the United Kingdom are not “elections” (since no candidate is elected). Unless otherwise specified, this Article nonetheless uses the terms “elections” and “election laws” to refer to the full set of primary and secondary legislation, court decisions, regulations, and rules governing both candidate and referendum campaigns.
introducing readers to the foundational rules governing campaign finance and political party funding in the United States and United Kingdom. Part IV, which is the heart of the Article, considers the most significant reform proposals being discussed in each country and organizes them into a taxonomy. As Part IV demonstrates, once the reform proposals are analyzed through this structure, it becomes clear that the two countries have much to learn from each other, despite their different legal rules. The Article concludes by highlighting several areas where further comparative consideration would be the most valuable, and encouraging regulators, lawmakers, and reform advocates in both countries to more fully engage with each other’s efforts in those areas.

II. THE 2016 ELECTIONS

The 2016 presidential election in the United States and the Brexit referendum in the United Kingdom have been extensively studied. In the United States, the Special Counsel to the U.S. Department of Justice, the intelligence community, and select committees of the Senate and the House of Representatives have all investigated the presidential election. In the United Kingdom, the Information Commissioner’s Office, the Election Commission, the Digital, Sports, Media and Culture Committee, and the Committee on Standards in Public Life have examined the Brexit referendum, which also was the subject of investigations by the Cabinet Office and the National Crime Agency.

---


The two countries have focused their efforts somewhat differently. The United States has focused mainly on foreign interference, while the United Kingdom has looked more at data breaches, micro-targeted advertising, and campaign funding improprieties. But their work reveals the same thing: changes in how paid and unpaid communications are purchased, targeted, and shared online have created an unprecedented ability for outside actors to influence domestic politics in ways our election rules did not fully anticipate and have not effectively responded to.

In the United States, the report released in May 2019 by Special Counsel Robert Mueller (the Mueller Report) details how Russian-affiliated actors engaged in an extensive online campaign to influence the 2016 presidential election. In the months before the election, this campaign was supported by a budget of more than $1,250,000 per month and included stealing online identities and information, training foreign actors to create and disseminate inflammatory messages about socially divisive issues, and using interconnected

---


8 According to an indictment filed February 16, 2018 by Special Counsel Robert Mueller, Russian intelligence worked through a Russian corporation—the Internet Research Agency (IRA)—to engage in online “information warfare against the United States.” Indictment ¶ (10)(c), United States v. Internet Research Agency LLC, No. 18-cr-00032-DLF (D.D.C. Feb. 16, 2018). The IRA employed hundreds of people to pose as Americans and comment on social media about U.S. politics. Id. ¶ (10)(a), 32–34. The IRA also engaged in data analytics to target Americans with political messages, and to create and spread “distrust towards the candidates and the political system in general” in the lead-up to the 2016 presidential election. Id. ¶ (10)(e). The foreign origin of these activities was masked, both online and in the underlying financial transactions. Id. ¶ 58.
and often automated networks to spread those messages to targeted audiences across social media platforms. There appears to have been less overt interference in the United Kingdom, but investigations since the Brexit vote have disclosed the prevalence and coordinated distribution of Kremlin-aligned media messaging online, a “Brexit Botnet” active during the referendum campaign, misleading and inflammatory online advertising campaigns targeted to select audiences, and concerns that existing law enabled foreign sources to fund certain online campaign activities.

These events revealed just how ill-equipped existing campaign laws are to deal with this type of activity. Political advertising has been migrating online for decades, but the regulatory systems in both the United States and United Kingdom lag well behind. The problems in each country are similar. Online election communications are subject to few or no transparency requirements, existing reporting rules make it difficult to trace online ads to their underlying funding source, expenditure thresholds triggering regulation do not capture either the low cost of online advertising or the organic way information travels online, microtargeted online advertising diminishes the effectiveness of both regulations and “counter-speech” responses, and rules intended to limit foreign influence in domestic elections are riddled with gaps and unresolved definitional issues.

Understanding how the United States and United Kingdom can learn from each other’s efforts to fix these problems requires first understanding the basic rules currently in place in each country. The following Part provides this, by explaining the foundations of the U.S. and U.K. campaign finance and political party funding laws.

III. FOUNDATIONS

A. Campaign Finance Law in the United States

The backbone of the United States campaign finance system is the Federal Election Campaign Act (FECA), as amended in 1974 in the wake of the Watergate scandal. Congress intended the new law to govern virtually all

---

9 Id. ¶ 11(b), 32–41.
10 DCMS REPORT, supra note 5, ¶ 240–49.
12 DCMS REPORT, supra note 5, ¶ 218–21.
13 Id. ¶ 31–40.
14 See id. ¶ 11–52.
15 Federal Election Campaign Act, 52 U.S.C. §§ 30101–30146 (2012 & Supp. IV 2017); see ROBERT E. MUTCHE, CAMPAIGN FINANCE: WHAT EVERYONE NEEDS TO KNOW 10 (2016). Watergate revealed loopholes in campaign finance rules that had allowed candidate campaign committees to collect large and often undisclosed donations, some of which came from foreign sources and many of which were routed through intermediaries to mask their true origins. See id. at 10–14.
aspects of how campaigns for federal office are regulated.\footnote{In the United States, Congress has full authority to legislate in relation to elections for federal offices and limited authority to regulate in relation to elections for state offices. U.S. CONST. art. I, § 4, cl. 1; see also U.S. GEN. ACCOUNTING OFFICE, GAO-01-470, THE SCOPE OF CONGRESSIONAL AUTHORITY IN ELECTION ADMINISTRATION 3–11 (2001), https://www.gao.gov/assets/240/230112.pdf [https://perma.cc/F7G2-ADJ3]. This Article addresses only federal law, but decisions of the U.S. Supreme Court interpreting the First Amendment limit both federal and state laws and regulations. U.S. CONST. amend. XIV, § 1; see, e.g., Gitlow v. New York, 268 U.S. 652, 666 (1925).} It imposed contribution and expenditure limits on candidates, political parties, and third-party campaigners,\footnote{Federal Election Campaign Act § 30116. The term “third-party campaigners” is used in both the United States and the United Kingdom to refer to groups (such as interest or pressure groups) other than political parties and candidates. See generally Andrew C. Geddis, Confronting the “Problem” of Third Party Expenditures in United Kingdom Election Law, 27 BROOK. J. INT’L L. 103, 107 (2001) (examining how U.K. legislation regulates “third party” expenditures on public messages).} bolstered regulatory reporting requirements, tightened rules against the solicitation and use of foreign funds, and created a federal regulatory body—the Federal Elections Commission (FEC)—to administer the new rules.\footnote{Federal Election Campaign Act §§ 30104, 30106, 30121.} The law was immediately challenged in court.\footnote{Id. at 11. See generally Brief of the Appellants, Buckley v. Valeo, 424 U.S. 1 (1976) (Nos. 75-436, 75-437), 1975 WL 441595, at *29.} The main argument made by the challengers was that virtually all the new rules infringed on political speech protected by the First Amendment.\footnote{Id. at 20–21.}

The resulting Supreme Court decision, \textit{Buckley v. Valeo}, remains the constitutional cornerstone of campaign finance regulation in the United States. To understand \textit{Buckley}, it is useful to keep in mind three things: (1) the difference between contributions (money given to others to spend) and expenditures (money an individual or group spends itself); (2) the difference between candidates and political parties on the one hand, and third-party campaigners (groups other than candidates or political parties) on the other; and (3) the difference between what came to be known as “express advocacy” (which directly calls for the election or defeat of a candidate for federal office) and “issue advocacy” (which does not).\footnote{Federal Election Campaign Act §§ 30101(2), (8)–(9), (16)–(17); see Buckley, 424 U.S. at 42–44; Frank Askin, \textit{Issue Advocacy}, \textit{FIRST AMEND. ENCYCLOPEDIA}, https://www.mtsu.edu/first-amendment/article/996/isue-advocacy [https://perma.cc/KS27-25XF] (defining “issue advocacy”).} The \textit{Buckley} Court treated each of these differences as important, and the constitutionality of a given rule today often hinges on how these categories are defined.

\textit{Buckley} is a complex opinion. The Court first addressed the difference between contributions and expenditures.\footnote{Buckley, 424 U.S. at 19–23.} It viewed contribution limits as only indirectly restricting speech.\footnote{Id. at 20–21.} While such limits restrict the amount of money a
donor can give to political parties, candidates, and third-party groups, they do not directly restrict the donor’s speech itself.\textsuperscript{24} As such, contribution limits need not be subject to the most robust judicial scrutiny and should be upheld as long as they are “closely drawn” to achieve a “sufficiently important interest.”\textsuperscript{25} The Court accepted that preventing corruption or the appearance of corruption was such an interest, and that contribution limits were sufficiently related to that interest.\textsuperscript{26} So the contribution limits imposed by FECA survived this relatively relaxed level of review.\textsuperscript{27}

The Court viewed expenditure limits differently. Unlike contributions limits, the \textit{Buckley} Court saw expenditure limits as directly limiting the ability of a speaker to communicate his or her own ideas.\textsuperscript{28} Limits on expenditures, therefore, are subject to strict judicial scrutiny, and any regulation of them needs to be narrowly drawn to advance a compelling interest.\textsuperscript{29} The expenditure limits in FECA failed this test.\textsuperscript{30} The government had defended the law’s comprehensive expenditure limits as necessary to promote political equality and to reduce the overall amount of money spent in political campaigns.\textsuperscript{31} The Court found both of these reasons constitutionally insufficient.\textsuperscript{32} Limiting the speech of some in order to enhance that of others was “wholly foreign” to the First Amendment, the Court said, and it was not up to Congress to determine whether the amount of money spent on political speech was wasteful or excessive.\textsuperscript{33} The Court also held that the anti-corruption interest, while compelling, could not justify expenditure limits: as long as political parties and candidates were spending money they had raised in compliance with the contribution limits, expenditure limits served no additional anti-corruption purpose.\textsuperscript{34} Third-party spending likewise posed no risk of quid pro quo corruption, the Court found, as long as done independently of candidates and parties.\textsuperscript{35} All of FECA’s

\textsuperscript{24} \textit{Id.} at 21.
\textsuperscript{25} \textit{Id.} at 25.
\textsuperscript{26} \textit{Id.} at 27–29.
\textsuperscript{27} \textit{Id.} at 143–44. \textit{Buckley} itself was somewhat opaque about the exact standard of review it was applying to contributions limits, which has been clarified by subsequent decisions. See \textit{Nixon v. Shrink Mo. Gov’t PAC}, 528 U.S. 377, 386 (2000) (“Precision about the relative rigor of the standard to review contribution limits was not a pretense of the \textit{Buckley per curiam} opinion.”).
\textsuperscript{28} \textit{Buckley}, 424 U.S. at 19–21.
\textsuperscript{29} See \textit{id.} at 44–45. The “exacting” or strict scrutiny applied by the Court in \textit{Buckley} requires the regulation to be “narrowly tailored” to advance a “compelling” interest. See \textit{First Nat’l Bank of Bos. v. Bellotti}, 435 U.S. 765, 786 (1978) (internal citations omitted); \textit{NAACP v. Button}, 371 U.S. 415, 438 (1963).
\textsuperscript{30} \textit{Buckley}, 424 U.S. at 58–59.
\textsuperscript{31} \textit{Id.} at 48–49, 57.
\textsuperscript{32} \textit{Id.}
\textsuperscript{34} \textit{Buckley}, 424 U.S. at 45–48, 53, 55.
\textsuperscript{35} See \textit{id.} at 45.
expenditure limits therefore were deemed unconstitutional infringements on political speech and struck down.\textsuperscript{36} The Court did uphold FECA’s transparency rules.\textsuperscript{37} The law included two types of transparency requirements. Disclosure rules requiring candidates, political parties, and some groups to publicly disclose most of their contributions and expenditures through mandatory reporting to the FEC,\textsuperscript{38} and disclaimer rules (called “imprint” rules in the United Kingdom) requiring certain spenders to identify on the face of a communication who had authorized and paid for it.\textsuperscript{39} These transparency regulations, the \textit{Buckley} Court held, only indirectly affected speech.\textsuperscript{40} Like contribution limits, they therefore would be subject to less rigorous judicial scrutiny and were sufficiently supported by both the interest in preventing corruption or the appearance thereof, and the interest in avoiding circumvention of the contribution limits the Court had just upheld.\textsuperscript{41}

The Court’s decision upholding disclosure and disclaimer requirements included an important caveat, however, involving the second and third distinctions set out above: the difference between candidates and political parties, and third-party campaigners; and the difference between express and issue advocacy. Under FECA, candidates and political parties are by definition entities whose primary purpose is to influence federal elections.\textsuperscript{42} The \textit{Buckley} Court therefore saw no difficulty in requiring them to regularly report their contributions and expenditures to the FEC. But third-party campaigners engage in many different types of activities, only some of which will influence federal elections.\textsuperscript{43} The Court therefore insisted that any scheme regulating the independent activity of third-party groups distinguish between “express advocacy” to influence federal elections and “issue advocacy” which the Court saw as the type of everyday advocacy around public policy issues that citizens should be able to engage in without becoming entangled in a complex regulatory

\textsuperscript{36} Id. at 58–59.
\textsuperscript{37} Id. at 84.
\textsuperscript{39} See Press Release, Cabinet Office & Kevin Foster MP, Government Safeguards UK Elections (May 5, 2019), https://www.gov.uk/government/news/government-safeguards-uk-elections [https://perma.cc/24YH-NYYW] (“Candidates, political parties and non-party campaigners will also be required to brand or ‘imprint’ their digital election materials, so the public is clear who is targeting them.”); Federal Election Campaign Act § 30120(a). For a discussion of disclosure and disclaimer rules, see R. Sam Garrett, Cong. Research Serv., IF10758, ONLINE POLITICAL ADVERTISING: DISCLAIMERS AND POLICY ISSUES (2019), https://fas.org/sgp/crs/misc/IF10758.pdf [https://perma.cc/W66X-M2UB]. The term “disclaimer” appears to have been coined because of the requirement imposed by some of these rules that the communication clearly state that it is has not been endorsed by a candidate. \textit{See id.}
\textsuperscript{40} Buckley, 424 U.S. at 65.
\textsuperscript{41} Id. at 66–68.
\textsuperscript{42} See Federal Election Campaign Act §§ 30101(2), (16).
\textsuperscript{43} Buckley, 424 U.S. at 67–68.
Moreover, in order to avoid “chilling” pure issue advocacy, the distinction between these two types of speech had to be clearly delineated. The Court acknowledged that creating a bright-line test of this sort would be challenging, but went ahead and did so anyway. In a footnote, it limited the relevant provisions of the law to communications including what came to be known as the “magic words.” Under this test, only communications including words such as “vote for,” “vote against,” “defeat,” or “reject” could be regulated. Consequently, under Buckley, third-party advocacy meeting the magic words test was subject to disclaimer and disclosure rules, but that which did not was not. As discussed below, subsequent legislation and court decisions have tweaked this dividing line, but its fundamental importance continues to drive U.S. law.

The practical effect of Buckley was to create a campaign finance regulatory system that no one intentionally designed and very few people actually like. A law calibrated to restrict both the ability to raise money (through contribution limits) and the need for it (through expenditure limits) became a system in which candidates have an escalating need for money but a tightly restricted ability to access it. Later decisions further held that while all money raised by candidates and political parties was subject to contribution limits, third-party “expenditure-only” groups eschewing express advocacy could raise money in unlimited and often undisclosed amounts. This had the predictable effect of channeling money away from candidates and political parties, and toward less publicly accountable third-party groups. It doing so, it also created a system in which much election-related activity operates outside the federal regulatory scheme.

44 See id. at 80.
45 Id. at 39–51.
47 Buckley, 424 U.S. at 44 n.52.
48 See, e.g., Citizens United, 58 U.S. at 439–40 (Stevens, J., concurring); McConnell, 540 U.S. at 126.
50 “Expenditure-only” groups are third-party campaign groups that spend money but do not make contributions to candidates, political parties, or other groups that do make such donations. See SpeechNow.org v. Fed. Election Comm’n, 599 F.3d 686, 694–96 (D.C. Cir. 2010).
51 In the 2016 U.S. election cycle, outside groups spent nearly $1.4 billion, surpassing the spending of both major parties (whose combined spending totaled only $290 million). Young Mie Kim et al., The Stealth Media? Groups and Targets Behind Divisive Issue Campaigns on Facebook, 35 POL. COMM. 515, 518 (2018).
B. Political Party Funding and Campaign Finance Law in the United Kingdom

Campaign finance regulation in the United Kingdom is in some ways the diametric opposite of that in the United States. Rather than control the availability of funds through limits on contributions to candidates and political parties, the U.K. system controls the need for them by restricting expenditures through spending limits and a ban on expensive broadcast advertising. It also tightly restricts third-party spending and imposes similar transparency rules on third-party campaigners as it does on candidates and political parties. This is very different than the United States system, which strictly limits contributions to candidates, political parties, and some third-party groups; restricts candidate and political party financing more tightly than that of third-party campaigners; and is constitutionally prohibited from limiting expenditures at all.

Three pieces of primary legislation structure the U.K. system. The Representation of the People Act 1983 (RPA) governs constituency-level spending; the Political Parties, Elections and Referendums Act 2000 (PPERA) governs national spending; and the Communications Act 2003 regulates political communications on broadcast television and radio. Most of the restrictions imposed under these statutes apply only during the “regulated period,” which is set by the U.K. Electoral Commission (EC) and usually covers a year prior to a general election or four months prior to a referendum or other election.

The RPA is the oldest of these laws, and many of its provisions have been in place in some form for decades. It focuses on the spending that happens for or against individual candidates within constituency districts. The law sets a

---


base level of permitted spending that is the same for all candidates in all districts, with additional spending permitted on a per-elector basis, allowing candidates in more heavily populated districts to spend more.\footnote{55} It also tightly restricts third-party spending for or against specific constituency-level candidates.\footnote{56}

When the RPA was enacted, its focus on constituency-level spending made sense. Historically, the national expenditures of the two main political parties were small compared to the spending done by candidates in individual districts.\footnote{57} But by the 1990s, political parties had grown in importance and spending had shifted away from constituency districts and to the national, party-driven campaigns. Advocates of reform began arguing that the RPA’s focus solely on constituency-level spending did not adequately address the reality of how U.K. campaigns were funded.\footnote{58} There also was concern about the influence on elected officials of large and undisclosed donations, and what was increasingly seen as excessive spending on campaigns.\footnote{59} PPERA responded to those concerns.

PPERA grew out of the Standing Committee on Standards in Public Life (the Neill Committee), which Parliament had tasked with recommending changes to the United Kingdom’s campaign finance laws.\footnote{60} The Neill Committee made more than 100 recommendations, most of which were adopted by Parliament in PPERA.\footnote{61} Despite the hopes of some reform advocates, these changes did not include imposing caps on contributions. Political contributions in the United Kingdom remain uncapped, and the major parties are funded by membership dues plus a relatively small number of very large donations.\footnote{62}


\footnote{57} NEILL REPORT, supra note 2, ¶ 10.16, at 114.

\footnote{58} See Geddis, supra note 17, at 110–15.

\footnote{59} NEILL REPORT, supra note 2, ¶10.2, at 110.


\footnote{61} NEILL REPORT, supra note 2, ¶¶ 1–100, at 4–14 (outlining the recommendations offered by the committee).

\footnote{62} The Committee tried again in 2011 after all three main British political parties made a commitment in principle to support contribution caps, but was unable to come up with a proposal all parties would agree to. COMM. ON STANDARDS IN PUB. LIFE, POLITICAL PARTY
PPERA did, however, require for the first time that contributions exceeding £5000 (later raised to £7500) be publicly disclosed.  

PPERA also imposed nationwide expenditure limits on political parties and third-party campaigners (the RPA continues to govern the expenditure limits imposed on individual candidates and third-party constituency-level spending). Different limits apply to different spenders in different elections. Calculating the applicable limit turns on things like when in the Parliamentary cycle the spending occurs, how many constituency districts are being contested, and in which part of the United Kingdom the spending takes place. Political parties have higher limits than third-party campaigners, and also have separate limits for general party advocacy versus candidate-specific advocacy. The applicable limits for referendums (like the Brexit vote) are set by the EC. As in the United States, to avoid circumvention of these limits, coordinated spending is treated differently. Spending done in coordination with a political party (or “lead campaign group” in a referendum) is counted toward the spending cap of the party or lead group. If other registered campaigners coordinate with each other, their spending is considered that of one group, and their combined total must stay within the applicable limit as such.

The third significant statute in the U.K. regulatory scheme is the Communications Act 2003. The 2003 law updated and continued the long-standing ban in the United Kingdom prohibiting political advertising on

FINANCE: ENDING THE BIG DONOR CULTURE, 2011, Cm. 8208, at 8–9 (UK) [hereinafter POLITICAL PARTY FINANCE].
63 STUART WILKS-HEEG & STEPHEN CRONE, FUNDING POLITICAL PARTIES IN GREAT BRITAIN: A PATHWAY TO REFORM 13 (2010).
65 Id. c. 41, §§ 79, 94, schs. 9, 10.
66 Id. These limits were tied to the proportion of the vote the party received at the most recent general election, which was the 2015 general Westminster Parliament election.
67 For example, in 2011 in a general election for the Westminster Parliament in which all constituency districts were being contested, the spending cap on each of the political parties was £19.5 million. POLITICAL PARTY FINANCE, supra note 62, at 30. The national spending cap for registered third parties (interest groups intending to spend more than £10,000 to attempt to influence an election) was just under £800,000 in England, and the third-party limit for candidate-specific spending in the constituency districts was £500. Id. at 31–32.
68 For referendum campaigns, the EC uses a statutorily defined process to identify a “lead campaign group” for each side of the debate. See THE ELECTORAL COMM’N, THE DESIGNATION PROCESS 4 (Mar. 2016), https://www.electoralcommission.org.uk/sites/default/files/pdf_file/Designation-process-for-the-EU-referendum.pdf [https://perma.cc/E73H-TRXV] [hereinafter DESIGNATION PROCESS]. The lead campaign group has a significantly higher spending limit than other registered groups. For Brexit, the lead campaigns had a limit of £7 million while the limit for other registered campaign groups was £700,000. 2016 EU REFERENDUM, supra note 5, at 91.
69 See DESIGNATION PROCESS, supra note 68, at 6.
70 See id.
broadcast media. Unlike the RPA and PPERA, this restriction is in effect at all times, not just during the regulated election period. Under the Act, recognized political parties contesting seats in a requisite number of constituency districts are given free broadcast time during the lead up to an election, but no other political ads can be legally broadcast on public or private television or radio at any time. The ban is far-reaching, and applies not only to political party ads, but also to third-party ads “directed towards a political end.”

The combined effect of RPA, PPERA, and the Communications Act 2003 is a regulatory system very different from that in the United States. Those differences are significant—the Communications Act 2003, for example, would be plainly unconstitutional under United States law. As shown below, however, the reform proposals being considered in each nation only rarely implicate these differences, allowing ample room for constructive comparative work.

IV. TAXONOMY OF REFORM PROPOSALS

Regulators, lawmakers, and reform advocates in the United States and United Kingdom have generated numerous proposals to more effectively regulate online campaigning and foreign interference in domestic elections. Some of these proposals—such as tightening data privacy rules and imposing antitrust restraints on social media companies—do not directly engage campaign laws and are not discussed here. Others, however, touch on controversies central to those laws, such as how to define campaign-related speech, how to balance personal privacy with public accountability, and how to ensure fair elections without infringing on the freedom of expression essential to a functioning democracy.

This Part addresses those proposals. In doing so, it organizes them into a common taxonomy, and examines the legal rules governing each class of proposal in each nation. Grouping the proposals into this organizational scheme brings a systemic coherence to comparative work in this area by contextualizing similar proposals within the specific legal rules under which they will be evaluated. This enables more of an “apples-to-apples” evaluation of the legal challenges each type of reform will face in each nation, allowing us to see more clearly which proposals are worthy of additional comparative study, and which are not.

The taxonomy classifies reform proposals according to their underlying goals. There are four principle goals: better educating the public about digital literacy (public education); enhancing the transparency of online campaigning (transparency); reducing the influence of foreign interests over voters’ choices

---


72 Id. §§ 319, 321, 333.

73 Id. § 321(2)(a)–(c).
(source exclusions); and excluding deceptive or otherwise harmful content from online distribution (content exclusions). Not all reform proposals, of course, fit neatly into this taxonomy, nor does the analysis that follows discuss every proposal made within each category; the taxonomy is exemplary, not comprehensive. But organizing the most common proposals this way efficiently facilitates comparative consideration.

As shown below, these four types of reform proposals can be visualized as a pyramid. The reforms at the base of the pyramid (public education) face few system-specific legal challenges and therefore are the most comparable across systems; reforms at the peak (content exclusions) face the most such challenges, and therefore offer fewer opportunities for productive comparative analysis (although even here there are areas in which the two nations can learn from each other’s experiences).

Public education proposals, at the base of the pyramid, do not directly involve campaign laws but are included because their goal is to make voters more critical consumers of political messages they see online. In the United States, these proposals include federal efforts such as the expansion of the State Department’s counterterrorism mission to include combating disinformation, and efforts such as those in California, Massachusetts, and Washington to teach digital media literacy in schools.74 In the United Kingdom, this category includes things like proposals to use a social media tax to fund online literacy

74 Alex Stamos et al., Combatting State-Sponsored Disinformation Campaigns from State-Aligned Actors, in STANFORD CYBER POLICY CTR., supra note 7, at 44 (internal citations omitted).
programs, and efforts to promote awareness of the standards of professional journalism. Public education proposals like these are unlikely to face significant legal challenges in either country and are only briefly discussed below.

Transparency proposals aim to help voters make more informed choices by ensuring they understand who is promoting or paying for the political messages they see. Transparency proposals are plentiful in both the United States and the United Kingdom, and include recommendations to require disclaimers (imprints) on online ads, require more detailed reporting of online expenditures, and change the type and nature of online spending disclosed to regulatory bodies. The laws governing transparency rules in the United States and United Kingdom are different, and these differences mean reforms in this category will face distinct legal challenges in each nation. Examining these proposals through the structure provided by the taxonomy allows us to see that these legal differences are not as relevant to the reforms being proposed as is frequently assumed. This category therefore provides extensive opportunity for genuinely valuable comparative consideration.

Source exclusions regulate political communications based on who is speaking, promoting, or paying for the communication. The goal of reform proposals in this class is to limit the influence of foreign interests on domestic elections by precluding foreign funding of election-related communications. Analyzing each nation’s law through the taxonomy reveals that source exclusions will face significantly different legal challenges in the United States and United Kingdom, but that there is sufficient common ground even here to make comparative study useful.

Content exclusions are designed to exclude or reduce harmful communications online. What is considered “harmful” varies in these proposals, and ranges from things already regulated if done offline (such as defamation, harassment, fraud, or abuse) to more controversial efforts to limit the online spread of propaganda, disinformation, and other content considered detrimental to democratic discourse. Context exclusions, at the apex of the pyramid, will face the most system-specific legal barriers, many of which are likely to be insurmountable in the United States.

The remainder of this Part will discuss in detail the current rules governing each of these categories in the United States and United Kingdom, and the legal

---


challenges reforms in each category of proposal are likely to face in each country. In doing so, it also will identify the specific areas where further comparative study will be most beneficial.

A. Public Education

Rather than directly regulate campaign speech, public education proposals seek to make voters more critical consumers of the political messages they see. This type of public education effort is unlikely to encounter significant legal barriers in either the United States or the United Kingdom. It nonetheless is included here because of its fundamental importance in combatting online and foreign election interference. Effective regulation of election-related speech is devilishly difficult, and will be so even if campaign laws are fully updated and regulators fully engaged. Increasing awareness of the problem through public education efforts is therefore critical.

Both the United States and the United Kingdom have taken small steps in this area. In the United States, the mission of the State Department’s Global Engagement Center has been expanded to include countering foreign disinformation more broadly.77 The U.S. Intelligence Community also has made efforts to increase public awareness of the problem. The former Director of National Intelligence,78 the Director of the Federal Bureau of Investigation,79 the former Homeland Security Secretary,80 and the former Special Counsel to the U.S. Attorney General all have made public statements warning Americans of ongoing efforts by foreign actors to use social media platforms to inflame political tensions and influence U.S. elections.81 In the United Kingdom, former


Prime Minister Theresa May has done so as well.\textsuperscript{82} Public education efforts also have been endorsed by most of the commissions and committees examining the issue in the United Kingdom, including the Electoral Commission,\textsuperscript{83} the Information Commission, the Committee on Standards in Public Life,\textsuperscript{84} and the House of Commons Digital, Culture, Media and Sports Committee.\textsuperscript{85}

While these efforts have varying degrees of political support, they raise few serious legal issues in either the United States or the United Kingdom. Within broad limits, governments in both countries are free to engage in public information campaigns. In the United Kingdom, this type of effort would encounter no discernable legal challenges.\textsuperscript{86} In the United States, public information campaigns are governed by the “government speech” doctrine. The core tenet of this doctrine is that when the government itself is speaking, it is allowed to advocate for its preferred position. So, for example, the government can fund an anti-smoking campaign without also having to fund pro-smoking messages. The doctrine is underdeveloped in several ways,\textsuperscript{87} but as long as it is clear the government is the entity speaking, a public information campaign designed to increase digital literacy among voters would be unlikely to encounter significant legal challenge even in the United States. The remainder of this Article therefore will focus on the other three classes in the taxonomy: transparency, source exclusions, and content exclusions.

\section*{B. Transparency}

Transparency is the largest and most diverse class in the taxonomy. When considering transparency proposals and the rules governing them, it is helpful to remember the distinction mentioned above between disclosure and disclaimer (imprint) rules. Disclosure rules require candidates, political parties, and some

\begin{footnotesize}
\begin{itemize}
  \item \textsuperscript{82} Prime Minister Theresa May, Speech to the Lord Mayor’s Banquet (Nov. 13, 2017), https://www.gov.uk/government/speeches/pm-speech-to-the-lord-mayors-banquet-2017 [https://perma.cc/S95H-WZ9P].
  \item \textsuperscript{85} The DCSM Report identifies digital literacy as a “fourth pillar” of education, along with reading, writing, and math. DCSM REPORT, supra note 5, at 87.
  \item \textsuperscript{86} DIGITAL CAMPAIGNING, supra note 83, ¶¶ 104–25, at 20–23. There are limits in the United Kingdom on governmental spending to promote particular outcomes in referendums. See generally 2016 EU REFERENDUM, supra note 5, at 90.
  \item \textsuperscript{87} There are unresolved questions about how to distinguish the government’s own speech from governmental funding for the speech of others, and whether it is legally relevant that the recipients of the message understand they are hearing a government-provided communication. There also are tangentially related prohibitions and norms against using federal funds to distribute propaganda directed at U.S. citizens.
\end{itemize}
\end{footnotesize}
groups to report their income and expenditures to a regulatory body. Disclaimer rules require certain communications to carry on their face information about who authorized or paid for the communication. Disclosure rules advance transparency by informing the public about who supports candidates for public office and who those candidates may be indebted to if elected; disclaimer rules advance transparency by informing the public about the source of the political messages they are seeing.88 The most prominent reform proposals in this class involve strengthening disclosure rules by requiring more detailed reporting about the financing of such communications, and expanding disclaimer rules to cover more online communications.

The breadth and variety of online communications potentially covered by disclosure and disclaimer rules is what creates the legal challenges in this category. Any transparency rule, whether it be about disclaimers or disclosures, must define the communications it covers. This is challenging even when targeting traditional campaign communications, and becomes more so in the fluid world of online social media. Relatedly, lawmakers also must decide if online transparency rules should mirror offline rules, or if the differences between formats warrant distinct regulatory approaches. As shown below, while the United States and United Kingdom regulate disclosures and disclaimers quite differently, each country struggles with these same questions.

1. Transparency in the United States

The federal statutory law governing disclosure and disclaimer in the United States is found in two statutes: the Federal Elections Campaign Act (FECA, discussed above), and the Bipartisan Campaign Reform Act (BCRA, which amended FECA in 2002).89 FECA requires political parties, candidates, and certain third-party campaigners to register and file regular reports with the FEC. Political parties and candidates running for federal office must register with the FEC when they raise or spend over a threshold amount in connection with a federal election.90 Third-party campaigners must register as “political committees” (more commonly known as political action committees or “PACS”) when their major purpose is to influence federal elections.91

Entities required to register with the FEC (FEC-registered groups) must file regular reports with the FEC specifying their contributions and expenditures. These reports are filed electronically at quarterly or monthly intervals, although

---


expenditures on certain types of communications must be reported within twenty-four hours.\textsuperscript{92} FEC-registered groups also must include disclaimers on most of their public communications, which are defined by the FEC as any general public political advertising, including broadcasts, newspaper and magazine ads, and internet ads “placed for a fee” on another person’s website.\textsuperscript{93} Consequently, groups that register with the FEC are by definition subject to extensive transparency rules. Their contributions and expenditures are disclosed in regular public reports, and most of their communications include disclaimers stating who authorized and paid for the communication.\textsuperscript{94}

But not all groups that make campaign-related communications are required to register as political parties or committees with the FEC and therefore are not subject to these regular disclosure and disclaimer rules.\textsuperscript{95} The most significant of these groups are “social welfare” groups.\textsuperscript{96} Social welfare groups are a category defined by the Internal Revenue Service (“IRS”, the U.S. federal tax agency) for taxation purposes.\textsuperscript{97} They often are referred to by the IRS code provision that defines them: 501(c)(4). Under the applicable statute, to qualify for 501(c)(4) status, an organization must not be organized for profit and must be operated “exclusively” to promote the social welfare.\textsuperscript{98} Regulations promulgated under the statute, however, permit social welfare groups to engage in political activity as long such efforts do not constitute their “primary” activity.\textsuperscript{99}

This has permitted extensive use of 501(c)(4) status by politically active groups wishing to avoid FEC regulation. Because they are not defined as political committees for purposes of federal campaign law, they are not subject


\textsuperscript{93} 52 U.S.C. § 30101(22) (Supp. IV 2017); 11 C.F.R. § 100.26 (1996). This provision is discussed in greater detail below.


\textsuperscript{95} See 52 U.S.C. § 30101(4) (defining a political committee); \textit{see also id.} § 30120 (requiring political committees to disclose certain funding and authorization sources for certain public communications).

\textsuperscript{96} See 26 C.F.R. § 1.501(c)(4)(a)(2)(i) (2019) (describing the nature of a “social welfare” organization as one being “primarily engaged in promoting in some way the common good and general welfare of the people”).

\textsuperscript{97} \textit{See id.} § 1.501(c)(4)(a)(1) (noting a disclosure and disclaimer exemption for civic organizations operating “exclusively for the promotion of social welfare”).


\textsuperscript{99} \textit{Lunder & Whitaker, supra} note 98, at 3.
to FECA’s regularized disclosure and disclaimer rules.\(^{100}\) Instead, their political communications are regulated by separate transparency rules developed by the FEC for unregistered groups that nonetheless engage in some election-related communications.\(^{101}\) Drawing on the distinction made by the Supreme Court in *Buckley*, these groups need only report expenditures for express advocacy and contributions “earmarked” for that advocacy.\(^{102}\) A communication is express advocacy for these purposes when it includes the *Buckley* magic words or the functional equivalent thereof.\(^{103}\) A contribution is earmarked when it is designated by the donor as given to fund a particular communication.\(^{104}\) FEC disclaimer rules developed for these groups track this paradigm, and apply only to those of their public communications that expressly advocate for the election or defeat of a candidate for federal office.\(^{105}\)

The result of all this is that under FECA, only FEC-registered groups are subject to regular disclosure and disclaimer requirements, while non-FEC registered groups such as 501(c)(4)s can engage in significant political communications while avoiding most disclosure and disclaimer rules. This creates an obvious transparency gap. Entities like social welfare groups can avoid regular FEC regulation and reporting requirements by limiting their election-related advocacy to less than fifty percent of their activity, while also avoiding targeted regulation by avoiding words of express advocacy even in communications intended to influence federal elections.

BCRA attempted to partially close this gap by bringing an additional category of speech into the U.S. disclosure and disclaimer regime. As required by *Buckley*, BCRA uses a bright-line test to define the category of speech being regulated. BCRA defines the communications it regulates—“electioneering communications”—as any “broadcast, cable, or satellite communication” that refers to a clearly identified candidate for federal office, is publicly distributed thirty days before a primary or sixty days before a general election, and is targeted to the relevant electorate.\(^{106}\) Under BCRA, electioneering communications, like express advocacy and the public communications of FEC-registered groups, must carry disclaimers identifying who is responsible and paying for the communication.\(^{107}\) Entities who spend $10,000 or more a year

---

100 See id.
101 Id. at 11. See generally Torres-Spelliscy, supra note 88 (discussing the history of campaign finance disclosure law and two exemptions to those disclosure laws).
103 This test, which extends slightly beyond the “magic words” test of *Buckley*, was developed by the Supreme Court in 2007, in *FEC v. Wisconsin Right to Life*, 551 U.S. 449, 453 (2007). *Wisconsin Right to Life* imposed a narrowing construction on the definition of electioneering communications, holding that the provision was unconstitutional unless read to apply only to expenditures that could not “reasonably be viewed” as anything other than urging the support or defeat of a candidate for federal office. *Id.* at 474.
104 11 C.F.R. § 110.6(b)(1) (2019).
107 *Id.* § 30104(f)(1)–(2).
producing or placing electioneering communications also must file a disclosure statement with the FEC identifying the names and addresses of those who have contributed more than $1000 to fund its communications.\footnote{Citizens United v. FEC and the Future of Federal Campaign Finance Reform, LIBR. CONGRESS, https://www.loc.gov/law/help/citizens-united.php [https://perma.cc/2EV7-MXDU] (last updated Aug. 16, 2019).}

The U.S. Supreme Court upheld these transparency provisions in \textit{Citizens United v. FEC}.\footnote{Citizens United v. Fed. Election Comm’n, 558 U.S. 310, 370–72 (2010).} Decided in 2010, \textit{Citizens United} is better known for striking down a ban on the use of corporate general revenue funds to fund a corporation’s independent expenditures.\footnote{Discussed \textit{infra} Part IV.C.1.} But the case also addressed the constitutionality of disclosure and disclaimer requirements as applied by BCRA to electioneering communications.\footnote{\textit{Citizens United}, 558 U.S. at 370–72.} The case is important for U.S. transparency law, because the core question presented was whether transparency regulations could be applied to political communications that did not constitute express advocacy under \textit{Buckley} and its progeny.\footnote{\textit{Id.} at 318–19.}

By an 8–1 vote, the Court upheld the transparency requirements.\footnote{\textit{Id.} at 370–72.} In doing so, it found that disclosure and disclaimer rules, like campaign contributions, are not direct prohibitions on speech and therefore need only be supported by a substantial (rather than compelling) interest.\footnote{\textit{Id.} at 366–67.} The Court further held that “shedding the light of publicity”\footnote{\textit{McConnell v. Fed. Election Comm’n}, 540 U.S. 93, 231 (2003) (quoting \textit{Buckley v. Valeo}, 424 U.S. 1, 81 (1976)).} on who is financing political speech is such an interest, as is providing the electorate with information sufficient to ensure that voters are fully informed about who is speaking.\footnote{\textit{Citizens United}, 558 U.S. at 370–71.} This type of information, the Court said, “enables the electorate to make informed decisions and give proper weight to different speakers and messages.”\footnote{\textit{Id.} at 371. An earlier decision, \textit{McIntyre v. Ohio Elections Comm’n}, 514 U.S. 334 (1995), had cast doubt on this informational interest, as applied to handmade leaflets distributed by an individual at a local meeting. The cost of preparing and distributing the leaflets at issue in \textit{McIntyre} was negligible, and would have fallen well below the applicable reporting thresholds upheld in \textit{Buckley} and subsequent cases. \textit{See McIntyre}, 514 U.S. at 337 (noting that the defendant had “composed and printed [the leaflets] on her home computer and had paid a professional printer to make additional copies”).} Importantly, the Court was clear that these transparency requirements can be imposed even when the communication being regulated does not constitute express advocacy and the group speaking is not otherwise regulated by the FEC.\footnote{\textit{Citizens United}, 558 U.S. at 369.}

Assuming appropriate exceptions are available to protect the privacy of smaller donors and to allow for as-applied challenges for entities for whom
disclosure poses a risk of serious harassment (limitations imposed by the Court in earlier cases), the U.S. Supreme Court has therefore allowed disclosure and disclaimer requirements to attach to a wide array of political communications. This may not continue: the composition of the Supreme Court is changing rapidly, and transparency rules are being challenged across the United States. But at least under existing law, disclosure and disclaimer requirements could be applied to many more online communications than they currently are.

This result has been stymied, however, by BCRA itself and by the FEC. The text of BCRA only applies to “broadcast” communications, and the FEC has interpreted the statute as not applying to other media, including newspapers, magazines, telephones, and the internet. This means the only statutory transparency requirements currently applicable to online communications are the more limited provisions found in FECA, which, as discussed above, only apply to the public communications of FEC-registered groups and the express advocacy of groups not otherwise regulated by the FEC. Neither the FEC nor Congress has as of yet expanded FECA’s coverage to include the broader category of electioneering communications as defined in BCRA. This means that online “issue ads” (ads not including words of express advocacy) run by groups not regulated by the FEC as political parties, candidates, or political committees are not subject to any disclosure or disclaimer laws.

Additionally, the FEC also has been slow to extend even the limited disclosure and disclaimer required by FECA to online communications. Instead, it has created a situation in which even ads that would require disclaimers if appearing offline (because they are the public communications of FEC-registered groups or the express advocacy of other groups) are not always required to carry disclaimers when distributed online. The FEC has

121 See Hasen, supra note 119, at 561–62 (illustrating what the Supreme Court in its changing composition has recently required to grant an as-applied exception to otherwise permissible disclosure requirements).
123 See Daniel W. Butrymowicz, Note, Loophole.com: How the FEC’s Failure to Regulate the Internet Undermines Campaign Finance Law, 109 Colum. L. Rev. 1708, 1709 n.4 (2009) (“Whatever the reason, the FEC has, over the last several years, shown a consistent desire to not regulate the internet.”).
accomplished this through a series of regulatory decisions and advisory opinions. One of the first such decisions came in 2002, when the FEC determined that paid text message ads were exempt from an otherwise applicable disclaimer requirement under a “small-items” exception developed for things like campaign buttons and bumper stickers. Two years later, the FEC expanded on this idea by arguing in a lawsuit that it had administrative discretion to categorically exclude all digital communications from FECA’s disclaimer rules. When that position was rejected in court, the FEC decided that FECA disclaimers would only be required on digital ads “placed for a fee” on the “website” of another. This meant that the requirement did not apply to communications distributed for free online, regardless of the cost of producing the content involved. It also meant the requirement did not extend to non-web-based platforms, such as mobile apps. That requirement was then even further diluted in 2011, when the FEC deadlocked on whether paid ads on Facebook required disclaimers. This non-decision allowed Facebook to host even express advocacy ads without disclaimers until 2017, when the FEC finally issued an opinion stating that paid advertisements on Facebook were required to carry disclaimers when constituting express advocacy or placed by an FEC-registered group.

Only in 2018, in the wake of revelations about the 2016 election, did the FEC slightly shift course and propose two draft regulations designed to expand the transparency of online election-related communications. The drafts are

---

124 Advisory Opinions apply only to the specific circumstances presented. They offer guidance to similarly situated entities, but do not have the certainty of law or promulgated regulations. GARRETT, supra note 39.


129 See Jennifer Valentino-DeVries, I Approved This Facebook Message—But You Don’t Know That, PROPUBLICA (Feb. 13, 2018), https://www.propublica.org/article/i-approved-this-facebook-message-but-you-dont-know-that [https://perma.cc/MG2Z-WYQN] (“[T]he six-person FEC couldn’t muster the four votes needed to issue an opinion, with three commissioners saying only limited disclosure was required and three saying the ads needed no disclosure at all, because it would be ‘impracticable’ for political ads on Facebook to contain more text than other ads.”).

130 A 2018 study by ProPublica indicated this rule was rarely followed and noncompliance was rarely punished. See id. (discovering that fewer than 40 of 300 Facebook ads had the FEC-required disclaimers).

131 Internet Communication Disclaimers and Definition of “Public Communication,” 83 Fed. Reg. 12,864 (proposed Mar. 26, 2018) (to be codified at 11 C.F.R. pt. 100). The Notice of Proposed Rulemaking requires all commenters to provide their name, city, and state. Id. at 12,864. Presumably, the Commission wants to know who is attempting to influence its
substantially similar. The most significant difference between the proposals is their alternative compliance standards for communications where full disclosure is considered impossible or impractical. \textit{Id.} at 12,879.

If adopted, either of these rules would bring more transparency to paid online political communications in the United States. Both drafts are limited in that they only extend to online ads the current disclaimer rules applicable under FECA, not the broader rules enacted in BCRA. So, the new online disclaimer requirements would apply only to the public communications of FEC-registered groups and the express advocacy of other groups, meaning entities like social welfare groups will continue to be able to run even paid online advertisements without disclaimers as long as they avoid words of express advocacy. The drafts are similarly limited on the disclosure side. Both drafts require disclosure of the identity of the entity paying for the ad, but continue the requirement that only earmarked contributions to that entity need be disclosed. So, while FEC-regulated groups would continue to have to disclose virtually all of their contributors, the draft rules would allow other groups to run even paid express advertisements without disclosing their underlying funders unless the funder specifically designates his or her donation as for a particular express ad.

An additional limitation of the draft rules is the continued application of the transparency rules only to online communications “placed for a fee” online. Restricting online transparency rules to paid placements means the only expenses counting toward the threshold—triggering regulation are those paid to the platform hosting the ad. Since online advertising is significantly less expensive than its offline counterparts, this means increasing numbers of even paid placements could fall below the reporting threshold. More significantly,
it means that extensive costs incurred in producing political communications—like the millions of dollars spent to train and employ workers at the St. Petersburg troll factory—do not count toward the applicable threshold. Other reform proposals being considered in the United States, such as the DISCLOSE Act and the Honest Ads Act, also hew to this more conservative path.

Again, this restrictive approach is not required by current U.S. law. As discussed above, *Citizens United* explicitly upheld the disclaimer requirements BCRA imposed on electioneering communications (broadcast communications that run in relevant time period and clearly identify a candidate for federal office) even when those communications do not constitute express advocacy. BCRA, unlike FECA, also uses cost of production—not cost of placement—as the triggering threshold for its transparency measures. So increasing transparency by extending BCRA’s definition of “broadcast communications” to include online communications should be within the constitutional parameters set by the Supreme Court in *Citizens United*. Lower federal court decisions recognize this. A federal district court recently struck down an attempt by the FEC to restrict BCRA’s transparency requirements to paid placements and exempt social welfare organizations from BCRA’s rules entirely. Additionally, the FEC’s earmarking rule, currently used by the FEC to restrict the reach of FECA’s disclosure requirements, has been struck down by a district

---

140 This “troll factory” refers to the Internet Research Agency in St. Petersburg, an agency who many commentators argue influenced the outcome of the UK-EU referendum. Clare Llewellyn et al., *For Whom the Bell Trolls: Shifting Troll Behaviour in the Twitter Brexit Debate*, 57 J. COMMON MKT. STUD. 1148, 1148 (2019).


144 *Shays v. Fed. Election Comm’n*, 337 F. Supp. 2d 28, 128 (D.D.C. 2004). The court held the statute clearly intended the rules to apply to unpaid placements as long as the overall expenditure involved in distributing the advertisement exceeded $10,000. *Id.* at 129. The court also struck down the FEC’s categorical exclusion of any electioneering communications placed by IRS regulated entities. *Id.* at 126–27; see also Del. Strong Families v. Att’y Gen. of Del., 793 F.3d 304, 308–09 (3d Cir. 2015) (“[I]t is the conduct of an organization, rather than an organization’s status with the Internal Revenue Service, that determines whether it makes communications subject to the Act.”).
court as inconsistent with FECA itself. These rulings indicate that there is space within existing law to significantly increase the transparency of online communications in the United States.

2. Transparency in the United Kingdom

Unlike their U.S. counterparts, legislators in the United Kingdom work relatively free of judicial restraints on their ability to impose transparency rules on election-related communications. They nonetheless have struggled with similar questions of how to define the communications they are regulating, and whether to treat online communications the same or differently than their offline equivalents. As discussed below, both systems also have failed to require as much transparency as would be legally permissible under their respective regulatory systems.

The Political Parties, Elections and Referendum Act 2000 (PPERA) is the most significant primary legislation regulating transparency in U.K. campaigns. As noted above, PPERA limits the campaign expenditures of national political parties and third-party campaigners. It also requires political parties and third-party campaigners to register with the EC, and to file regular disclosure reports itemizing their contributions and expenditures.

Whether a third-party campaigner is required to register with the EC depends on whether the group intends to spend more than a threshold amount on “regulated campaign activity.” Regulated campaign activity includes activities that pass the “purpose test.”


147 Id. §§ 41–69.

148 Id. § 28.

149 Id. § 62.


151 Id. A 2016 review of the regulation of third-party campaigners in the United Kingdom proposed replacing the purpose test with an actual intent requirement. The Lord Hodgson of Astley Abbotts, supra note 139, at 6. Parliament has not acted on this proposal.

category of candidate; and that those efforts are aimed at, seen, heard by, or involving the public. If an activity meets this test, costs counting toward the reporting threshold include not just the cost of placing the ad, but all costs involved in its production, publication and distribution.

The Communications Act 2003 also defines campaign-related activity, for the purpose of enforcing its prohibition on the broadcast of political advertising. As noted above, the scope of communications prohibited by the broadcast ban is broad, and includes not just communications from political parties, but any communication intended to influence elections, legislators, or the public on matters of public dispute. This is the most expansive definition of election-related communications in U.K. election law. As such, it offers a useful test of the willingness of British and European courts to tolerate far-reaching regulation of political communication.

It was subject to just such a test in 2008, when a group called Animal Defenders International challenged the broadcast ban under Article 10 of the European Convention on Human Rights. Article 10 protects freedom of expression, and was incorporated into U.K. domestic law through the Human Rights Act 1998.

Animal Defenders wanted to broadcast an advertisement on the BBC. The ad featured an image of a girl chained in a cage morphing into a chimpanzee while a voiceover provided information about the similar capabilities of chimpanzees and young children. The objective of the campaign, according to the group, was not to influence elections but to draw public attention to the use of primates for research and recreational purposes. When the BBC refused to air the ad, Animal Defenders sued.

The U.K. Supreme Court (sitting at the time as the Lords of Appeal in the House of Lords) held the ban was not incompatible with Article 10, even as

---

153 OVEVIEW OF REGULATED NON-PARTY CAMPAIGNING, supra note 150, at 4–5. In the Brexit referendum, it was defined as activity “intended to, or are otherwise in connection with, promoting or bringing about a particular outcome in the referendum.” JOHNSTON & WOODHOUSE, supra note 152.

154 See OVEVIEW OF REGULATED NON-PARTY CAMPAIGNING, supra note 150, at 10 (noting that spending on social media that meets the purpose test will require accounting of all the described costs).


156 OVERVIEW OF REGULATED NON-PARTY CAMPAIGNING, supra note 150, at 4–6 (overviewing the regulations applying to third-party campaigner communications).

157 R (Animal Defenders International) v. Secretary of State for Culture, Media and Sport [2008] UKHL 15, [1], [2008] 1 AC 1312 (appeal taken from Eng.).


161 Id. at [50].

162 Id. at [3].

163 Id. at [4].

164 Id. at [46].
applied to Animal Defenders’ proposed communication. The court recognized that public scrutiny of different “views, opinions and policies” is essential to the democratic process, but believed that such scrutiny was best achieved by allowing Parliament to enact legislation ensuring a balanced presentation of competing ideas, especially on television. The court did recognize, consistent with European jurisprudence, that Article 10 requires restrictions on expression be proportionate to their goals, and that the expansive definition of “political advertising” in the Communications Act could be considered overly broad. But the difficulty of drawing clear lines in this area, the court said, meant that the court should defer to the considered judgment of Parliament. The European Court of Human Rights agreed, and allowed the broadcast ban to stand.

The key point of Animal Defenders for current purposes is that a very expansive definition of political advertising was upheld by both U.K. and European courts, using a deferential standard of review. This means equally expansive disclosure and imprint (disclaimer) laws—which are much less restrictive of political expression than the broadcast ban—would likely face few judicial barriers in the United Kingdom. Despite this, lawmakers in the United Kingdom, like their counterparts in the United States, have not yet expanded their transparency rules to capture the full scope of communications decisions like Animal Defenders leave open to them.

In regard to disclosure, U.K. regulators limit the scope of the nation’s reporting regime by only requiring political parties and third-party campaigners intending to spend more than a set amount on regulated campaign activities to register with the EC and provide regular reports of their contributions and expenditures. While these registration thresholds are in the same range as their U.S. counterparts, the U.K. disclosure system overall is quite different. Regulated groups in the United Kingdom are only required to disclose the

---

165 Id. at [36].
166 Animal Defenders International [2008] UKHL 15, [28]; see also Rowbottom, supra note 2, at 1–2.
168 See id. at [6], [31].
169 Id. at [33].
identity of donors who donate more than £7500 to the group. This is much higher than the $200 triggering public disclosure in the United States. The timing of the required reporting also is different. Political parties in the United Kingdom must provide regular contribution reports during the run-up to a general election, but they do not need to report expenditures until three or six months after an election (with groups that spend more having longer to report). Registered third-party campaigners, in turn, are only required to report contributions and spending during the regulated period, and third-party campaigning outside the regulated period is not reported at all. Current reporting rules also allow expenditures to be lumped in unhelpful ways, making it difficult to trace online spending through the reports, and the EC has only limited authority to compel third-party campaign groups to disclose the underlying source of the funds they receive.

There also are significant differences between the United States and United Kingdom in regard to imprint (disclaimer) requirements. Imprint rules in the United Kingdom are in several ways more extensive than those found in the United States. Any person or group distributing to the public material meeting the purpose test must include an imprint on the material, whether or not they are required to register with the EC. There is no minimum spending threshold triggering this obligation. Because the purpose test itself is quite broad, this means that imprints are required on many types of communications in the United Kingdom than in the United States. But there is one big exception to this: as of this writing there is no imprint requirement applicable to online

---

175 OVERVIEW OF PARTY CAMPAIGN SPENDING, supra note 53, at 15; see also DIGITAL CAMPAIGNING, supra note 83, ¶ 76.
177 See DIGITAL CAMPAIGNING, supra note 83, at 3 (recommending further itemized reports to increase third-party transparency in digital spending).
179 See id. The test requires that covered material be distributed to the public, which imposes a practical limitation on the breadth of the rule. Id.
communications. The U.K. government has agreed to develop rules requiring online imprints, but has not yet done so. As the robust online campaign during the Brexit referendum made clear, this leaves a gaping hole in the United Kingdom’s transparency regime.

Scotland experimented with plugging that hole during the 2015 vote on Scottish independence. The parliaments of Scotland, Wales, and Northern Ireland enjoy certain devolved powers, including the ability to regulate local elections. Exercising this power, the Scottish Parliament required any material “wholly or mainly related” to the independence referendum to include an imprint, regardless of whether the material was distributed on or off line. The “wholly or mainly related” test was even broader than the purpose test used in nationwide elections, and there was no threshold spending requirement—any communication meeting the test was required to carry an imprint. This created a sweeping online imprint requirement.

It was only partially successful. In a report submitted to Parliament after the Scottish referendum, the EC noted that the scope of the disclosure requirement meant that a potentially wide amount of online material was captured by the rule, and that this had created confusion among campaigners about what communications were in fact required to carry imprints. In particular, the EC reported receiving questions about whether communications posted on personal Facebook and Twitter accounts were within the scope of the rule, and if so where on these pages the imprints should appear. The EC’s response was to advise campaigners that social media accounts “focused primarily on campaigning” needed to carry imprints on their Facebook homepage or Twitter profile, but that...
individuals or organizations who were “just expressing their views” through their own accounts were not covered by the requirement.\textsuperscript{187}

As the EC acknowledged in its post-election report, the distinction between using a social media account to focus primarily on campaigning versus using one to simply express your own views is far from crisp, and will require clarification if online imprint requirements become the norm in U.K. elections.\textsuperscript{188} Nonetheless, the Scottish referendum provides a useful first look at a real-world effort to provide greater campaign transparency online.

3. Transparency Recap

As evidenced by the number and prominence of proposed reforms in this category, increasing transparency around online communications is an important tool in combatting disinformation and foreign interference in democratic elections. Fortunately, U.S. and U.K. laws regulating transparency are sufficiently similar to allow each country to learn from the other in this area. In both countries, disclaimer and disclosure rules can be triggered either by who you are or what you say. Both countries also require that candidates, political parties, and some third-party campaigners report their income and expenditures to an agency overseeing election activity. Both countries likewise require many political communications to carry disclaimers, but not all election-related communications are included in those requirements, especially when occurring online.

The legal challenges facing each country in this area also are surprisingly similar. Both countries are struggling with how to design online disclaimer and disclosure requirements that balance the regulatory burden imposed by such requirements with the need for public awareness of who is authorizing and paying for online campaigns. The United States can gain valuable insights by in-depth study of the groundbreaking Scottish experiment with online imprints, while the United Kingdom has much to learn from careful consideration of the long path already trod by U.S. courts and legislatures trying to distinguish regulated campaign communications from unregulated political expression. These are not easy issues, and the experiences of the two countries have much to offer each other in evaluating them.

The United States also could learn from the United Kingdom’s experience using pay-to-produce rather than a pay-to-place thresholds for online disclosure and disclaimer requirements. Online advertising is significantly less expensive than its offline counterparts, and often is shared organically between users rather than by the initial distributor. Developing regulatory thresholds that capture the actual expenses incurred in implementing these types of campaigns is critical to increasing transparency, but is tricky to do well. The United Kingdom has more

\textsuperscript{187} \textit{Id.}

\textsuperscript{188} The report advised against changing the rule immediately because Brexit campaigning had already begun. \textit{Id.} at 111 n.41.
experience than the United States in using this measure, and U.S. lawmakers can benefit from that experience as they consider the appropriate reach of proposed reforms like those found in the DISCLOSE and Honest Ads Acts.

There are other insights to be gleaned as well. In the United States, the reports regularly filed by FEC-registered groups provide more timely and detailed information than those collected by the EC. The U.S. system therefore could provide useful guidance as the United Kingdom considers whether and how to change the timing and nature of its own reporting system to increase its value to voters. The United States, in turn, has much to gain by careful review of the decisions the United Kingdom has made in regard to balancing transparency with the privacy interests of smaller donors. This issue is becoming more important in the United States as disclosure rules come under increased pressure; the United Kingdom experience using a significantly higher public disclosure threshold than that found in the United States may hold valuable lessons for the U.S. reform community.

C. Source Exclusions

Unlike transparency rules, source exclusions prohibit some entities from participating in a nation’s democratic deliberations at all, even if their participation is fully disclosed. Most countries that regulate money in politics prohibit or restrict the election-related activities of foreign entities in some way, and reform proposals in this category are generally geared toward strengthening those rules. In the United States and United Kingdom, these proposals include increasing and better enforcing restrictions on political expenditures by foreign actors, non-citizens, and subsidiaries of foreign-owned corporations; requiring online platforms to verify the domestic identity of entities purchasing online campaign communications; and tightening prohibitions on foreign contributions to candidates and political parties.

Despite their ubiquity, source exclusions raise a host of problems. To start with, defining “foreign” can be surprisingly problematic, particularly in regard

---


190 See, e.g., id. ("[F]oreign nationals are prohibited from . . . [m]aking any contribution or donation of money or other thing of value, or making any expenditure, independent expenditure, or disbursement in connection with any federal, state or local election in the United States . . . ").

191 See, e.g., id. ("[F]oreign nationals are prohibited from . . . [m]aking any disbursement for an electioneering communication . . . ").

192 See, e.g., id. ("[F]oreign nationals are prohibited from . . . [m]aking any contribution or donation to any committee or organization of any national, state, district, or local political party . . . ").
to multi-generational diaspora communities and transnational corporations. It also can be challenging to operationalize source exclusions online, where information readily crosses borders and original sources are easily obscured. Finally, it can be difficult to articulate a legally compelling reason why a voter should be prevented from hearing a properly attributed message just because of the foreign status of the messenger—and not everyone agrees such restrictions are appropriate. Addressing these issues is critical to crafting and defending workable foreign source exclusion rules in both the United States and the United Kingdom.

1. Source Exclusions in the United States

Restrictions on foreign involvement in domestic elections have been part of U.S. law since at least 1938, when Congress enacted the Foreign Agents Registration Act (FARA) in response to fears that German nationals were being paid to distribute Nazi propaganda in the United States. The restrictions were tightened in the 1960s, after a Senate investigation revealed that campaign contributions had been channeled to congressional candidates by Filipino sugar industry magnates, and again in the 1990s after reports of Chinese nationals using “soft money” donations to gain access to high-level government officials in the Clinton Administration.

The distinction developed in Buckley between campaign contributions and expenditures shapes this legal landscape as well. In regard to contributions,

---

193 See Joo-Cheong Tham, Of Aliens, Money and Politics: Should Foreign Political Donations Be Banned?, 28 KING’S L.J. 262, 267–68 (2017) (describing countries’ different definitions of “foreign” actors as being undergirded by different integrity-preservation concerns).


197 Teachout, supra note 196, at 172.

candidates and political parties are categorically prohibited from soliciting, accepting, or receiving donations from foreign nationals. It is likewise illegal for a foreign national to directly or indirectly contribute money or other “things of value” to a candidate in any federal, state, or local election. FECA also prohibits knowingly assisting foreign nationals in circumventing the foreign source exclusion ban, while FARA prevents circumvention by prohibiting U.S. agents from making contributions on behalf of their foreign principals. These rules have consistently been upheld under the Buckley paradigm subjecting regulations of contributions to a relatively relaxed level of constitutional scrutiny.

Source-based expenditure bans present a more complicated legal question in the United States. The Supreme Court dealt with source-based expenditures bans most recently in *Citizens United*. As discussed above, *Citizens United* upheld the broad disclosure and disclaimer requirements imposed by BCRA on broadcast electioneering communications. But the case is better known for striking down a source-based ban on the use of corporate general revenue funds to fund the independent expenditures of corporations. Prior to *Citizens United*, U.S. law permitted corporations to use these funds to purchase “issue ads” but they were prohibited from using them to purchase express advocacy (under FECA) or electioneering communications deemed the functional equivalent of express advocacy (under BCRA and subsequent court decisions). The Court held in *Citizens United* that this source-based expenditure ban violated the First Amendment.

In doing so, the *Citizens United* Court used sweeping language to condemn source-based regulations of political speech. The government, the Court said, may not “deprive the public of the right and privilege to determine for itself what speech and speakers are worthy of consideration.” This strong rhetoric

200 Id.; 11 C.F.R. 110.20(b) (2019). BCRA also includes a sentencing enhancement for campaign finance violations involving foreign contributions specifically from or directed by foreign governments. Ciara Torres-Spelliscy, *Dark Money as a Political Sovereignty Problem*, 28 KING’S L.J. 239, 253 (2017).
202 In May 2019, the Supreme Court declined to review a Massachusetts law prohibiting such contributions. 1A Auto, Inc. v. Dir. of Office of Campaign & Political Fin., 105 N.E.3d 1175, 1181–82 (Mass. 2018), cert. denied, 139 S. Ct. 2613 (2019).
204 See 52 U.S.C. §§ 30101–30145 (Supp. IV 2017); Bipartisan Campaign Reform Act of 2002, Pub. L. No. 107-155, § 201, 116 Stat. 81; Buckley v. Valeo, 424 U.S. 1 (1976). This is a narrower definition of BCRA’s “electioneering communications” than is applicable to transparency laws. As discussed above in Part I, the Court has made clear that this narrower definition is not constitutionally compelled in regard to transparency requirements.
205 *Citizens United*, 558 U.S. at 372.
207 *Citizens United*, 558 U.S. at 341.
cast constitutional doubt on all source-based expenditure restrictions, including the foreign-source bans found in both FECA and BCRA. So it was not surprising when the federal district court for the District of Columbia was asked just a year after *Citizens United* was decided to consider whether those bans were also unconstitutional.

The case, *Bluman v. FEC*, involved two foreign nationals legally living in the United States who wanted to make express advocacy expenditures supporting candidates for federal office but were prevented from doing so by BCRA’s foreign expenditure ban. Writing for the majority, then-Judge Brett Kavanaugh upheld the law. In doing so, he brought into play a line of Supreme Court decisions, present but not emphasized in *Citizens United*, establishing the constitutional permissibility of excluding foreigners from certain tasks considered central to democratic self-government. The gist of these cases is that excluding foreign nationals from activities “intimately related” to the process of self-government is constitutionally acceptable, even when such distinctions would not be tolerated elsewhere. As the *Bluman* court noted, this doctrine has been used by the Supreme Court to uphold laws excluding foreign nationals from serving as jurors, holding elected office, and voting. The power to exclude foreign nationals from such areas, the *Bluman* court said, is a core component of a nation’s right and duty to preserve itself as a political community; to define, in other words, who is and is not a member of a given *polis*.

Drawing on this line of cases, *Bluman* held that the First Amendment did not bar the government from trying to restrict foreign influence over how voters cast their ballots, at least to the extent presented in the case before the court (which was limited to express advocacy). Elections, the court said, are “an integral aspect” of the process of self-governance, and spending money to influence voters is “at least as (and probably far more)” closely related to democratic self-government than other tasks the Supreme Court had applied the doctrine to, such as serving as a probation officer or public school teacher. Nothing in *Citizens United* precluded this result, according to the *Bluman* court.

---

208 See *Levitt*, *supra* note 206, at 222.
210 *Id.* at 292.
211 *Id.* at 287.
212 *Id.* (quoting *Bernal v. Fainter*, 467 U.S. 216, 220 (1984)).
213 *Id.*
214 *Id.* at 287–88.
215 *Bluman*, 800 F. Supp. 2d at 292.
216 *Id.* at 288–89.
and the self-governance line of cases dictated it. Therefore, limiting foreign influence over American democracy was a sufficiently compelling interest to justify the foreign source expenditures ban being challenged.

The Supreme Court summarily affirmed Bluman. But there is considerable tension between Citizens United, which is deeply sceptical of the legitimacy of source-based bans and reserves to voters the power to critically assess political information, and the rationale of Bluman, which emphasizes the power and duty of sovereign states to protect voters from foreign influence even when properly disclosed. This uncertainty caused difficulties almost immediately after Bluman was decided when the FEC was presented with a complaint alleging that a foreign national spent $327,000 opposing a California ballot referendum.

The FEC commissioners could not agree whether to act on the complaint. The disagreement turned on whether Bluman’s validation of foreign expenditure bans in a candidate election logically extended to a ban on foreign expenditures in a referendum. The legal distinction between candidate elections and referendums dated to before Citizens United, when the Supreme Court had held in First National Bank of Boston v. Bellotti that corporations could not be prohibited from using general corporate revenue funds to fund expenditures related to a referendum campaign. The Bellotti Court’s reasoning was that there was no candidate in a referendum campaign who could be “corrupted” by corporate money, and therefore no risk of quid pro quo corruption justifying the restriction.

The FEC commissioners in favor of dismissing the California complaint argued that Bellotti, not Bluman, governed because the complaint involved a ballot referendum, not a candidate election. The FEC commissioners in favor of advancing it argued that Bluman’s language regarding the importance of protecting democratic self-governance established a distinct compelling interest justifying foreign source exclusions that applied with equal force in candidate

\[218\] Bluman, 800 F. Supp. 2d at 289.
\[219\] Id. at 288.
\[223\] Id. at 2.
\[224\] Id.
\[226\] Id. at 790.
\[227\] MindGeek, MUR 6678, at 1–2; Statement of Reasons of Chair Ann M. Ravel at 2, In the Matter of MindGeek, MUR 6678 (Fed. Election Comm’n 2015); Statement of Reasons of Commissioner Ellen L. Weintraub at 1, In the Matter of MindGeek, MUR 6678 (Fed. Election Comm’n 2015).
and referendum campaigns.\textsuperscript{228} Bluman, those commissioners argued, had carved out a “no-go” zone for foreign participation in American politics that applied regardless of the nature of the underlying vote.\textsuperscript{229} Because the commissions deadlocked over this issue, the FEC took no action on the complaint.\textsuperscript{230}

The dispute over the California complaint reveals an important and unresolved issue in U.S. law regarding foreign source exclusion bans. Since Buckley, opponents of campaign finance regulation have argued that the only constitutionally acceptable justification supporting such regulations is the interest in preventing quid pro quo corruption or the appearance thereof.\textsuperscript{231} Except for a brief deviation in 1990, the Supreme Court has generally seemed to agree.\textsuperscript{232} The unresolved issue is whether that limitation also applies to foreign source bans. The commissioners’ disagreement in the California case pushes this point. If the only constitutionally acceptable reason to permit foreign source bans is to prevent the appearance or actuality of quid pro quo corruption, then such a ban could not be constitutionally applied to a referendum campaign where there is no candidate to corrupt. If, on the other hand, the Bluman rationale about protecting voters from foreign influence is a constitutionally acceptable justification distinct from concerns about quid pro quo candidate corruption, then it should apply with equal force to both candidate and referendum campaigns. As the dispute between the FEC commissioners in the California case demonstrates, the lack of clarity on this point leaves unclear both the scope and justification of U.S. foreign source exclusion laws.\textsuperscript{233}

\section*{2. Source Exclusions in the United Kingdom}

The law governing source exclusions in the United Kingdom is similar to that in the United States. Once again, the relevant primary legislation is PPERA. When PPERA was enacted, there was significant concern about foreign donations made in the 1990s to British political parties, particularly Prime Minister John Major’s Conservative Party, which had been criticized for

\textsuperscript{228} Statement of Reasons of Commissioner Ellen L. Weintraub at 3–4, In the Matter of MindGeek, MUR 6678 (Fed. Election Comm’n 2015).
\textsuperscript{229} Id. at 3.
\textsuperscript{230} MindGeek, MUR 6678 at 2.
\textsuperscript{233} L. PAIGE WHITAKER, CONG. RESEARCH SERV., R45320, CAMPAIGN FINANCE LAW: AN ANALYSIS OF KEY ISSUES, RECENT DEVELOPMENTS, AND CONSTITUTIONAL CONSIDERATIONS FOR LEGISLATION 24–25 (2018).
accepting donations tied to Serbia, Cyprus, and Russia. The Labour Party leveraged those concerns in the 1997 general election by including a foreign source ban in its election manifesto. When Labour won, the new government asked the Neill Committee to include the ban in its study of political party funding reforms. The PPERA, as noted above, was the end result of the Neill Committee’s work.

Under PPERA, contributions to political parties and registered third-party campaigners can only be accepted from “permissible donors.” Permissible donors include individuals registered on a U.K. electoral register, U.K.-registered political parties, U.K.-registered business organizations, and subsidiaries of foreign corporations registered and doing business in the United Kingdom. British citizens living abroad are permissible donors, as are non-citizen residents legally living in the United Kingdom. Political parties and registered third parties are statutorily responsible for verifying that the donations they accept are from permissible donors, and they are legally obligated to return donations that cannot be verified. Additionally, only entities who are themselves permissible donors can register as third-party campaigners. Because they cannot register as third-party campaigners, foreign entities also therefore cannot legally engage in regulated campaign spending exceeding the registration threshold.

In recommending these rules, the Neill Committee, like the court in Bluman, relied on concepts of national self-governance and the democratic process. British political parties, the Committee reasoned, are chosen by and responsible to British citizens, and their actions should not be influenced by outsiders with no “genuine stake” in the country. Therefore, only those who live, work, or

238 Id.
239 See id.
240 See id. § 56. These permissible donor rules also apply to individual candidates and registered third-party campaigners. See id. §§ 22–24.
241 Id. § 56.
242 Id.
243 See Neill Report, supra note 2, at 68.
244 Id. Worldwide, foreign donation bans are one of the most common campaign financing restrictions. See Is There a Ban on Donations from Foreign Interests to Political
do business in the United Kingdom should be “entitled to support financially the operation of the political process.” But the Neill Committee had to grapple with some sensitive issues in deciding how to implement this idea in the United Kingdom. It struggled with two questions in particular, both of which would become relevant years later in the Brexit campaign: how to define “foreign” in a country comprised of distinct nations with large diaspora populations and varying levels of devolved power, and how to prevent foreign corporations from channeling outside funds through subsidiaries registered in the United Kingdom.

Defining “foreign” in the United Kingdom context was the Neill Committee’s first challenge. If everyone who was not a U.K. citizen and resident was considered an impermissible donor, the Scottish National Party and Plaid Cymru (the Scottish and Welsh independence parties) would be disproportionately—and in their view unfairly—disadvantaged. Under such a ban, their candidates would face Labour and Conservative Party opponents who could be funded by English money while being unable to raise competing funds from their own supporters abroad. These parties historically had been supported financially by their expatriates, and they considered those expats to be no differently situated than English nationals who donated money to pro-union candidates standing for election in Scotland or Wales. The relationship of Northern Ireland to the Republic of Ireland raised additional problems, both in terms of who would be considered “foreign” for purposes of donating to political parties in Northern Ireland, and whether those donations should be publicly disclosed in a still-volatile political environment.

These concerns led the Neill Committee to recommend a nuanced approach to foreign source bans. It decided against creating general exceptions for foreign supporters of national independence parties, but partially appeased the concerns of those parties by including U.K. citizens living abroad as permissible donors. The Neill Committee went further in regard to Northern Ireland, concluding that both pragmatic concerns and the 1998 Good Friday Agreement warranted allowing Irish citizens and companies doing the business in the Republic of Ireland to be treated as not foreign for purposes of donations to Northern Irish political parties. The Committee also recommended that

---

245 NEILL REPORT, supra note 2, at 70.
246 Id. at 64, 67–68.
247 Id. at 70.
248 Id.
250 NEILL REPORT, supra note 2, at 48–49.
251 Id. at 72.
252 See id. at 76 (The Good Friday Agreement recognized the right of all people on the Irish island to identify themselves as Irish, British, or both.).
political donations in Northern Ireland be given a short-term exemption from otherwise applicable public disclosure rules. These recommendations were accepted by the government and adopted in PPERA, with the additional caveat that Northern Irish political parties were prohibited from making donations to parties and other regulated entities in the greater United Kingdom to avoid circumvention of the foreign source ban outside of Northern Ireland.

Corporate money presented a second challenge to the Neill Committee. Unlike the United States, the United Kingdom has never had a generally applicable ban on corporate contributions or expenditures. The Committee struggled with whether or not U.K.-registered subsidiaries of foreign companies should be considered permissible donors. Because such subsidiaries were presumed to be doing business in the United Kingdom, they were directly affected by U.K. law and arguably part of the U.K. political community. But because they were subsidiaries of foreign corporations, they also were possible conduits of foreign money promoting foreign interests. Partnerships comprised of international members and limited liability companies presented variations of the same problem.

The Neill Committee adopted a compromise approach here as well, recommending that U.K.-registered subsidiaries be included as permissible donors but that they be required to demonstrate that they had sufficient U.K.-based business activity to independently fund their U.K. donations. The legislation ultimately accepted the Committee’s recommendation that subsidiaries of foreign corporations must be both registered and doing business in the United Kingdom to qualify as permissible donors, but did not require such entities to demonstrate that their U.K. donations could be supported by revenue generated within the country.

Under PPERA, U.K. law bans foreign contributions but defines “foreignness” in ways designed to be sensitive to the history and circumstances

---

253 See id. at 61.
255 See id.
256 British election law does not regulate corporate contributions differently than other contributions, but British corporate law requires such contributions to be disclosed to and authorized by shareholders. Ciara Torres-Spelliscy & Kathy Fogel, Shareholder Authorized Corporate Political Spending in the United Kingdom, 46 U.S.F. L. Rev. 525, 526 (2011).
257 See Neill Report, supra note 2, at 68.
258 Id. at 74.
259 Id. at 73.
260 Id.
261 See 2016 EU Referendum, supra note 5, at 101; see also Digital Campaigning, supra note 83, ¶ 94–97.
of the United Kingdom’s constitutive parts. For enforcement purposes, the law puts the burden on parties, candidates, and registered third-party campaigners to verify that they only accept donations from permissible donors, and requires that they return funds whose origins cannot be ascertained. It also prohibits foreign entities from registering as third-party campaigners, thereby rendering it illegal for such entities to spend above the triggering threshold on any regulated campaign activities, which includes most activity intended to influence voters during an election period.

Despite this relatively nuanced approach, the issues that troubled the Neill Committee reappeared during the Brexit referendum. In fact, the largest scandal to emerge from the Brexit campaign involved fears that foreign money had been channeled to several pro-Leave campaign groups through a web of interconnected corporations. The concerns centered on U.K. businessman Arron Banks and an entity he created called Leave.EU. Leave.EU registered with the EC as a third-party campaigner and reported receiving millions of pounds of donations from Banks. It then used these funds to finance not only its own campaign, but also to make large donations to five separately registered pro-Leave groups.

When these transactions became public, concerns were raised about whether Banks, whose business ventures appeared from public records to have limited cash flow and high debt loads, had actually been the source of these funds. In response, the EC opened an investigation. The investigation revealed that Leave.EU had operated its campaign through a separate entity, Better for the Country Limited (BFTC), and that BFTC had received £6 million in funding to pay Leave.EU’s referendum expenses and an additional £2 million to use for other referendum spending in the form of donations to other registered campaign groups. BFTC reported that this entire amount—£8 million—was received from Banks in the form of donations or loans from him or his U.K.-based insurance companies. BFTC was incorporated by Banks in the United

---

263 Id. §§ 54–56.
264 Id. § 88.
266 Investigation into Payments, supra note 5.
267 Id.
268 Id.
269 Id.
270 Id.
271 Id.
272 Investigation into Payments, supra note 5.
Kingdom in May 2015, and it was not improper or illegal for Leave.EU to run its referendum campaign through BFTC.273 What the EC was concerned about was the involvement of Rock Holdings Limited, another Banks entity incorporated in the Isle of Man.274 Rock Holdings is not registered or doing business in the United Kingdom, and is not a permissible donor under U.K. law.275 So money or loans supplied by Rock Holdings could potentially violate the U.K. foreign source ban. The EC’s post-Brexit report to Parliament itemized these concerns.276 The EC reported that it believed there were reasonable grounds to suspect that Rock Holdings was a party to the donations and loans made by Banks to BFTC and used to fund Leave.EU and, through it, other pro-Leave campaigners.277 The report also stated that the EC had reason to believe that Banks knew this, that he intentionally used prohibited foreign sources to fund campaign activities, and that he concealed the true source of those funds in violation of U.K. law.278 The EC referred the matter to the National Crime Agency,279 which subsequently cleared Banks and Rock Holdings of criminal wrongdoing.280 The EC in response has recommended additional election law changes to close what it sees as a loophole in the United Kingdom’s foreign source funding ban.281

The Banks matter is not the only foreign funding scandal to emerge from Brexit. There also was concern during the referendum that the non-disclosure

273 Id.
274 Id.
275 Id. The Isle of Man is a British Overseas Territory, but is not part of the United Kingdom and does not have EU membership through its affiliation with Britain. Isle of Man entities are not permissible donors under PPERA and were not made so by the amendments regarding the Brexit referendum. The Electoral Comm’n, Donations and Loans: Guidance for Regulated Donors in Great Britain 26 (rev. Jan 2010), https://www.electoralcommission.org.uk/__data/assets/pdf_file/0019/13708/026-regulated-donors-guidance-final.pdf [https://perma.cc/99FJ-D9KT].
276 Investigation into Payments, supra note 5.
277 Id.
provision of PPERA in relation to contributions made to political parties in Northern Ireland was used to shield the source of funds used to purchase anti-
EU advertising throughout the United Kingdom.282 These concerns involved an
entity called the Constitutional Research Counsel (CRC), an unincorporated
organization based in Scotland.283 During the Brexit referendum, the CRC gave
£435,000 to the Northern Ireland Democratic Unionist Party (DUP).284 The
DUP is the dominant party in Northern Ireland’s devolved parliament, and the
only Northern Irish party to sit in the Westminster Parliament.285

The CRC’s £435,000 donation to the DUP was the largest ever made in
Northern Ireland.286 But because at the time of the Brexit campaign PPERA
prohibited public disclosure of political donations in Northern Ireland, the
underlying funding sources of CRC itself were not publicly disclosed.287
Disclosure reports made by the DUP to the EC showed that much of this money
was used to fund anti-EU messaging throughout the United Kingdom, including
a full-page newspaper ad in London and targeted Facebook ads arranged by
Aggregate IQ (a Canadian data firm linked to Cambridge Analytica and also
implicated in the allegedly improper use of Facebook user data).288

Restrictions on public disclosure of Northern Ireland political party funding
have been lifted since the Brexit vote, but the EC remains prohibited by law
from providing information about donations that were legally confidential at the
time they were made.289 The EC’s Chief Commissioner told a Parliamentary
committee that the EC had verified that the donors listed on the DUP election
reports were permissible, but was prohibited by law from elaborating further.290
This lack of public disclosure about the CRC’s underlying funding has spurred
continuing suspicions about the propriety of these transactions.291

282 DCMS REPORT, supra note 5, at 64–65.
283 Id. at 64.
284 Id. at 64–65.
285 Sinn Féin, the Irish republican political party, successfully elects Westminster MPs
from constituency districts in Northern Ireland, but the party’s MPs abstain from taking their
seats in protest over Britain’s continuing claims of sovereignty over Northern Ireland. Paul
Maskey, Editorial, I’m a Sinn Féin MP. This Is Why I Won’t Go to Westminster, Even over
Brexit, GUARDIAN (Mar. 6, 2018), https://www.theguardian.com/commentisfree/2018/
mar/06/sinn-fein-mp-british-parliament-irish-republicans-brexit [https://perma.cc/6D5V-
TT6T].
286 DCMS REPORT, supra note 5, at 64–65.
287 See id.
288 Id. at 45–48; see also INFO. COMM’R’S OFFICE, supra note 5, at 49–50.
289 DCMS REPORT, supra note 5, at 65.
290 Conclusion of Assessments into Allegations Regarding Certain EU Referendum
ission.org.uk/i-am-a/journalist/electoral-commission-media-centre/referendums-to-keep/
conclusion-of-assessments-into-allegations-regarding-certain-eu-referendum-campaigners
[https://perma.cc/39Q5-52HM] [hereinafter Assessment Conclusions].
291 See id.
The Brexit campaign revealed additional gaps in the United Kingdom’s foreign source exclusion laws. As noted above, PPERA requires any third-party campaigner intending to spend over the threshold amount on regulated activities to register with the EC. supra note 5. Impermissible donors, including foreign entities, cannot legally register and therefore cannot legally engage in spending over the threshold amount. supra note 5. But it appears there is no provision in U.K. law preventing foreign actors from spending under the registration threshold. Consequently, in the Brexit referendum, foreign actors could spend up to £10,000 on election activity without running afoul of the law. supra note 5. A similar issue exists on the donation side: foreign entities are not permissible donors, but “donations” are defined as contributions of £500 or more. supra note 5. Since contributions under £500 are not “donations,” they also are not governed by the permissible donor requirements. supra note 5.

These gaps mean that much Brexit-related content promoted on social media by foreign actors did not violate U.K. law. While foreign-funded Facebook activity seems to have played less of a role in the Brexit referendum than in the United States presidential election, reviews conducted after the referendum indicate that Russian entities did purchase some Brexit-related advertisements, and data released by Twitter shows accounts affiliated with the St. Petersburg operation promoted Brexit tweets. supra note 5. But as long as the cost of these communications was under the third-party campaigner registration threshold of £10,000, these expenditures would have been legally permissible.

supra note 5.

supra note 5.

supra note 5.

supra note 5.

supra note 5.

supra note 5.

supra note 5.

supra note 5. More recently, anti-EU political parties in the United Kingdom have been accused of taking advantage of this gap on the contribution side, by accepting foreign sourced donations under the £500 threshold. Rajeev Syal, Brexit Party at High Risk of Accepting Illegal Funds, Says Watchdog, GUARDIAN (June 12, 2019), https://www.theguardian.com/politics/2019/jun/
Another problem in the U.K. regulatory scheme became apparent only after the Brexit referendum, as Parliament struggled to approve a plan operationalizing the United Kingdom’s exit from the European Union. Because PPERA’s third-party campaigner expenditure rules only apply during the regulated period leading up to an election and only cover communications intended to influence voters, paid advertisements placed outside that period purportedly targeting elected officials are effectively unregulated. What this meant during the Brexit negotiations was that in early 2019, as the U.K. Government tried to obtain approval of its exit plan, hundreds of thousands of pounds worth of social media ads placed on Facebook and other platforms pushing for a “no-deal” exit were not covered by U.K. disclosure or disclaimer laws. Transparency measures voluntarily adopted by Facebook show that a group called “Britain’s Future” purchased the ads, but because the ads were placed outside of a regulated election period and purportedly targeted ministers rather than voters (by telling them to “vote no” on the deal), the British people were completely in the dark about who was funding the online campaign.

The United Kingdom, like the United States, is considering various ways to address these issues. The prohibition on public disclosure of political contributions in Northern Ireland has already been lifted, and the EC has asked the Government to allow it to retroactively disclose what it knows about Brexit referendum funding in Northern Ireland. The U.K. Information Commissioner, as well as the EC, has recommended that the Banks/Rock Holdings situation be addressed by adopting the Neill Committee’s original proposal of requiring U.K.-based subsidiaries of foreign corporations to demonstrate that their U.K. businesses generate sufficient revenue to cover their political spending. There also are proposals to improve the ability of political parties, candidates, and third-party campaigners to verify the source of the funds they receive, to tighten rules regarding the flow of money between Northern Ireland and the rest of the United Kingdom, and to close the gaps created by the

---


301 Rajan, supra note 194.

302 Assessment Conclusions, supra note 290.

303 DIGITAL CAMPAIGNING, supra note 83, ¶¶ 93–97.

wording of the third-party campaigner registration threshold and foreign donation ban. The EC also has recommended studying whether and how the regulated period should be adjusted, to better capture the nature of today’s ongoing online influence campaigns.

3. Foreign Source Exclusion Recap

Once again, there is much in these reform proposals warranting comparative consideration, and very little that is barred by current U.S. constitutional law. Legislators in the United States are struggling with the same question as their U.K. counterparts regarding how to define foreignness in regard to subsidiaries of foreign held corporations.305 The Neill Committee’s careful and comprehensive consideration of the underlying justifications for foreign source bans, and its nuanced application of that justification when defining the relevant political communities in Northern Ireland, Scotland, and Wales, surely will be of value to U.S. advocates as cases like Bluman wind their way to the Supreme Court. The Northern Ireland experience also offers a useful comparison as U.S. courts work to define the appropriate scope of the abuse and harassment exceptions to U.S. disclosure laws. Finally, the FEC, which has struggled with the broader question of how to justify a foreign source ban in referendums (when there is no candidate to “corrupt”) would benefit from the rich discussion of this issue engaged in by the Neill Committee and continuing today as the United Kingdom modifies its rules for non-candidate referendums and perpetual campaigns.306

The United Kingdom, in turn, could learn from the extensive experience the United States has had with a regulatory system not restricted to a defined pre-election period.307 The U.S. system has grappled for decades with the problems associated with balancing this type of ongoing regulation with the need to maintain both donor privacy and space for robust political speech.308 Lawmakers in the United States have done this in the shadow of the Supreme Court’s First Amendment jurisprudence, but the same normative challenges exist regardless of that judicial overlay.309 The United Kingdom could benefit from this experience when contemplating whether and how to extend its own regulatory regime.

305 See DISCLOSE Act of 2017, S. 1585, 115th Cong. § 101(a)(3) (2017) (using an unwieldy definition to articulate foreignness). The FEC determined that domestic subsidiaries of foreign corporations can engage in some election-related activities but cannot do so using funds from foreign nationals. Martin, supra note 189. Foreign nationals also may not participate in the decision-making process regarding any such activities. Id.
306 See supra Part IV.B.3.
307 See supra Part III.A.
308 Id.
309 Id.
D. Content Exclusions

Unlike source exclusions, content exclusions restrict communications on the basis of what they say.\(^{310}\) They are the most controversial type of regulation, and invariably generate accusations of censorship, particularly in the United States.\(^{311}\) This also is the class in the taxonomy in which U.S. and U.K. laws are the most distinct. But it oversimplifies both U.S. and U.K. law to believe that the two systems are so different that there is no fruitful ground, even here, for comparative study. As shown below, despite its frequent insistence to the contrary, U.S. First Amendment law tolerates content-based regulation in many contexts, and recent content-based reform proposals in the United Kingdom are far less far-reaching than they can sound (to American ears) on first impression.

1. Content Exclusions in the United States

It is black letter law that the First Amendment requires strict scrutiny of regulations distinguishing speech on the basis of its content.\(^{312}\) Little else is clear about this area, however, including the basic question of how to determine whether a law is or is not content-based.\(^{313}\) There also are numerous exceptions to the strict scrutiny rule, and sometimes even content-based regulations subjected to strict scrutiny survive and are upheld as constitutional.\(^{314}\)

These legal nuances take several forms. First, there are several content-based types of speech, such as perjury and fraud, that are considered categorically outside the protection of the First Amendment.\(^{315}\) Regulation and even prohibitions of these categories of speech are constitutionally acceptable as long as they are viewpoint neutral.\(^{316}\) Second, even constitutionally protected speech can be regulated in some situations if the regulation advances a substantial interest unrelated to the suppression of speech, such as prohibiting the use of profanity in governmental buildings or prohibiting harassment in public workplaces.\(^{317}\) Third, speech can be regulated on the basis of its content if it presents a clear, serious, and imminent danger to people or property—the classic example of which is falsely crying “fire” in a crowded theater.\(^{318}\)


\(^{311}\) Id. at 2226 (“Content-based laws—those that target speech based on its communicative content—are presumptively un-constitutional . . .”).

\(^{312}\) Id.

\(^{313}\) See, e.g., id. at 2227 (“Some facial distinctions based on a message are obvious, defining regulated speech by particular subject matter, and others are more subtle, defining regulated speech by its function or purpose. Both are distinctions drawn based on the message a speaker conveys, and, therefore, are subject to strict scrutiny.”).


\(^{316}\) See id.


\(^{318}\) Schenck v. United States, 249 U.S. 47, 52 (1919).
Additionally, in exceptional cases, even viewpoint-based regulations can survive strict scrutiny review and be upheld as constitutional, as long as the regulation is sufficiently narrowly drawn to advance a compelling interest, such as criminalizing “true threats” against the life of the President.319

Each of these doctrines is complex and contested; the important point for current purposes is to illustrate that despite its frequently absolutist rhetoric, U.S. constitutional law permits content-based restrictions for a variety of reasons under a variety of doctrines. These doctrines have been applied even in the realm of political communications.320 The clearest example of this is the long-standing acceptance of state laws banning campaigning in or near polling stations.321 All fifty U.S. states have such laws, many of which have been in place for more than 100 years.322 Although its reasoning has varied, the Supreme Court has consistently upheld these laws as long as they are reasonable, viewpoint neutral, and not overly broad.323

Efforts to regulate campaign speech on the basis that it is misleading or untruthful have had more mixed success. The Supreme Court’s most recent decision in this area, United States v. Alvarez, involved the Stolen Valor Act, a federal statute that penalized falsely claiming to hold certain types of military honors.324 In a split decision, a plurality of justices struck down the law.325 In doing so, they refused to accept the government’s argument that false speech, like perjury and other existing content-based exclusions, is categorically unprotected by the First Amendment.326 The plurality also refused to apply a reduced standard of review just because the speech being regulated was verifiably false.327 Instead, the plurality held that laws punishing false statements will be subject to strict scrutiny unless there is an additional “legally cognizable harm” associated with the false statement justifying less rigorous review.328 The plurality opinion gave several examples of laws that fulfill this criteria, including perjury laws (which protect the integrity of the courts), laws prohibiting misrepresenting yourself as a representative of the government (which protect the integrity of governmental processes), and anti-fraud laws (which protect against consumer and financial harms).329

Alvarez is a mixed bag for content-based election law reform proposals. The plurality was plainly skeptical of applying anything less than strict judicial

321 See id.
322 Id.
325 Id. at 730.
326 Id. at 719–20.
327 Id. at 716.
328 Id. at 719.
329 Id. at 720.
scrutiny to new types of content-based regulations of speech. The Justices—including the dissenting Justices—expressed particular concern that content-based rules in the political realm risk chilling protected political speech. Nonetheless, the plurality’s recognition that preventing fraud, misrepresentation, and the integrity of governmental processes are cognizable harms justifying less rigorous judicial scrutiny may leave open a path for a carefully crafted false statement rule even in the electoral realm.

An earlier Supreme Court case, Brown v. Hartlage, shows one possible approach. In Hartlage, the Court considered a Kentucky law that punished a candidate (Brown) for promising to take a smaller salary than entitled to if elected. Brown apparently made the statement without realizing that state law prohibited him from taking a salary reduction. Brown won the election, and the losing candidate claimed Brown’s statement violated the state Corrupt Practices Act, which among other things penalized candidates for state office from making false or misleading statements. The state courts sided with Brown’s opponent, and nullified the election. Brown appealed to the U.S. Supreme Court, arguing that applying the Corrupt Practices Act to his situation violated the First Amendment. The Supreme Court agreed with Kentucky that protecting the integrity of elections was a compelling interest and that “demonstrable falsehoods,” even in the context of political campaigns, do not have the same protection as truthful statements. Nonetheless, the Court invalidated the application of the law to Brown’s statement because there had been no showing that it had been made in bad faith, with knowledge that it was false, or with “reckless disregard” of its truth.

This language borrows heavily from U.S. defamation law. Defamation law has had a constitutional component in the United States since 1965 when the Supreme Court held in New York Times v. Sullivan that defamation law could not be used to punish minor, unintentional untruths about public officials. Instead, the Sullivan Court said, the First Amendment protects even defamatory statements about public officials as long as the statements are not made with “actual malice.” As echoed by the Court in Hartlage, “actual malice” is

---

331 Id. at 750–51 (Alito, J., dissenting).
332 Id. at 720.
334 Id.
335 Id. at 48.
336 Id. at 49. The argument was that the statement was misleading because the salary was set by law and the winning candidate could not change it. Id. at 50.
337 Id. at 50–51.
338 Id. at 52.
339 Brown, 456 U.S. at 52, 60.
340 Id. at 61–62.
342 Id. at 279–80.
343 See Brown, 456 U.S. at 61.
defined as making false statements of material fact with either actual knowledge or reckless disregard of the statement’s truth or falsity. Speakers are thereby constitutionally protected from liability for unintentional or even negligent errors, minor untruths, and non-factual statements such as satire, opinion, and hyperbole. This doctrine does not protect intentional lies. Indeed, it does just the opposite: an intentional lie about a verifiable material fact is precisely what can be punished under Sullivan.

States have relied on this distinction when defending state laws regulating untrue statements in election campaigns. The federal appellate courts, however, appear—at least at first blush—to be unconvinced. The Eighth and Sixth Circuit courts of appeals have struck down prosecutions under false speech laws in Minnesota and Ohio, respectively. In each case, state courts had interpreted the state statutes as applying only to speech meeting the Sullivan standard, but the federal courts nonetheless held that the First Amendment barred the state prosecution.

But these decisions are less determinative than they appear, for two reasons. First, in each case, the underlying statement involved was not obviously verifiable as true or false. In Ohio, the underlying statement accused a member of Congress of voting for “taxpayer-funded abortion” by supporting the Affordable Care Act—a complex statute that expanded reproductive health care generally but did not specifically appropriate federal funds for abortion care. The Minnesota case was more complicated, but the court seemed skeptical that the underlying statements giving rise to the complaint had been demonstrably false. Neither case, therefore, presented a crisp opportunity for a court to evaluate the core question of whether a narrowly drawn statute applied only to false statements of verifiable fact would pass First Amendment scrutiny.

Second, even while striking down the statutes in front of them the courts in both cases affirmed that states are not powerless to regulate all false or

---

346 See id. at 283.
347 Id.
348 281 Care Comm. v. Arneson, 638 F.3d 621, 636 (8th Cir. 2011).
349 Susan B. Anthony List v. Driehaus, 814 F.3d 466, 476 (6th Cir. 2016).
350 See id. at 472–76; 281 Care Comm., 638 F.3d at 636. In the United States, a state supreme court has authority to determine the meaning of a state law, but federal courts determine whether the statute so interpreted violates federal statutory or constitutional law. See, e.g., N.Y. Times v. Sullivan, 376 U.S. 254, 264–66 (1964) (finding Alabama defamation law unconstitutional); see also State Courts vs. Federal Courts, JUDICIAL LEARNING CTR., https://judiciallearningcenter.org/state-courts-vs-federal-courts/ [https://perma.cc/4GPG-NQUG].
351 Susan B. Anthony List, 814 F.3d at 470.
352 The Eighth Circuit opinion put the word “false” in quotations when discussing the statements underlying the earlier proceeding. See 281 Care Comm., 638 F.3d at 626.
misleading statements in the political realm. The Eighth Circuit unequivocally rejected any such interpretation of its decision, stating “[w]e do not, of course, hold today that a state may never regulate false speech in this context.”

The Sixth Circuit likewise emphasized that the plaintiff was not asserting (and it was not affirming) a constitutional “right to lie” in election campaigns. Instead, both courts held only that false campaign statements are not categorically unprotected by the First Amendment—a relatively unremarkable holding, echoed by the Supreme Court in Alvarez. Additionally, both courts readily acknowledged that preserving the integrity of democratic elections is a compelling governmental interest, even though the states had failed to sufficiently narrow the reach of the statutes in the cases presented.

These cases leave open the possibility that a carefully drawn and applied regulation of false or misleading campaign speech could be upheld, even if subjected to heightened review. Which is not to say it would be: laws restricting political speech rarely survive judicial review in the United States, and a deep skepticism of any governmental process adjudicating the truth or falsity of political statements undergirds all of these opinions. But a carefully crafted law advancing a well-defined interest in protecting the integrity of elections could address some of the worst abuses in the U.S. system.

There are examples of what such a law might look like. In striking down the Ohio law, the Sixth Circuit provided a detailed critique of why the Ohio statute was overly broad, and what a more narrowly tailored statute would look like. Election law scholar Richard Hasen has argued that a statute regulating only intentionally false statements involving easily verifiable factual statements (such as intentionally misstating the date of the election) could be sufficiently narrowly tailored to advance the compelling interest in protecting the right to vote. Targeting actual “fake news”—completely made up stories designed for commercial rather than political purposes—is another possible approach, as is

---

353 Id. at 636.
354 See Susan B. Anthony List, 814 F.3d at 473.
355 Id. at 473; see 281 Care Comm., 638 F.3d at 636; United States v. Alvarez, 567 U.S. 709, 722 (2012).
356 281 Care Comm., 638 F.3d at 636; Susan B. Anthony List, 814 F.3d at 473–74. The Sixth Circuit in particular emphasized the importance of “protecting ‘voters from confusion and undue influence,’ and ‘ensuring that an individual’s right to vote is not undermined by fraud.’” Id. at 473 (quoting Burson v. Freeman, 504 U.S 191, 199 (1992)).
357 Susan B. Anthony List, 814 F.3d at 474. Specifically, the court found the Ohio law was insufficiently narrowly tailored in four ways: (1) it could take up to six months to resolve a complaint, which undercut arguments that the law protected election integrity; (2) anyone could initiate a complaint, and there was no screening for frivolous complaints, meaning the process could be weaponized by competing campaigns; (3) the law applied to non-material statements; (4) the law applied to “commercial intermediaries” such as billboard companies who were messengers rather than speakers. Id. at 474–76.
limiting a false statement law to micro-targeted online advertising not seen beyond its target audience and therefore less amenable to more narrowly tailored counter-speech remedies.\textsuperscript{359} Designing and enacting any such rule would be challenging, but nothing in current U.S. law takes it completely off the constitutional table.

2. Content Exclusions in the United Kingdom

Content-based exclusions do not face the same judicial scrutiny in the United Kingdom as in the United States, but even in the United Kingdom such laws do face some judicial skepticism. Courts in the United Kingdom have imposed reasonableness and proportionality restraints on speech regulations enforcing primary legislation, and Article 10 of the European Convention on Human Rights creates a judicially protected freedom of expression right applicable in the United Kingdom through the Human Rights Act 1998.\textsuperscript{360} More generally, U.K. courts, like their U.S. counterparts, recognize the special importance in a democracy of free and robust discussion of political issues, and are sensitive to that when evaluating content-based regulations of political speech.\textsuperscript{361}

Nonetheless, it is plainly true that U.K. lawmakers have more leeway to regulate false or misleading campaign speech than do their U.S. counterparts. The RPA prohibits the use of fraud or “undue influence” in political campaigns, and U.K. courts have permitted prosecution under those provisions for things like distributing an election flyer falsely claiming to have been created by a rival party, and making false statements that a candidate has withdrawn from an election.\textsuperscript{362} But the most extensive judicial treatment of content-based restrictions on political speech in the United Kingdom has come under

\textsuperscript{359} See Susan B. Anthony List, 814 F.3d at 474 (holding that laws may not “pass constitutional muster because they are not narrowly tailored in their (1) timing, (2) lack of a screening process for frivolous complaints, (3) application to non-material statements, (4) application to commercial intermediaries, and (5) over-inclusiveness and under-inclusiveness”).

\textsuperscript{360} Human Rights Act 1998, c. 42, § 1.3, sch. 1 (UK), http://www.legislation.gov.uk/ukpga/1998/42/contents [https://perma.cc/Q33Z-NQ5R]. Under the HRA, a declaration of incompatibility does not invalidate primary legislation, nor does such a declaration affect the parties’ obligation to comply with it. Id. § 4. Instead, it triggers a review process through which the Government may amend the legislation. Id. The declaration itself entails no legal obligation on the Government or Parliament to make any such amendments. Id.


\textsuperscript{362} See generally PROTECTING THE DEBATE, supra note 5 (recommending sanctions for the electoral offence of intimidation); see also R v. Rowe ex parte Mainwaring [1992] 1 WLR 1059, 1059 n.1.
Section 106 of the RPA. Section 106 prohibits any person from attempting to influence an election by making or publishing “false statement[s] of fact” in relation to a candidate’s “personal” character. Candidates who violate this provision are subject to civil fines and can be barred from standing for elective office. Courts also have the authority to set aside elections tainted by violations of the statute.

Prosecutions under Section 106 have been few, but they have been sustained by U.K. courts. The decision in Watkins v. Woolas is illustrative. In Woolas, the Parliamentary Election Court considered whether to set aside an election result on the grounds that the winning candidate violated Section 106. The underlying allegation involved three statements made by the winning candidate (Woolas) about his opponent (Watkins): that Watkins had failed to condemn the actions of violent Muslim extremists, had actively solicited the support of such extremists, and had reneged on a promise to live in the constituency district in which he was standing.

The court determined that all three statements, in context, were intentionally false statements of fact. It further found that Woolas had no reasonable grounds for believing they were true and that he did not in fact believe them to be so. The court then found the first two statements were made with intentional dishonesty. Finally, the court held that the first two statements were about Watkins’s personal character (which are covered by Section 106), rather than his political behavior or opinions (which are not). The court

---

364 Id.
367 Rowbottom, supra note 365, at 509 n.10.
368 See Watkins v. Woolas [2010] EWHC (QB) [208].
369 Id. at [1]–[6]. Woolas was heard by the U.K. Parliamentary Election Court, a court constituted under U.K. law for the purpose of hearing certain election disputes. Woolas v. Parliamentary Election Court, [2010] EWHC (Admin) 3169, [2011] 2 WLR 1362, 1363–64 (Eng.). Decisions of the court are appealable only as to errors of law. Id.
371 Id. at [207]. The statements were made in leaflets and said that “extremist Muslim activists” had targeted Woolas with violence. Id. at [64]. They included photographs of Watkins next to photographs of demonstrators holding placards calling for the beheading of people who insult Islam. Id. A second leaflet stated that “one extremist website” had created a competition for the most “imaginative ways to kill [Phil Woolas]” and that Watkins had not condemned the website or the group that created it. Id. at [96]–[97].
372 Id. at [207].
373 Id. at [195], [203].
374 Id. at [82], [94]–[95].
therefore determined that the first two statements had violated Section 106 and that the election should be set aside.\textsuperscript{375}

Woolas appealed, arguing that applying Section 106 to his statements violated the freedom of expression protected by Article 10 of the European Convention on Human Rights.\textsuperscript{376} The U.K. appellate court disagreed.\textsuperscript{377} Citing an earlier case (\textit{Bowman v. United Kingdom}\textsuperscript{378}), the court agreed with Woolas that free elections and freedom of expression are the “bedrock” of democracy, and that Article 10 required judges to carefully review laws restricting a candidate’s speech during an election period.\textsuperscript{379} But, the court went on, free and fair elections also require protecting the right of the electorate to make its choices based on the candidates’ competing positions and policy arguments, not on false statements about a candidate’s personal character.\textsuperscript{380} Section 106, with its limited focus on exactly that, was therefore a proportionate limitation on the freedom of expression protected by Article 10.\textsuperscript{381}

The U.K. courts’ more permissive approach to content-based regulations of political speech has been tested elsewhere as well. \textit{R (ProLife Alliance) v. British Broadcasting Corporation} involved the BBC’s refusal to show a broadcast submitted by the Pro Life Alliance, a U.K. political party.\textsuperscript{382} As a political party registered under PPERA, the ProLife Alliance was entitled to make a broadcast on the BBC during the election period.\textsuperscript{383} But the ad it submitted included what the court described as “prolonged and graphic images of . . . mangled and mutilated” aborted fetuses.\textsuperscript{384} The BBC refused to broadcast the ad, citing its statutory power and duty to maintain standards of taste, decency, and non-offensiveness in its programming.\textsuperscript{385}

The ProLife Alliance sued, arguing that the BBC’s refusal to broadcast the ad because of its content violated Article 10.\textsuperscript{386} The case reached the U.K. Supreme Court, which upheld the BBC’s decision.\textsuperscript{387} Lord Nicholls’ lead opinion acknowledged that Article 10 required that the content-based restrictions imposed by the BBC be justified, particularly when used to censor

\textsuperscript{375} \textit{Id.} at [207]–[208].
\textsuperscript{377} \textit{Id.} at 1363–64.
\textsuperscript{379} Woolas [2010] EWHC 3169 [89].
\textsuperscript{380} \textit{Id.} at 1391.
\textsuperscript{381} \textit{Id.} at [91].
\textsuperscript{382} \textit{R (ProLife Alliance) v. British Broadcasting Corporation} [2003] UKHL 23, [2], [2004] 1 AC 185 (appeal taken from Eng.).
\textsuperscript{383} \textit{Id.} at [4]–[9].
\textsuperscript{384} \textit{Id.} at [3].
\textsuperscript{386} \textit{ProLife Alliance} [2003] UKHL 23, [2].
\textsuperscript{387} \textit{Id.}
the broadcast of a political party during an election campaign. But, Lord Nicholls went on, nothing in U.K. law or Article 10 jurisprudence entitled political parties to an exemption from a generally applicable prohibition on the broadcast of offensive material. A concurring opinion by Lord Hoffman elaborated on this point, noting that the primary right protected by Article 10 in the context of a generally applicable prohibition is a right to not be denied access to the airways on “discriminatory, arbitrary or unreasonable” grounds. Because the BBC’s decision was based in preexisting and non-discriminatory standards, it was lawful under the ECHR’s Article 10 jurisprudence.

The upshot of these cases is that while U.K. courts, like their U.S. counterparts, recognize the importance in a democracy of robust discussion of political issues, and consequently engage in more searching review of content-based restrictions involving political speech, they also are more accepting of restrictions designed to protect democratic discourse from false, misleading, or offensive campaigning.

Reform proposals in the United Kingdom calling for content-based prohibitions on false or harmful election communications draw on this leniency. Many of these proposals came together in an April 2019 report, the “Online Harms White Paper,” presented to Parliament by the Home Department working in conjunction with the Digital, Culture, Media & Sport Committee. The report identifies four kinds of online harms: online harassment and bullying, terroristic propaganda and recruitment, political disinformation, and gang and criminal glorification. The core proposal in the report is that Parliament should impose a statutory duty of care on large social media and digital data companies requiring them to adopt risk-based and proportionate policies addressing the most egregious online harms immediately, while gradually developing better practices in regard to lesser harms. The law would not require companies to engage in any particular specified acts. Instead, the legal obligation imposed would require companies to create, follow, and enforce

---

388 Id. at [6]–[8].
389 Id. at [10].
390 Id. at [62].
391 Id. at 185. In a dissenting speech, Lord Scott disagreed. Id. at [83]–[99]. Lord Scott would have required the BBC directors to conduct a proportionality review before rejecting the broadcast, and argued that the interest of preserving public decency and avoiding offense was insufficient to justify restricting the speech of a political party and therefore incompatible with Article 10. Id.
394 Id. at 5.
395 Id. at 7–8.
396 See id. at 7.
practices sufficient to meet the statutory duty of care. An independent regulator would oversee compliance, generate public reports regarding company practices, and impose fines and liability on non-compliant companies.

Despite its rather ambitious scope, many of the recommendations in the report are relatively benign. In several areas, the report’s most important contribution is to highlight existing regulatory gaps allowing activity to go unpunished online even when it would be regulated or prohibited under current law if occurring offline. For example, as discussed above, the RPA already prohibits fraud and undue influence in elections, and penalizes candidates for making false statements of fact about the character of their opponents. Campaign statements in the United Kingdom also are subject to criminal and civil laws restricting copyright, libel, contempt and obscenity, and similar existing prohibitions against the incitement of racial or religious hatred. Yet as the “Online Harms” report highlights, these generally applicable rules are rarely enforced online. Several of the suggestions in the report aim merely to close that gap.

In regard to digital disinformation, the report goes further and suggests that the rapid spread of online disinformation designed to mislead voters or undermine democratic processes may call for new regulatory controls. It grounds this observation in the “unprecedented effectiveness” of online actors to use false information online to manipulate public opinion through automation, anonymity, and fraud. Even here, though, the report notes that a precedent already exists in British law for regulating covert efforts to manipulate public opinion, and proposes that those same principles guide any new approach to online regulation.

The “Online Harms” report concludes by making several relatively modest recommendations about how social media companies could meet any statutory duty of care imposed by future legislation. It suggests they adopt terms of service prohibiting users from misrepresenting their identity on social media for the purpose of disseminating or amplifying disinformation, use automated AI techniques to find and remove fake news items, develop systems to evaluate the trustworthiness of content providers, make more reliable content more visible to users, and require users to be notified when they are interacting with

---

397 Id.
398 Id. at 59.
399 See HM Gov’t, supra note 393, at 5–10.
400 See id.
402 See id.
403 HM Gov’t, supra note 393, at 24.
404 Id.
405 Id. The report points to the obligation of television regulators under the Broadcasting Act 1990 to ensure broadcasts do not include subliminal messages designed to influence individuals without them being aware. Id.
406 Id. at 70–71.
automated accounts or looking at paid content. None of these suggestions are legally extraordinary, and all deserve at least some comparative consideration.

3. Content Exclusions Recap

The law regulating content-based exclusions is meaningfully different in the United States and United Kingdom, and these differences are highly relevant to the reform proposals under consideration in each country. Regulators in the United Kingdom have much more flexibility than their U.S. counterparts in devising content-based strategies to deal with online electioneering and foreign interference. But this does not mean that there is no value in comparative study, even here. Courts in the United States have done a great deal of thinking about the value and risks of content-based regulation of political speech, and U.K. regulators would surely benefit from careful consideration of this jurisprudence. The U.S. experience here is extensive, and valuable even to those who disagree with the ultimate conclusions reached.

At least some recommendations in the “Online Harms” report also should have value to U.S. reformers, especially those that resonate within U.S. consumer protection, fraud, and defamation law. Paying attention to how these suggestions are operationalized in the United Kingdom may stimulate new thinking about how they could be implemented within existing U.S. law. Even recommendations unlikely to be constitutional in the United States if imposed on social media companies by law may nonetheless provide case studies for social media companies looking for ways to better manage their online space. Facebook, for example, already is experimenting with providing information about the reliability of content providers, and Twitter has taken some steps toward eliminating imposter accounts and bot-nets. Thus, while the legal environment around content-based regulations is very different in the United States and United Kingdom, the two countries nonetheless have things to learn from each even here.

V. Conclusion

The 2016 elections in the United States and United Kingdom revealed how challenging it is in both nations to effectively regulate online and foreign election interference under current law. Despite the urgent need to remedy this, widespread assumptions about America’s “First Amendment Exceptionalism”

---

407 Id.
have left regulators, lawmakers, and reform advocates in both countries relatively uninterested in each other’s efforts. This Article hopes to change that. Evaluating the reform proposals being considered in each nation through a common taxonomy enables us to see more clearly where differences between the legal systems of the two nations do and do not pose system-specific challenges to different types of reform proposals, and to direct future comparative efforts accordingly.

As the above analysis shows, there are several areas where further comparative study would be valuable. In regard to public education, few legal barriers hinder reforms in either country, and advocates in the United States and United Kingdom should freely share and learn from each other’s efforts. In regard to increased transparency, the United States can learn from the United Kingdom’s experience with lower disclosure thresholds for individual donors and its wider experience with pay-to-produce rather than pay-to-place thresholds for transparency rules. The United States also could benefit from Scotland’s experience in actually implementing a comprehensive online imprint requirement. The United Kingdom, in turn, can learn from the decades of experience the United States has in trying to fairly and effectively distinguish regulated campaigning from unregulated political speech. The United Kingdom also may benefit from studying the successes and failures of the more detailed and ongoing reporting requirements common in the U.S. system.

Similar opportunities for constructive comparison exist in regard to foreign source and content exclusions. Both nations are struggling with the definition of “foreign” when defining a political community. There will be much to be learned here as each country continues exploring the parameters of and justifications for their respective foreign source bans. The United States also could learn from the United Kingdom’s longer history of regulating corporate campaign spending, including the spending of domestic subsidiaries of foreign-held entities. Even in regard to content exclusions, where the basic law of the two nations is the most different, there are areas where comparative study would be valuable. Courts in the United Kingdom have done quite a bit of thinking about how false statements of verifiable facts interfere with the right of voters to make informed decisions based on accurate information. The United States could learn from this when considering whether and how to craft narrow laws penalizing egregious misrepresentation and fraud in the electoral realm. The United Kingdom, in turn, may benefit from the extensive judicial and scholarly literature in the United States documenting the risks of being too eager to adopt broad prohibitions in this area.

The United States and the United Kingdom share a long tradition of insightful comparative scholarship. While there are real and meaningful differences between the legal regulation of campaign communications in the United States and the United Kingdom, those differences present few barriers to many of the most significant types of reform proposals being considered in each country, so that tradition can and should continue through additional comparative work in this critical area.
Two recent changes to U.S. federal law threaten the viability of colleges and universities. President Trump’s signing of the Tax Cuts and Jobs Act (TCJA) into law at the close of 2017 signified a continued trend of decreasing funds previously available to higher education. National and state funding cuts are resulting in a cost-value educational crisis in the United States, with tuitions increasing, students needing to borrow more, and academic programs and faculty lines being cut. In addition, college athletic departments assert that the United States Supreme Court’s holding in Murphy v. National College Athletic Ass’n, which declared that the Professional and Amateur Sports Protection Act (PASPA) unconstitutionally commandeered state legislatures into maintaining laws prohibiting sports wagering, could jeopardize the integrity of college sports. Only by examining the foundational relationships between gambling, taxation, and higher education in the United States is it evident that Murphy could actually be the catalyst for generating revenue back into colleges and universities rather than the apocalyptic threat some have predicted.

Following Murphy, more than thirty states have introduced bills legalizing sports gambling. Of those, only two have legislatively earmarked a portion of the projected revenue back into higher education specifically. States’ de minimis interest in redirecting some of this revenue stream to colleges and universities is surprising for two reasons. First, public institutions directly finance the very sports that are helping to drive the newly legalized sports betting market. In addition, there exists a strong historical affiliation between gambling, tax revenue, and education in the United States. Consequently, college administrators should embrace sports wagering as a means of increasing the integrity of college sports and capitalize on this opportunity to recoup revenue lost due to the TCJA and states’ continued reallocation of funding away from higher education.

The intersection of sports gambling, taxation, and higher education is an undertheorized area of law. This Article highlights the historic...
connections that bind them and proposes a mechanism that public institutions and state legislators can adopt to better monitor sports-related integrity concerns while supplementing college and university budgets. This Article is the first to harmonize the relationship between education and gambling post-Murphy; the first to introduce various state tax frameworks surrounding legalized sports gambling in the United States; and the first to introduce a fee model to help redirect funding from legalized sports wagering back into higher education.

TABLE OF CONTENTS

I. INTRODUCTION ..........................................................467

II. MURPHY v. NCAA—OPENING NEW DOORS FOR SPORTS WAGERING ..................................................470
   A. Pre-Murphy ..............................................................471
   B. Murphy: The Long Road to “Unconstitutional Commandeering” ..................................................473
   C. Post-Murphy ..............................................................475
   D. Murphy’s Impact on Collegiate Athletics .................479

III. THE EVOLUTION OF GAMBLING FUNDING EDUCATION ........480
   A. Early History ............................................................483
   B. Contemporary History .............................................484
   C. Money Getting Reallocated? ....................................486

IV. MURPHY, THE TCJA, AND COLLEGE SPORTS ........486
   A. The Impact of the TCJA on Higher Education ..........490
      1. Charitable Giving and the Standard Deduction ....491
      2. Athletic Seating & Fringe Benefit Deductions ....492
      3. Excise Tax on Executive Compensation .........493
      4. Excise Tax on Private Institutions’ Endowments ....495
      5. Unrelated Business Income Tax .................496
   B. State Tax Frameworks for Sports Gambling .........496
      1. Taxing Sports Gambling in the United States ....498
      2. Taxing State Gambling Revenue—An Overview ....509

V. DIVIDING THE SPORTS BETTING PIE ..................514
   A. Integrity Fees ....................................................515
   B. Integrity of Collegiate Sports? ..........................516
   C. Gambling Interchange Fees ............................518

VI. CONCLUSION ..........................................................521
I. INTRODUCTION

College athletic departments claim that two recent changes to federal law threaten their viability.\(^1\) First, on December 22, 2017, President Donald Trump signed into law the Tax Cuts and Jobs Act (TCJA);\(^2\) second, in 2018, the Supreme Court of the United States declared that the twenty-six-year-old federal prohibition on sports wagering unconstitutionally commandeered state legislatures.\(^3\) The TCJA was hailed by President Trump and his administration as a simpler tax process benefiting the majority of individuals and businesses in the country.\(^4\) Likewise, Murphy v. National Collegiate Athletic Ass’n was celebrated by states seeking new revenue streams to fight against depleted coffers.\(^5\) Supporting neither decision, however, were college athletic departments.\(^6\) A major aspect of collegiate funding was poised to suffer a significant blow as a result of the TCJA’s tax overhaul.\(^7\) In addition, some college administrators deemed Murphy a major threat to the integrity of competition and the fight to prevent gamblers from manipulating college game results.\(^8\)

---


3. See Murphy v. Nat’l Collegiate Athletic Ass’n, 138 S. Ct. 1461, 1478 (2018). While the Act struck down the Professional and Amateur Sports Protection Act (PASPA), referred to as a prohibition, it was technically more of a freeze, as states that had offered sports wagering in 1992, when the statute was passed, were allowed to continue doing so. See John T. Holden, Prohibitive Failure: The Demise of the Ban on Sports Betting, 35 GA. ST. U. L. REV. 329, 330 (2019).


7. See Murschel, supra note 1 (noting that for some schools, some of which saw a tax change under the new law, donations make up approximately fifty percent of their athletics budget); see also Kisska-Schulze, supra note 2, at 348 (pertaining to the repeal of the charitable deduction).

8. John Keilman, Should Sports Gambling Become Legal in Illinois, Will You Be Able to Wager on Illini, Wildcats or Huskies? Don’t Bet on It., CHI. TRIB. (May 9, 2019),
Although commentators predict the TCJA will have a direct negative impact on college athletic departments, it signifies a greater trend of decreased public university funding throughout the United States. Colleges have expressed reservations that increased legal gambling may escalate the risks of games being fixed. These concerns may, however, be misguided. This Article argues that college administrators should redirect attention towards embracing legalized sports wagering as a means of increasing the integrity of collegiate sporting events, expanding partnerships, and seeking a new source of revenue to supplement income lost due to the TCJA. Murphy opened the door for colleges and universities to capitalize on gambling tax revenue.

Taxing gambling has a contentious history, nearly as controversial as legalizing gambling. The first federal gambling excise tax passed in 1951 at a rate of ten percent of betting handle. It was subsequently reduced to two


13 See Jill R. Dorson, How States Are Spending Their Sports Betting Tax Revenue, SPORTSHANDLE (Oct. 25, 2018), https://sportshandle.com/how-states-are-spending-their-sports-betting-tax-revenue/ [https://perma.cc/2UNU-VWKR] (discussing how states that have legalized sports wagering in the wake of the Murphy decision have decided to allocate the new revenue).

percent, and again to 0.25% in 1984, which remains in effect today.¹⁵ Federal excise tax revenue is allocated to uses that the federal government deems necessary, though in 2018 Senators Schumer and Hatch introduced legislation designating tax revenue for use by a national gambling monitoring body.¹⁶ Allocating gambling funds to sources, including education, remains possible due in part to that bill failing in 2018.¹⁷

Indeed, there has long been a connection between gambling revenue and education.¹⁸ Historically, the challenge for educators has been to keep gaming money from being deployed to other state interests.¹⁹ Of the states that have legalized sports wagering, many are directing excess revenue to general funds; however, Nevada and Washington, D.C. have specifically earmarked sports betting revenue for educational purposes.²⁰ While states often reallocate money reserved for education to other causes,²¹ the nascent sports betting market is relatively well positioned for the imposition of a small tax or fee on sports bets made in order to benefit institutions of higher education.

This Article proposes that public colleges and universities actively lobby for state legislatures to impose a gambling interchange fee (GIF) on wagers made on college sporting events to compensate for economic losses suffered as a result of the TCJA, as well as states’ reallocation of funding away from higher education.²² Unlike “integrity fees,” which have been unsuccessfully promoted by professional sports leagues that already receive significant subsidies from taxpayers and private companies, a GIF would attach only to wagers made on

---

¹⁶ Holden, supra note 5, at 591–92.
¹⁹ Id.
²⁰ Dorson, supra note 13.
²² While there is not an intellectual property right associated with the playing of a sporting event that mandates compensation, unlike professional team sports which are operated by private entities who seek to maximize revenues, public colleges and universities serve a public interest and do not seek profit maximization, and thus have greater justification in seeking an earmark than professional sports leagues, despite the fact that neither is entitled by law to a payment as result of wagering being offered on the contests that they facilitate. See generally Ryan M. Rodenberg et al., “Whose” Game Is It? Sports-Wagering and Intellectual Property, 60 VILL. L. REV. TOLLE LEGE 1 (2014).
college sporting events and be earmarked for disbursement back into states’ underfunded public institutions. These fees would not only benefit colleges and universities directly, but provide revenue to account for expenses tied to monitoring sports-related integrity concerns.

To fully evaluate these issues and proposal, this Article is divided into four substantive sections, followed by a conclusion. Part I examines the sports gambling landscape and what lies ahead following the Murphy decision. Part II analyzes the historical connection between gambling and educational funding. Part III provides an overview of the state tax frameworks surrounding sports gambling, and the TCJA’s impact on higher education. Part IV proposes that colleges and universities pursue a stake in funding from states now legalizing sports gambling through the imposition of a GIF. In conclusion, this Article suggests that colleges and universities could recoup some of their economic losses suffered if, instead of opposing sports wagering outright, they act strategically to attract a portion of this newfound revenue.

II. Murphy v. NCAA—Opening New Doors for Sports Wagering

When Justice Alito wrote that the Professional and Amateur Sports Protection Act (PASPA) unconstitutionally commandeered state legislatures into maintaining laws prohibiting sports wagering, a wave of states expressed interest in authorizing sports betting. States’ interests had little to do with giving individuals access to an activity that many were already illegally engaged in, but much to do with seeking a new revenue stream. Seemingly gone from the contemporary discussion about expanding sports betting are moral concerns that drove much of PASPA’s initial support. The pre-Murphy era represented a very different time for how Americans and politicians perceived sports betting.

---

23 The integrity fee is a fee that the NBA and Major League Baseball have repeatedly sought from states seeking to legalize sports wagering. It was initially suggested that the fee would compensate the leagues for added costs resulting from threats to the integrity of sporting events, like fixed games, as a result of legalized gambling. But, the language surrounding the fee changed to “royalty,” despite there not being a property right associated with sports gambling that would typically receive such a fee. In reality, what the leagues sought was a private tax of sorts. See John Holden, When They Say Integrity Fee, Are Pro Sports Leagues Really Asking for a Private Tax?, LEGAL SPORTS REP. (Nov. 30, 2018), https://www.legalsportsreport.com/26361/pro-sports-leagues-integrity-fee-private-tax/.

24 See infra Part V.C.


To better understand Murphy’s evolution and ultimate impact, this Part proceeds as follows: Section A discusses the history of gambling in the United States pre-Murphy, Section B analyzes the dichotomy of disputes leading to the U.S. Supreme Court’s 2018 decision in the case, Section C examines states’ momentum to legalize sports gambling post-Murphy, and Section D opines on the impact of Murphy on higher education.

A. Pre-Murphy

America’s connection with gaming began more than 100 years before the establishment of the nation itself. The Jamestown lottery, as it was known at the time, funded the first English colony in North America. Lotteries were far from the only gambling activities with which Americans engaged in prior to independence. New York first established horse racing in 1666. Lotteries during the pre-Revolution era were the most feasible means of funding both private and public infrastructure projects, as banks were not sufficiently widespread to make them accessible to many colonial areas. In the years leading to the Civil War, gambling in America fell out of favor, and many states embedded anti-gambling provisions into their new constitutions. After the Civil War, much of the opposition to gambling and lotteries was dismissed, particularly since Reconstruction funds were needed and a lottery was preferred over a legislated tax.

During this period, Western expansion brought significant levels of gambling to the West; however, as cities developed, gambling became abbreviated. In 1890, transportation improvements benefitted lotteries, and congressional concerns about the distribution of interstate lottery materials resulted in a federal ban on mailed lottery circulations. Scandals associated with the lotteries, combined with Congress’s clampdown, led to a legal gambling depression until 1931, when Nevada re-legalized casino gambling. It was not until 1964 that New Hampshire became the first state to re-authorize a state lottery.

The 1960s were an important era for federal anti-gambling legislation; while most states maintained gambling prohibitions, organized crime developed a

30 Id.
32 Id.
33 Id. at 369.
34 Id. at 370.
35 Id. at 370–71.
36 Id. at 372.
37 Rose, supra note 31, at 374.
38 Id.
monopoly on the industry, and in particular, they focused on bookmaking. In 1961, Congress passed the Wire Act based on recommendations from the Kefauver Committee, which conducted a fourteen-city tour of the country investigating organized crime. Along with the Travel Act and the Wagering Paraphernalia Act, Congress sought to target organized crime’s money-making operations. The legislature passed an omnibus crime bill in 1970, adding to the 1960s legislation that included one of the federal government’s most powerful tools for disrupting gambling activities, the Illegal Gambling Business Act.

In 1990, Congress first attempted to pass a bill that would eventually become PASPA. Originally, it was not a criminal law or civil prohibition; instead, it was a piece of legislation regulating intellectual property rights. Two years later, the intellectual property bill morphed into a statute referred to as an “oddbity” and “Orwellian.” PASPA was no longer an intellectual property statute; it instead unleashed a sports gambling ice age. The bill froze state laws in place, allowing states that were already operating sports betting schemes to continue, while estopping states that had not been offering sports betting at the time of passage. There was, however, an exception made for a single jurisdiction that was granted a one-year window to enact a law allowing sports wagering. The one-year window for Atlantic City, New Jersey closed without the state exercising its right, leaving Nevada as the sole jurisdiction with sportsbook style wagering.

---

39 See generally David G. Schwartz, Cutting the Wire: Gaming Prohibition and the Internet (2005).
43 Id. § 1953.
44 Id. § 1955. The Illegal Gambling Business Act effectively criminalizes intrastate gambling activity provided that the business involves more than five people, is in substantially continuous operation for thirty days, or has gross receipts of $2000 in a single day. Id.
45 Holden, supra note 3, at 337.
46 Id.
50 Id. § 3704(3). While the statute makes no geographic reference by the terms used, only Atlantic City, New Jersey qualified. See id.
51 Holden, supra note 3, at 354. Various states had limited exemptions; most commonly Delaware, Oregon, and Montana are cited for their limited exemptions, but certain activities
After Congress passed PASPA, Atlantic City—otherwise known as “America’s Playground”—struggled financially.52 Gambling revenues decreased, as did airport and road traffic.53 The city faced competition from a variety of places, including Pennsylvania, which legalized casino gambling after PASPA passed.54 It was Atlantic City’s decline that primed New Jersey to seek legalized sports gambling in an effort to boost tourism back into the city.55

New Jersey lawmakers sought to recapture a portion of the purported hundreds of billions of illegally wagered dollars.56 The exact size of the illegal sports wagering market is not known, but estimates range from $80 billion to $650 billion.57 State lawmakers hoped that regulated gambling would draw funds otherwise routed to illegal bookmakers, even if only a small percentage of market reserves were recaptured.58 In 2012, Chris Christie, then-Governor of New Jersey, signed a law authorizing the state’s casinos and racetracks to offer sportsbook-style betting.59

B. Murphy: The Long Road to “Unconstitutional Commandeering”

Almost immediately following Governor Christie’s signing of the 2012 bill into law, a quintet of sports organizations—the National Collegiate Athletic Association (NCAA), the National Football League (NFL), the National Basketball Association (NBA), the National Hockey League (NHL), and Major League Baseball (MLB)—sued the Governor and other New Jersey officials under a PASPA provision granting sports organizations the same power to enforce the statute as the U.S. Attorney General.60 The district court opinion, in what would become known as Christie I, was a near-total success for the sports organizations.61 The plaintiffs prevailed on PASPA being a rational exercise of

like wagering on Keirin bicycle racing is legal in New Mexico. Id.; see also Rodenberg & Holden, supra note 47, at 15–16.


53 Id.

54 Id.

55 Brent Johnson & Keith Sargeant, Here’s All You Need to Know About N.J. Sports Betting Before It Launches on Thursday, NJ.COM (June 12, 2018), https://www.nj.com/politics/2018/06/all_you_need_to_know_as_nj_legalizes_sports_betting.html [https://perma.cc/V7TN-AZP7].


58 Holden, supra note 5, at 577–78.

59 Holden, supra note 3, at 356.


61 See id. at 579 (holding for the sports organizations).
Congress’s Commerce Clause powers that were found to neither offend the Tenth nor Fourteenth Amendments, nor violate the “Equal Footing Doctrine.” The defendants appealed the decision to the Third Circuit Court of Appeals. The panel returned a two-one decision in favor of the sports organizations. The majority held that the leagues have standing to sue the state to stop implementation of the law, and that PASPA was a valid exercise of Congress’s Commerce Clause powers. The Third Circuit further held that PASPA did not impermissibly commandeer the states, stating, “PASPA does not require or coerce the states to lift a finger—they are not required to pass laws, to take title to anything, to conduct background checks, to expend any funds, or to in any way enforce federal law.” The Supreme Court’s decisions in *New York* and *Printz* formed the basis for the New Jersey defendants’ arguments, as it was their contention that PASPA required the state legislature to refrain from exercising their sovereign power to pass and repeal laws, without offering a supplemental federal regulatory scheme on which it could rely if the state chose not to enforce its laws any longer. Nor did the Third Circuit find support for the defendant’s argument that PASPA violated the Tenth Amendment, or that the law violated the equal sovereignty principle that states must be treated substantially the same. Despite a dissenting opinion by Third Circuit Judge Vanaskie, who argued that PASPA violated the anti-commandeering principle, the sports leagues prevailed on appeal in *Christie I*, and again when the Supreme Court denied certiorari.

Not to be deterred, in 2014 New Jersey repealed its laws prohibiting sports betting as proposed by the Solicitor General’s office. Following the repeal, the sports organization quintet sued the state for a second time. District Court Judge Michael Shipp—who oversaw *Christie I*—found that the law partially repealing New Jersey’s 2012 law still ran afoul of PASPA and struck it down.

---

62 *Id.*  
64 *Id.* at 241.  
65 *Id.* at 240–41.  
66 *Id.* at 231 (emphasis omitted).  
67 *Id.* at 227–32.  
68 *Id.* at 236–38.  
69 Nat’l Collegiate Athletic Ass’n v. Governor of New Jersey, 730 F.3d at 241–51.  
72 Brief for the United States in Opposition at 11, Nat’l Collegiate Athletic Ass’n v. Governor of New Jersey, 730 F.3d 208 (3d Cir. 2013).  
74 *Id.* at 503–06.
Fuentes—who wrote the majority opinion in *Christie I*—that the 2012 repeal was PASPA-compliant. In October of 2015, the state’s petition for rehearing en banc was granted. Despite being granted the somewhat extraordinary rehearing en banc, New Jersey lost for the sixth time in six attempts. The state’s 9-3 loss at the Third Circuit en banc hearing left the state with one final attempt to appeal the decision to the U.S. Supreme Court for a second time.

The Supreme Court granted certiorari on the question of whether PASPA impermissibly coerced state legislatures into maintaining laws in June of 2017, and heard oral arguments in December that same year. On May 14, 2018, the Supreme Court issued a decision in the *Murphy* case (*Christie II*), holding that “[the PASPA provision at issue here—prohibiting state authorization of sports gambling—violates the anticommandeering rule. That provision unequivocally dictates what a state legislature may and may not do.” The Supreme Court further held that PASPA could not be saved by severing the portions of the law that impermissibly commandeered state legislatures, as without those, the statute would necessarily fail; unlike under permissible preemption schemes, there was no federal framework that the state could default to for enforcement. In concurrence, Justice Thomas further opined that while PASPA impermissibly commandeered the state’s legislature, he was unconvinced that sports gambling was an interstate commerce issue. The *Murphy* decision resulted in PASPA being deemed unconstitutional and, consequently, interested states could pass laws regulating sportsbook style sports wagering outside of Nevada for the first time since 1992.

C. Post-Murphy

Even prior to the Supreme Court issuing the *Murphy* decision, states had begun preparing for a post-PASPA landscape. New York and Pennsylvania

---

76 Nat’l Collegiate Athletic Ass’n v. Christie, 832 F.3d 389, 402 (3d Cir. 2016) (en banc).
77 See id.
80 Id. at 1484.
81 Id. at 1485 (Thomas, J., concurring) (“Unlike the dissent, I do ‘doubt’ that Congress can prohibit sports gambling that does not cross state lines.”). This position was raised in one of the amicus briefs filed in the case. See Brief of Amicus Curiae Researcher John T. Holden in Support of Petitioners at 29–33, *Murphy v. Nat’l Collegiate Athletic Ass’n*, 138 S. Ct. 1461 (2018).
passed legislation to legalize sports gambling in the event of PASPA’s failure, whether it came judicially or legislatively.\textsuperscript{83} While there was great exuberance from coast to coast following the removal of the federal prohibition, the reality of passing gaming legislation slowed much of the early state momentum.\textsuperscript{84} Still, despite the challenges of crafting sports betting regulations, with various stakeholders seeking a portion of the finite revenues, more than thirty states have now introduced bills to legalize sports betting since May 14, 2018.\textsuperscript{85}

Aside from Nevada, which has experienced legalized sportsbook-style betting since 1931, several states took early advantage of the newfound opportunity.\textsuperscript{86} While observers assumed New Jersey would offer sportsbook-style wagering shortly after the \textit{Murphy} decision,\textsuperscript{87} Delaware accepted the first wagers.\textsuperscript{88} New Jersey began accepting wagers shortly thereafter, and became the first state outside Nevada to allow mobile sports wagering.\textsuperscript{89} Included in the


\textsuperscript{84} Gaming authorization is a politically controversial topic in many states, and involves numerous stakeholders, including negotiations with tribal governments in a number of states. Tribal gaming negotiations for Class III activities, like sports betting, typically require the state and tribes to sit down and negotiate a compact, which is a process that has been historically contentious in many states. See Holden, supra note 5, at 607–08.


initial wave of states to legalize sports betting were Mississippi, West Virginia, Pennsylvania, and Rhode Island.

In addition, Arkansas legalized sports betting via a referendum in 2018, Iowa, Indiana, and Montana all legalized sports betting in May 2019. Following these early movers, Tennessee passed a sports betting law that was significant, as it was the first to mandate that sportsbook operators purchase official data from sports leagues—a legally dubious mandate—without a commercial justification. After failing to secure an official data order in the first year after Murphy, the sports leagues achieved a second victory in Illinois

when the state legislature granted an official mandate involving certain types of in-play wagers.\textsuperscript{101} While numerous states have expressed an interest in offering sports wagering, many hold complicated relationships with Tribal governments that dictate mandatory negotiating between the state-tribe and approval of the Secretary of the Interior before the state can offer sports wagering.\textsuperscript{102}

For example, New Mexico, which sought an exemption for Keirin bike racing during PASPA’s legislative hearings,\textsuperscript{103} has launched sports wagering at five tribal casinos.\textsuperscript{104} The Santa Ana Star Casino and Hotel was the first to introduce sports wagering in New Mexico, which was the result of permissive language used in the gaming compact between the state and the Tribe.\textsuperscript{105} In fact, the gaming compact grants virtually unfettered ability to New Mexico tribes to offer Class III gaming activities like sports betting, stating, “Permitted Class III Gaming. The Tribe may conduct, only on Indian Lands, subject to all of the terms and conditions of this Compact, any or all forms of Class III Gaming.”\textsuperscript{106} The agreement between state and Tribal governments seemingly grants the tribes, subject to the compact, extensive authority to authorize Class III gaming at their facilities.\textsuperscript{107} The New Mexico tribal compact is very permissive, allowing the Santa Ana Pueblo and subsequently Pojoaque Pueblo\textsuperscript{108} to begin accepting wagers. The compact’s language is also exceptional, as most compacts specifically address which activities are permissible.\textsuperscript{109} Sports wagering negotiations between Tribes and other states like Florida and Oklahoma, which have significant Tribal gaming infrastructure but lack a large


\textsuperscript{102} Holden, supra note 5, at 607–08.

\textsuperscript{103} See Rodenberg & Holden, supra note 47, at 15–16.


\textsuperscript{106} id.

\textsuperscript{107} See id.


Tribal gaming footprint, will likely fall behind states like New Mexico.\textsuperscript{110} Indeed, despite the gaudy numbers often associated with sports betting, the activity is a relatively low-margin product for gaming operators.\textsuperscript{111} Tribes must carefully evaluate whether to open up Tribal gaming compacts and risk giving up something to state governments in order to offer a product that will not likely add significant value to a casino’s bottom line.\textsuperscript{112}

Despite the relatively low profit margins, legalized sports wagering has attracted a number of foreign sports wagering companies, as well as welcomed the conversion of the two largest daily fantasy sports providers into sports wagering operators.\textsuperscript{113} These companies have also moved from online-only fantasy sports platforms to mobile and brick-and-mortar sports betting operators.\textsuperscript{114} Both FanDuel and DraftKings have had immediate success, particularly in the mobile betting space, and quickly acquired more than seventy percent of New Jersey’s online gambling market share.\textsuperscript{115} While states have been slow to adopt regulations and launch mobile wagering, the potential to capitalize on additional revenue is clear.\textsuperscript{116} However, collegiate athletics administrators remain adamant in opposing legalized sports wagering.\textsuperscript{117}

D. Murphy’s Impact on Collegiate Athletics

The impact of Murphy on collegiate sports remains to be seen in the protracted litigation against the State of New Jersey.\textsuperscript{118} College athletic administrators have long opposed legalized gambling for fear that gamblers would induce student-athletes into manipulating matches.\textsuperscript{119} College athletes are perhaps vulnerable to match-fixers by virtue of their poverty; as amateur

\textsuperscript{110} See Holden, supra note 5, at 619.
\textsuperscript{112} See Holden, supra note 5, at 607–08.
\textsuperscript{116} By November 2019, mobile wagering accounted for approximately eighty percent of New Jersey’s sports betting market. See id.
athletes they may not receive the multi-million dollar salaries of professional athletes.\(^{120}\) Since match-fixers have historically targeted student-athletes, college administrators may harbor legitimate fears.\(^ {121}\)

Concerns that legalized gambling will promote match-fixing and increase gambling amongst collegiate athletes are largely speculative.\(^{122}\) Some states have acknowledged college athletic departments’ apprehensions by placing bans on in-state collegiate team wagering in order to limit student-athletes’ exposure to would-be match-manipulators.\(^{123}\) The challenge for college administrators is that while such fears are theoretical, athletic departments and universities will face increased monitoring and compliance costs until sports gambling is proven safe. Part III addresses the evolving relationship between gambling funding and education in the United States.

### III. The Evolution of Gambling Funding Education

Following September 11, 2001, the U.S. economy lapsed into a recession that eventually culminated in a period referred to as the Great Recession.\(^ {124}\) The nation’s financial erosion caused Congress to curtail state and local governmental funding, forcing states to procure alternative revenue sources to offset lost subsidies.\(^ {125}\) Taxes on alcohol, cigarettes, strip clubs, soda, and marijuana emerged, furnishing innovative opportunities to boost state revenue production streams.\(^ {126}\) In addition, some states launched significant legal battles to expose fresh avenues of revenue, including successes in collecting sales taxes from nonresident online retailers and the newly expanded legalized gaming

---

\(^{120}\) Holden & Rodenberg, supra note 11, at 465.

\(^{121}\) See id. at 464–65. College athletes have been the most frequent human targets of match-fixers, and horse racing is fixed at a higher rate than any human based sport. See id. at 458.

\(^{122}\) Regulated gambling markets have protections that unregulated markets do not have, though some studies have shown a proclivity of NCAA athletes towards sports wagering, which with increased availability may or may not exacerbate issues. See Jeffrey L. Derevensky & Tom Paskus, Mind, Body and Sport: Gambling Among Student-Athletes, NCAA, http://www.ncaa.org/sport-science-institute/mind-body-and-sport-gambling-among-student-athletes [https://perma.cc/KH7B-FQA2].

\(^{123}\) See, e.g., John Holden, How NJ Sports Betting Set Up a Potential Constitutional Problem for Everyone, LEGAL SPORTS REP. (June 7, 2019), https://www.legalsportsreport.com/32820/in-state-nj-sports-betting/ [https://perma.cc/A4SR-V9E8]. This practice is likely futile, nonsensical, and unconstitutional. Not only does this have little effect when one looks at states like New Jersey and Pennsylvania, where New Jersey bans wagering on in-state teams, but those same bettors can cross a bridge and bet on those games in Pennsylvania, banning wagering on in-state teams because of a threat to game integrity raises significant problems under a dormant commerce clause analysis. Id.


\(^{125}\) Id. at 326–27.

\(^{126}\) Id. at 327–30.
activities. Revenue from these various sources ultimately falls into one or more state buckets—either a general fund to be used for broad discretionary purposes, or into numerous designated (earmarked) reserves for use toward specified public services.

To a certain degree, every state earmarks a percentage of tax revenue to fund governmental activities. Earmarked beneficiaries vary extensively but tend to include highways and roadways, local government programs, education, health and welfare, and environmental causes. Likewise, the sources of dedicated revenue are diverse, often complementing the political atmosphere of the jurisdiction. These sources may take the form of either “sin taxes” on items such as tobacco, alcohol, or soda, or Pigouvian taxes on select pollutants, like motor fuel.

Funding higher education has long challenged state governments. While a significant portion of earmarked revenue supports state K-12 programs, the Great Recession forced jurisdictions to reduce their fiscal support of post-secondary education. This alteration placed a hefty burden on colleges and universities, requiring them to seek unique outlets to supplement about forty-

---


131 See Tahk, supra note 129, at 766.


seven percent of their total overhead costs. As a result, tuitions magnified, online enrollments increased, federal aid programs expanded, students borrowed, programs and faculty lines were cut, and charitable donations became critical to the vitality of higher education. These changes have resulted in a cost-value educational crisis in the United States, with public institutions becoming less affordable and accessible amid quality erosion. There is no shortage of academic scholarship devoted to addressing this cataclysm in higher education.

In the ten years following the 2008 recession, states collectively spent $9 billion less on higher education than they did in the year preceding its start. While five states spent more money per student in 2017 than they did in 2008, most states decreased funding by about sixteen percent per student during that period, and funding in eight states fell by more than thirty percent. Budget cuts have largely been passed onto students, who today each contribute roughly $1800 more toward university budgets than in the previous ten year period. Governmental spending on higher education has diminished by twenty-five

---

136 See Michael Mitchell et al., Ctr. on Budget & Policy Priorities, A Lost Decade in Higher Education Funding: State Cuts Have Driven Up Tuition and Reduced Quality 1–29 (Aug. 2017), https://www.cbpp.org/sites/default/files/atoms/files/2017_higher_ed_8-22-17_final.pdf (noting that states provide about fifty-three percent of the costs associated with higher education, requiring that institutions themselves make up the difference through tuition and/or budget cuts).

137 See Barr & Turner, supra note 135, at 170–86 (discussing various enrollment responses to the Great Recession); see also Mitchell et al., supra note 136, at 1–29 (documenting that budget cuts have been instituted due to state funding cuts); Kisska-Schulze, supra note 2, at 371 (noting that in 2017 charitable donations made to U.S. colleges and universities rose six percent, resulting in $43.6 billion in revenue that year).


140 Mitchell et al., supra note 136, at 1.

141 Id. at 2.

percent. Financial challenges have long been part and parcel of operating higher education institutions, and history supports that state lottery revenues have helped keep higher education in operation. To examine the developing partnership between lotteries and higher education, this Part proceeds as follows: Section A discusses the early history of the American lottery system, Section B examines contemporary applications of lottery revenue for educational purposes, and Section C scrutinizes state initiatives to reallocate lottery funds away from education.

A. Early History

While modern gambling in America largely began around 1950, several hundred years earlier gambling played a very different role in the United States. The first lottery in England is believed to have been drawn in 1569 and was intended to generate funds for harbor repairs and other public works projects. English lotteries evolved, with the Crown finding them useful revenue mechanisms for a variety of projects, including plantation funding in the Virginia colony. As Europeans colonized America, they brought their English gaming traditions with them, though Puritan rule resulted in statutes dating to 1631 that banned gambling in states like Massachusetts. Lotteries, however, received different treatment than card and dice games. Lotteries were a necessity for American colonists. Prior to 1790, the country had only three incorporated banks, and the Crown kept a tight grip on funding going to the New World. As a result, colonists were forced to find alternative ways to generate revenue in order to develop, and provide for, colonial defense.

Lotteries were initially used as a private tool but quickly spread as a popular funding mechanism for both public and private uses. While Massachusetts imposed a twenty-pound fine for running a lottery in 1719, colonists were able to avoid the fines by entering lotteries in other colonies that were less averse to the contests. Despite the passage of law meant to rein in New York lotteries in 1721, lotteries remained popular, and many were even state-run despite the

---

144 See infra Part III.B.
146 Id. at 228.
147 Id.
148 Id. at 237.
149 See id. at 238.
150 See id.
151 Blakey, supra note 145, at 238.
152 Id.
153 Id.
154 Id. at 240–41.
restriction. Indeed, it was during this era that one of the first linkages between the lottery and higher education took place, as it was by means of a lottery that Kings College (now Columbia University) was established. Pennsylvania similarly utilized a lottery for the construction and initial operation of Dickinson College. New York and Pennsylvania were not isolated in their use of lotteries for establishing universities: Harvard, Yale, Dartmouth, Princeton, and William and Mary were all at least partially established using funds from lotteries.

Following the Civil War, lotteries and gambling taxes became an important means of generating revenue for the reconstruction of the South and the establishment of more than fifty universities. Lotteries in the 1800s were typically privately run via a license obtained from the state. The 1800s also saw a rise in the number of scandals associated with lotteries, leading some lawmakers to question their continued operation. It was the Louisiana lottery that generated a scandal of such proportion that Congress took interest in banning the above discussed use of the mail for distribution of lottery materials, resulting in a seventy-year freeze on legal state-run lotteries.

B. Contemporary History

In March of 1964, New Hampshire cities and towns voted to allow sweepstakes ticket sales. Two days later, the state became the first to sell lottery tickets in almost seven decades. Since the launch of the 1964 sales, forty-three other states, as well as the District of Columbia and Puerto Rico, implemented state-run lotteries. In 2008, Americans spent $60.6 billion on lottery tickets, and spent “well over a trillion dollars” across all gambling activities. Lotteries, in their contemporary iteration, serve two predominant

155 Id. at 248.
156 Id.
157 Blakey, supra note 145, at 253.
159 Id.
160 Rose, supra note 31, at 368–69.
161 Id. at 369.
163 About Us, NH LOTTERY, https://www.nhlottery.com/About-Us [https://perma.cc/5WZK-UMU4].
164 Id.
165 Id.
166 Rose, supra note 31, at 361–62.
state purposes: first, to supplement the state’s general fund; and second, to supplement education funding.167

Educational programs—particularly primary and secondary education—are paramount beneficiaries of various state earmarked tax revenues.168 For example, in 2005 Alabama earmarked forty-percent of its tax revenue from beer sales to fund public education.169 The first $40 million of Colorado’s annual tax revenue from marijuana sales goes to the Colorado Public School Capital Construction Assistance Fund, which subsidizes the Building Excellent Schools Today (BEST) grant program.170 Another major source of educational funding comes from gaming tax revenue.171 In particular, state lotteries serve as attractive options for profit realization dedicated to funding public education.172 An example of this is the Helping Outstanding Pupils Educationally (HOPE) scholarship program, which is wholly funded by the Georgia Lottery.173

The practice of earmarking lottery proceeds for a specific cause within the state budget is common.174 One challenge facing lawmakers and funding recipients is maintaining additional revenue outlets, since states have a historical tendency to replace, rather than supplement, gaming revenues.175 Despite this setback, one study found that earmarking increases educational spending by more than thirty-five cents when compared to one non-earmarked dollar, suggesting the net benefit of gaming funds earmarked for educational funding has a positive effect.176 Presently, at least twenty states direct all of their lottery revenue to education, with others allocating at least a portion of the revenue

168 PÉREZ, supra note 130, at 3 (noting that thirty-five states dedicate selective tax revenues for educational funding).
170 COLO. CONST. art. XVIII, § 16(5)(d).
172 See Crowley & Hoffer, supra note 169, at 121.
173 RUSSELL, supra note 171, at 3.
175 Id. at 300–01.
directly to education funding. However, the practice of earmarking has not resulted in funding actually reaching the state’s educational coffers.

C. Money Getting Reallocated?

While lawmakers tout that state gambling dollars fund education, the reality is that the money earmarked for education is often usurped by other governmental agencies and expenses. The State of Maryland is a prime example. After legalizing casino gambling in 2008, Maryland’s casinos deposited $1.7 billion into the state’s education trust account. While the casino money is purportedly going to education, the state has reallocated some of that money to other state causes. Maryland is not an isolated case. However, some state lotteries that fund scholarship programs, like Georgia’s HOPE Scholarship, continue to survive and retain promised lottery funding.

While sports betting currently remains a low-revenue gaming product for commercial operators, its legalization creates opportunities for increased state educational income sources. Facing an already steady decline of state funding for higher education, college athletic departments face new financial perils with the signing of the TCJA into law. In Part IV, we discuss the financial impact of the TCJA on college sports and the broader spectrum of higher education, and present current data regarding states’ initiatives to tax gambling revenue post-Murphy.

IV. MURPHY, THE TCJA, AND COLLEGE SPORTS

College sports are not immune to the effects of state funding cuts. Intercollegiate athletics are subsidized through two sources—allocated and generated revenues. Allocated revenues derive from student fees, institutions,
and state and local governmental support.\textsuperscript{187} Generated revenues are earned directly by athletic departments, and include income from ticket sales, broadcasting contracts, alumni donations, and royalties.\textsuperscript{188} Cumulatively, allocated and generated revenue subsidize coaching salaries, student-athlete grants-in-aid, recruitment expenditures, building and grounds maintenance, travel expenses, and uniforms.\textsuperscript{189} The revenue needed to cover these costs varies significantly across institutions, with most athletic programs falling in the red.\textsuperscript{190} In 2014, the average Football Bowl Subdivision (FBS) athletic department subsidy from their parent institution was $14.7 million.\textsuperscript{191} With institutions regularly facing budget deficits, internal measures to offset athletic expenditures are growing, and include increased student fees, purging non-revenue sports, and shifting budgets away from select academics.\textsuperscript{192}

Although not all agree that allocated revenue—like mandatory student fees—should be earmarked for college sports,\textsuperscript{193} successful university athletic

\textsuperscript{187} Id. Direct institutional support includes funds directed specifically to athletics, while indirect institutional funds include payment of utilities, maintenance, and support staff salaries. Id.

\textsuperscript{188} Id.


\textsuperscript{191} The 25 Universities that Spend the Most on Athletics, SPORTS MGMT. DEGREE HUB (July 24, 2018), https://www.sportsmanagementdegreehub.com/the-universities-that-spend-the-most-on-athletics/ [https://perma.cc/CN53-ZDT9].

\textsuperscript{192} See David Ridpath, Who Actually Funds Intercollegiate Athletic Programs?, FORBES (Dec. 12, 2014), https://www.forbes.com/sites/cap/2014/12/12/who-actually-funds-intercollegiate-athletic-programs/#429a297f71af [https://perma.cc/E4W5-JUCN] (documenting that student fees have increased at a rate thirteen percent higher than tuition); see also Taurus Myhand, A Dream Still Deferred: The Unlawful Use of Student Fees for Instructional Technology in an Alabama Public School Causing a Disparate Impact for Minority Children, 19 Rutgers Race & L. Rev. 77, 82 (2018) (documenting that budget deficits are resulting in increased student fees); Lora Wuerdeman, Note, Sideline Business in Intercollegiate Athletics: How the NCAA Can De-escalate the Arms Race by Implementing a Budgetary Allocation for Athletic Departments, 39 N.C. Cent. L. Rev. 85, 87 (2017) (noting that institutions have taken drastic measures to offset college athletic costs).

\textsuperscript{193} See, e.g., Jon Marcus, Should College Students Have to Pay Fees that Go to Groups They Don’t Support?, WASH. POST (Oct. 29, 2017), https://www.washingtonpost.com/local/education/should-college-students-have-to-pay-fees-that-go-to-groups-they-dont-support/2017/10/29/79276a24-bbe3-11e7-9e58-e628854af98_story.html?utm_term =.08ce66c7bd [https://perma.cc/Z915-J9HM] (reporting on proposals to give students the right to opt out of mandatory student fees geared towards certain groups, including
programs have positive impacts on institutional reputation.\textsuperscript{194} For some schools, university success is directly linked to athletic performance.\textsuperscript{195} Strong athletic programs boost enrollment, amplify college profiles, promote an internal culture of school spirit, and warrant long-term alumni relations.\textsuperscript{196} While most NCAA athletic programs depend on allocated revenue, self-sufficiency is increasing.\textsuperscript{197} Football and men’s basketball programs—the revenue-generating sports at most institutions—have evolved into a billion-dollar industry.\textsuperscript{198}

The history of amateur sports commercialization dates back to the early part of the twentieth century.\textsuperscript{199} Since then, the economic demands of college sports have amplified their commercialization, particularly with respect to the “money sports” subsidizing non-revenue generating athletics.\textsuperscript{200} The discourse surrounding the big business of college sports exploded after Johnny Manziel’s “show me the money” gestures, Ed O’Bannon’s antitrust litigation, and Northwestern University football players’ attempt to unionize and bargain collectively.\textsuperscript{201} Although dialogue surrounding pay-for-play, professionalizing...
 colleges sports, and the continued arms race of coaching salaries continues to grow,
the reality is that only a small fraction of university athletic programs generate revenue. Instead, like most academic and non-academic programs housed within college and university systems, allocated revenues are the lifeblood of college sports.

President Trump’s signing of the TCJA revolutionized the U.S. Tax Code. The law includes certain provisions that directly and negatively impact tax-exempt organizations, which include most colleges and universities. Certain changes made to individual and corporate taxes could further exacerbate financial ramifications for higher education. In addition, select provisions

---


203 Darren A. Heitner, Money & Sports: Economic Realities of Being an Athlete, 8 DEPAUL J. SPORTS L. & CONTEMP. PROBS. 161, 170 (2012) (noting that most collegiate athletic departments do not generate positive revenue); see also Kevin Brown & Antonio Williams, Out of Bounds: A Critical Race Theory Perspective on ‘Pay for Play,’ 29 J. LEGAL ASPECTS SPORT 30, 32 (2019) (documenting that the primary revenue-generating collegiate sports are NCAA Division I Football Bowl Subdivision (FBS football) and Division I men’s basketball).

204 Heitner, supra note 203, at 170 (discussing that most athletic departments rely on their parent institutions for funding); see also Berry, supra note 189, at 199 (noting that most athletic departments lose money).

205 Kisska-Schulze, supra note 2, at 347.

206 See infra Parts IV.A.1–5.
directly target college sports. Combined, these changes could have serious impacts on higher education and collegiate athletics, resulting in the need for universities and their affiliated sports programs to secure new methods of funding. The Murphy decision could promulgate innovative channels for state governments to reinforce financial support to college athletic programs and higher education as a whole. To better analyze this initiative, it is important to first address the TCJA’s effect on colleges and universities, as well as the evolving state tax frameworks post-Murphy. To examine these issues, this Part proceeds as follows: Section A discusses the impact of the TCJA on higher education, to include collegiate athletic programs, and Section B introduces the current status of state tax gambling initiatives.

A. The Impact of the TCJA on Higher Education

The TCJA significantly reformed U.S. federal tax legislation for the first time in over thirty years, affecting virtually every echelon of the American economy. As part of the overhaul, individual income tax rates and brackets were reduced. The corporate tax rate transitioned from a graduated scale to a flat twenty-one percent. The standard deduction substantially increased, while the highest individual tax rate dropped to thirty-seven percent. Moving costs and alimony payments became non-deductible, as did business-related entertainment expenses. Numerous changes were made to taxing foreign income.

In addition, the TCJA introduced new provisions that could directly or indirectly affect higher education and the college sports arena. Five provisions will change how colleges, universities and athletic departments operate. These include: (1) increasing the standard deduction for individual taxpayers; (2) eliminating the deduction previously available to charitable donors’ receiving rights to athletic seating; (3) establishing a new excise tax on tax-exempt

---

207 See infra Parts IV.A.1–5.
210 Tax Cuts and Jobs Act §§ 11001(a)–(b).
211 Id. § 13001(a)–(b).
212 Id. § 11021(a).
213 Id. § 11001(a).
214 Id. § 11051(a), 131 Stat. at 2089 (repeal of 26 U.S.C. § 215); see also § 11049(a), 131 Stat. at 2088 (codified at 26 U.S.C. § 217) (stating that the deduction for moving expenses is suspended for 2018–2025 (except for active-duty military)).
organizations’ executive compensation; (4) including a new excise tax on private institutions’ endowments; and (5) providing new tax provisions for colleges and universities earning unrelated business income (UBI). Each of these provisions is discussed below separately.

1. Charitable Giving and the Standard Deduction

Following Congress’s reconstruction of the Tax Code, changes made that directly impact individual taxpayers could also have an indirect effect on charitable contributions made to higher education. Through 2025, individual taxpayers can elect to take a significantly higher standard deduction on their individual income tax returns than previously available. For most individuals, electing to itemize is no longer the optimal choice. A 2019 study found that eighty-eight percent of taxpayers will now elect the standard deduction in lieu of itemizing. Congressional data further estimates that 18 million households elected to itemize in 2018, down from 46.5 million in 2017.

A negative side effect of this change is that taxpayers electing the standard deduction may moderate or eliminate their charitable contributions moving forward. Without itemizing, 28.5 million households will lose the charitable tax benefit associated with gifting. Empirical evidence suggests that tax incentives directly affect gifting by lower and middle-class taxpayers. In fact, early indicators predict that doubling the standard deduction could result in a $12 to $20 billion reduction in annual gifting. Although recent data found

---

217 See infra Parts IV.A.1–5.
218 Tax Cuts and Jobs Act § 11021(a), 131 Stat. 2054, 2072 (2017) (codified at 26 U.S.C. § 63(c)(7)(A)(ii)). In 2017, the maximum standard deduction amounts were set at $6350 for individuals, $9350 for heads of household, and $12,700 for married filing jointly. The TCJA almost doubled those amounts, increasing the maximum thresholds to $12,000, $18,000, and $24,000 respectively.
222 IND. UNIV. LILLY FAMILY SCH. OF PHILANTHROPY, supra note 220, at 7.
223 Kisska-Schulze, supra note 2, at 370 (citing to Alice Gresham Bullock, Taxes, Social Policy and Philanthropy: The Untapped Potential of Middle- and Low-Income Generosity, 6 CORNELL J.L. & PUB. POL’Y 325, 340 (1997)).
that contributions to colleges and universities actually exceeded record highs in 2018, it also established that the total number of individual donors declined, suggesting that middle-income taxpayers were less inclined to make charitable donations last year as compared to wealthy donors.\footnote{Nick Hazelrigg, \textit{Larger Donations, Fewer Donors}, \textsc{Inside Higher Ed} (June 20, 2019), https://www.insidehighered.com/news/2019/06/20/donations-colleges-are-number-donors-down [https://perma.cc/E3VA-GBR7].}

2. Athletic Seating & Fringe Benefit Deductions

Collegiate athletic departments depend on private contributions from alumni and boosters to subsidize their sports programs.\footnote{Javonte U. Lipsey, \textit{Assessing Fundraising Practices of Intercollegiate Athletic Departments: An Empirical Analysis of Tiered Reward Systems} 1 (2019) (unpublished Master’s thesis, University of North Carolina at Chapel Hill) (on file with \textsc{Ohio State Law Journal}).} Annually, NCAA DI FBS fundraising accounts for $26.72 million.\footnote{Id.} Earmarked gifts from private donors to college athletic departments often support the building, renovation, or expansion of athletics facilities.\footnote{Athletics, \textsc{Inside Philanthropy}, https://www.insidephilanthropy.com/campus-cash/athletics [https://perma.cc/F96F-SR3V].} Smaller gifts help fund named scholarship endowments, academic opportunities for student-athletes, and athletic leadership programs.\footnote{Id.}

Post-TCJA, donations made directly to college athletic departments could diminish due to the law’s repeal of the deduction allowed on contributions made in exchange for athletic seating rights.\footnote{See Tax Cuts and Jobs Act of 2017, Pub. L. No. 115-97, § 13704, 131 Stat. 2054, 2169 (codified at 26 U.S.C. § 170(l)).} Particularly, the TCJA disallows the so-called “80-20 rule,” which permitted boosters an eighty percent deduction on gifts made in exchange for the right to buy tickets or seating at college athletic events.\footnote{26 U.S.C. § 170(l)(B) (2018).} Faced with the unpredictability of how the new law would impact their budgets, some universities suggested that donors make multi-year contributions at the close of 2017 to mitigate immediate budget shortfalls and ensure that priority seating rights remained.\footnote{Darren Rovell, \textit{Tax Reform Nixes Season-Ticket Donation Deductions; Schools Scramble into Action}, \textsc{ESPN} (Dec. 20, 2017), https://www.espn.com/college-football/story/_/id/21827570/tax-reform-bill-removes-deduction-donations-season-tickets-forcing-universities-make-other-plans [https://perma.cc/XK49-9FBY] (further noting that Southern Methodist University, Florida State, and the University of Oklahoma offered the multi-year contribution option to donors).} In addition, the TCJA disallows companies to deduct fifty percent of the purchase price of tickets and stadium
suites to concerts and sporting events for client entertainment purposes. Eliminating these deductions could result in a decreased number of individual booster and corporate donations.

3. Excise Tax on Executive Compensation

The TCJA imposed a new twenty-one percent excise tax (equivalent to the new flat corporate tax rate) on the top five highest paid employees of tax-exempt organizations whose annual compensation exceeds $1 million. The tax likewise applies to excess parachute payments. The purpose of this tax was to better align compensation paid by tax-exempt organizations with for-profit companies. Tax-exempt organizations include corporations organized for educational purposes, as well as organizations fostering “national or international amateur sports competition.” Many public colleges, universities, and collegiate athletic programs could feel its impact.

Today, compensating collegiate coaching staff in excess of $1 million annually is the norm rather than the exception. The average college football coaching salary in 2018 was $2.6 million, with the top overall coaches’ salaries across the United States concentrated in football and men’s basketball. In fact, eight college basketball coaches and thirty-one college coaches...
football coaches were the highest paid public employees in their state in 2017.242 Lucrative coaching contracts propel the financial and successful endeavors of many college athletic departments, resulting in continued scholarly debates surrounding the arms race in collegiate coaching salaries.243 At the close of 2017, the Joint Committee on Taxation estimated that over a ten-year period this new excise tax will provide $1.8 billion in revenue to the federal government.244

Despite initial concerns regarding this provision, on December 31, 2018 the IRS issued interim guidance, establishing that the twenty-one percent tax "will not apply to salaries of more than $1 million paid to coaches of sports teams at state colleges and universities that abandon 501(c)(3) [tax-exempt] status but remain exempt from federal income tax as state instrumentalities."245 Per this guidance, public colleges and universities can rely on the doctrine of implied sovereign immunity for tax-exempt status.246 The guidance clarified that institutions receiving a determination letter from the IRS recognizing their tax exempt status can proactively revoke such status, thus exempting them from the excise tax.247 However, a loophole was inadvertently created by Congress when drafting this provision, as public universities need not fall within the parameters of 501(c)(3) to enjoy tax-exempt status.248 Instead, as compared to private institutions, public colleges and universities are considered long-arms of the state and are therefore exempt from federal taxation.249 While public institutions

---

244 See Joint Comm. on Taxation, Estimated Budget Effects of the Conference Agreement for H.R. 1, the "Tax Cuts and Jobs Act"—Fiscal Years 2018–2027, JCX-67-17 (2017) (for fiscal years 2018 to 2027).
247 Reid & Dimopoulos, supra note 245.
248 See id. (noting that Congress inadvertently created a “gap” which allows state colleges and universities not to pay the tax).
may currently rely on this interim guidance and avoid the twenty-one percent tax on coaching salaries, the Treasury anticipates promulgating forthcoming proposed regulations that could ultimately close this loophole.

4. Excise Tax on Private Institutions’ Endowments

The TCJA imposed a 1.4 percent excise tax on the investment income of certain private colleges and universities. This so-called “Harvard tax” could impact numerous private institutions that will be required to pay millions of dollars on previously untaxed funds. While not a direct tax on college athletic programs, the forfeiture of institutional revenue could impact future funding of more than 8000 student-athletes who compete in Division I athletics in Ivy League athletic programs. In conjunction with the $1.8 billion that the federal government is expected to derive from the twenty-one percent excise tax on compensation, this endowment tax is also estimated to produce $1.8 billion in federal revenue.

---

from gross income “(1) income derived from any public utility or the exercise of any essential governmental function and accruing to a State or any political subdivision thereof, or the District of Columbia; or (2) income accruing to the government of any possession of the United States, or any political subdivision thereof”.


251 Tax Cuts and Jobs Act of 2017 § 13701, 131 Stat. 2054, 2167 (codified at 26 U.S.C. § 4968). The tax applies to private institutions that enrolled at least 500 students during the preceding tax year, have more than fifty percent of its tuition-paying students in the United States, and at least $500,000 in aggregate fair market value of assets per student. Id. § 4968(b)(1)(A)–(D).


5. Unrelated Business Income Tax

As noted, public colleges and universities are generally considered tax-exempt entities. \(^{255}\) However, any revenue derived from unrelated business income (UBI) is taxable at a rate of twenty-one percent. \(^{256}\) UBI is income earned by a tax-exempt organization from a trade or business that is not substantially related to the organization’s exempted purpose. \(^{257}\) Activities that may be considered UBI include “advertising and exclusive provider arrangements, sports management agreements, facility rentals, arenas, food service, golf courses, hotels, recreation centers and programs, parking lots, commercial research, and bookstores.” \(^{258}\)

In addition, effective January 1, 2018, tax-exempt organizations must now include amounts paid or incurred for qualified transportation fringe benefits, including any parking facilities used for parking or for on-premises athletic facilities. \(^{259}\) Further, the TCJA requires that UBI be calculated on a per-business basis instead of a per-entity basis. \(^{260}\) Essentially, the new law disallows tax-exempt organizations from using net operating losses from one unrelated business activity to offset income earned from another of its unrelated business activities. One study determined that the new separate reporting requirement will annually diminish about $12,000 in funds from each affected nonprofit’s mission. \(^{261}\)

B. State Tax Frameworks for Sports Gambling

State expenditures on higher education remain at a historic low. \(^{262}\) In addition, the TCJA will negatively impact funding to collegiate athletic programs. There is no doubt that Congress purposefully targeted higher education and college athletics in an effort to redirect revenue to the federal

---

\(^{255}\) See 26 U.S.C. § 512 (2018); see also INTERNAL REVENUE SERV., PUBLICATION 598: TAX ON UNRELATED BUSINESS INCOME OF EXEMPT ORGANIZATIONS 2 (Mar. 2012) (documenting that UBI is taxed at regular corporate tax rates).

\(^{256}\) See 26 U.S.C. § 512; see also INTERNAL REVENUE SERV., supra note 255, at 2 (documenting that UBI is taxed at regular corporate tax rates).


\(^{260}\) See id. § 13702, 131 Stat. at 2168 (codified at 26 U.S.C. § 512(a)).


\(^{262}\) MITCHELL ET AL., supra note 136, at 1.
However, with colleges and universities facing continued tuition increases, faculty cuts, and course reductions, state funding remains a critical component to help supplement—and perhaps now even substitute—lost internal revenue. Expanding funding models to address these financial pressures should be a strategic goal of state legislatures. Murphy could serve as the catalyst for generating new revenue back into colleges and universities.

Post-Murphy, fourteen states and the District of Columbia permit sports betting and are collecting tax revenue, while more than thirty states have introduced some form of pending legislation. The tax rates imposed on sports gambling vary across jurisdictions, from a low of 6.75 percent to a high of fifty-one percent. Similarly, the types of transactions upon which these levies are imposed differ across state lines. To better understand the national landscape surrounding the taxation of sports gambling, subsection (i) provides a chart identifying the current status of legislation, and subsection (ii) offers a more thorough discussion of the evolving measures taken by states to tax sports gambling in the United States.


264 See MITCHELL ET AL., supra note 136, at 1 (identifying specific negative impacts of decreased funding following the Great Recession).

265 While both New York and New Mexico have very limited sports betting offerings currently, given that the offerings are not statewide, we have excluded them from the table. See Adam Candee, Without Mobile, NY Sports Betting Finally Launches Upstate at Rivers, LEGAL SPORTS REP., https://www.legalsportsreport.com/34442/ny-sports-betting-launch/ [https://perma.cc/8XXF-QWYJ] (last updated July 17, 2019); Holden, supra note 105. New York, by virtue of a 2013 law, allows in-person sports wagering at a small number of upstate casinos. See Candee, supra note 265. Whereas, by virtue of specific language in tribal compacts, a few New Mexico tribal casinos are able to offer sports wagering. See Holden, supra note 105.


267 RICHARD C. AUXIER, TAX POLICY CTR., STATES LEARN TO BET ON SPORTS: THE PROSPECTS AND LIMITATIONS OF TAXING LEGAL SPORTS GAMBLING 5 (May 2019).
1. Taxing Sports Gambling in the United States\textsuperscript{268}

<table>
<thead>
<tr>
<th>State</th>
<th>Legislation Specifications</th>
<th>Rate</th>
<th>Statute/Bill</th>
<th>Date Passed/Proposed</th>
<th>Projected Revenues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legalized Sports Betting</td>
<td></td>
<td>13%\textsuperscript{270}</td>
<td>Ark. Const. Amend. 100</td>
<td>Passed 11/6/2018 272</td>
<td>$31,700,000\textsuperscript{273}</td>
</tr>
<tr>
<td>Arkansas\textsuperscript{269}</td>
<td>≤ $150,000,000 net casino gaming receipts</td>
<td>&gt; $150,000,000 net casino gaming receipts</td>
<td>20%\textsuperscript{271}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorado\textsuperscript{274}</td>
<td>N/A</td>
<td>10%\textsuperscript{275}</td>
<td>HB 1327</td>
<td>Engrossed on 04/24/2019 276</td>
<td>$104,000,000</td>
</tr>
</tbody>
</table>

\textsuperscript{268} This subsection provides a chart summarizing the current status of states’ sports gambling tax legislation as of March 1, 2020. As this is a quickly evolving area of law, readers should be aware that after March 1, 2020, information provided within this chart may have changed.

\textsuperscript{269} ARK. CONST. amend. 100, § 5.

\textsuperscript{270} Id.

\textsuperscript{271} Id.


\textsuperscript{273} We used the Oxford Economics estimate for Arkansas and cut it in half, as the model Arkansas adopted does not make sports betting as widely available as the Oxford Economics estimates. See OXFORD ECON., ECONOMIC IMPACT OF LEGALIZED SPORTS BETTING 36 (May 2017); see also Evin Demirel, Don’t Bet on It: Impact in Arkansas of Legalized Sports Gambling Unclear, TALK BUS. & POL. (Aug. 13, 2018), https://talkbusiness.net/2018/08/dont-bet-on-it-impact-in-arkansas-of-legalized-sports-gambling-unclear/ [https://perma.cc/A4TN-ZZ68].


\textsuperscript{275} Id. § 44-30-1508(1).


\textsuperscript{277} We are extrapolating projected $10.4 million of tax revenue at a ten percent tax rate. See Kim MacCormack, Colorado Online Sports Betting, ROTOGRINDERS, https://roto grinders.com/sports-betting/colorado-online-sports-betting [https://perma.cc/QGF4-UTNK].
<table>
<thead>
<tr>
<th>State</th>
<th>Revenue Sharing Model</th>
<th>Online Launch</th>
<th>Revenue Share</th>
<th>Bill Number</th>
<th>Passed Date</th>
<th>Potential Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delaware</td>
<td>Onsite (online launch pending)</td>
<td>50%</td>
<td>HB 100</td>
<td>05/12/2009</td>
<td>$15,997,230.86</td>
<td></td>
</tr>
<tr>
<td>Indiana</td>
<td>On site and online</td>
<td>9.5%</td>
<td>HB 1015</td>
<td>05/08/2019</td>
<td>$122,763,300</td>
<td></td>
</tr>
<tr>
<td>Iowa</td>
<td>On site and online</td>
<td>6.75%</td>
<td>SF 617</td>
<td>05/13/2019</td>
<td>$49,128,081</td>
<td></td>
</tr>
</tbody>
</table>

---

279 Id. § 22.
280 Id.
282 We extrapolated Delaware’s potential revenue based on its first seven months of operation.
284 Id. at ch. 5, § 12.
285 Id. at ch. 10, § 1.
287 We extrapolated Indiana’s potential revenue from the first three months of operation by multiplying the results by four. *See USA Legal Sports Betting States, Sites & Revenue 2020*, ACTION RUSH, https://actionrush.com/sports-betting/ [https://perma.cc/PVV3-DBAU].
289 Id. § 8.
290 Id. § 15.
292 We extrapolated Iowa’s potential revenue from their first four months of operation by multiplying by three. *See USA Legal Sports Betting States, Sites & Revenue 2020*, supra note 287.
<table>
<thead>
<tr>
<th>State</th>
<th>Mobile and online</th>
<th>Revenue Share (%)</th>
<th>Bill Number</th>
<th>Passage Date</th>
<th>Revenue (2019)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michigan</td>
<td>Mobile and online</td>
<td>8.4%</td>
<td>HB 4311</td>
<td>12/31/2019</td>
<td>$500,000,000</td>
</tr>
<tr>
<td>Mississippi</td>
<td>On site only</td>
<td>8%</td>
<td>HB 967</td>
<td>03/13/2017</td>
<td>$100,800,000</td>
</tr>
<tr>
<td>Nevada</td>
<td>On site and online</td>
<td>6.75%</td>
<td>NRS § 463 et seq.</td>
<td>Legalized 1949</td>
<td>$301,048,000</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>Reserving right to set revenue-sharing agreements individually with licensees</td>
<td>N/A</td>
<td>HB 480-FN</td>
<td>07/16/2019</td>
<td>$15,000,000</td>
</tr>
</tbody>
</table>

---

295 Michigan Gaming Control and Revenue Act § 432.212(16).
299 Id. § 11(3).
300 Miss. CODE. ANN. §§ 75-76-177(1)(a)–(c) (2013).
302 We extrapolated Mississippi’s revenues out based on its March 2018 revenue of $8.4 million, but recognize this may result in a slightly higher projection as a result of the NCAA basketball tournament typically being one of the most popular sports wagering events, which occurs during March. See Jeff Amy, Mississippi Casinos Win More as Sports Betting Extends, ASSOCIATED PRESS (Apr. 30, 2019), https://www.apnews.com/1277d8e3c048fd28a459aa8567c1508 [https://perma.cc/ZYU8-3NWT].
304 Id.
305 Id. § 463.370(1)(c).
306 We used Nevada’s 2018 revenue, as it was the only state offering sports wagering for the entire year. Brian Pempus, Nevada Smashes Sports Betting Revenue Record with $300 Million Year, US BETS (Jan. 31, 2019), https://www.usbets.com/nevada-2018-sports-betting-revenue/ [https://perma.cc/U8TQ-H4LY].
308 Id. § 287-I:3.
309 Id.
310 We extrapolated $15 million from New Hampshire projecting $7.5 million of revenue from fifty percent revenue sharing agreement. See Lisa Kashinsky, Chris Sununu Says N.H. Sports Betting Wagers Net $15.8 Million in One Month, BOS. HERALD (Jan. 28,
<table>
<thead>
<tr>
<th>State</th>
<th>Mode</th>
<th>Revenue</th>
<th>Tax Rate</th>
<th>Bill</th>
<th>Passed Date</th>
<th>Revenue Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Jersey</td>
<td>Land based Casino</td>
<td>$199,176,986.40</td>
<td>8.5%</td>
<td>AB 4111</td>
<td>06/11/2018</td>
<td>$199,176,986.40</td>
</tr>
<tr>
<td></td>
<td>Online by Racetrack</td>
<td></td>
<td>13%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>14.25%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oregon</td>
<td>Online app</td>
<td>N/A</td>
<td>N/A</td>
<td>App</td>
<td>launched 10/16/2019</td>
<td>To date, Oregon has lost $5.3 million since launching the app</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>On site and online</td>
<td>$808,400,000</td>
<td>34%</td>
<td>HB 271</td>
<td>10/30/2017</td>
<td>$808,400,000</td>
</tr>
</tbody>
</table>

---


312 Id. § 7.
313 Id.
314 Id.
315 Id.
316 We extrapolated New Jersey’s projected revenue based on its year-to-date revenue total as of the end of May 2019. See New Jersey Sports Betting Revenue, supra note 115.
321 Id. § 13C11(A)(1)(I).
322 Id. § 13C62(A).
324 To estimate Pennsylvania’s projected revenue, we used a 2017 Oxford Economics report based on convenient availability of the product, because as of the time of writing, Pennsylvania had not yet launched mobile wagering, though it was slated to during Summer 2019. Though Oxford Economics used a lower tax rate than Pennsylvania actually
Rhode Island\textsuperscript{325} | Revenue sharing model between state, IGT (partner company), and casinos | On site only\textsuperscript{326} | 51% to the state, 32% to the partner, and 17% to the casino | S 0037A | Passed 03/25/2019 | $11,500,000\textsuperscript{329} |

Tennessee\textsuperscript{330} | Online\textsuperscript{331} | 20\%\textsuperscript{332} | HB 1 | Passed 06/04/2019 | $29,750,000\textsuperscript{334} |

West Virginia\textsuperscript{335} | On site and online\textsuperscript{336} | 10\%\textsuperscript{337} | SB 415 | Passed 04/23/2018 | $23,864,696.60 |

implemented, we have used the Oxford Economics projections for simplicity. See Oxford Econ., supra note 273, at 39.
\textsuperscript{326} Id. § 42-61.2-5(a).
\textsuperscript{327} Id. § 42-61.2-5(a)(1)-(3).
\textsuperscript{328} Rhode Island Senate Bill 37, LEGISCAN, https://legiscan.com/RI/text/S0037/id/1851946 [https://perma.cc/5R6P-RG64].
\textsuperscript{331} Id.
\textsuperscript{332} Id. § 4-51-304(b).
\textsuperscript{335} S.B. 415, 2018 Leg., Reg. Sess. (W.V. 2018).
\textsuperscript{336} Id. § 29-22D-15(a)-(c).
\textsuperscript{337} Id. § 29-22D-16(a).
\textsuperscript{339} We extrapolated West Virginia revenue based on revenue earned through the week of June 29, 2019. See West Virginia Sports Betting Revenue, PLAYWV!, https://www.playwv.com/revenue/ [https://perma.cc/AS9T-SYYN].
Washington, D.C. 2020

<table>
<thead>
<tr>
<th>State</th>
<th>Legislation Introduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Washington</td>
<td>Tax on private gross sports wagering revenue</td>
</tr>
<tr>
<td></td>
<td>On site and online                                                                ----------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>10% 342 CB B22-0944                                                                ----------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Passed 01/23/2019 343</td>
</tr>
<tr>
<td></td>
<td>$92,000,000 344</td>
</tr>
</tbody>
</table>

### Legislation Introduced

<table>
<thead>
<tr>
<th>State</th>
<th>Legislation Introduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>N/A 10% HB 315 Introduced 04/02/2019; died in committee 347</td>
</tr>
<tr>
<td>Alaska</td>
<td>Establishing Alaska Lottery Corporation that would be permitted to institute video and sports betting 349</td>
</tr>
<tr>
<td></td>
<td>N/A HB 246 SB 188</td>
</tr>
<tr>
<td></td>
<td>Introduced 02/12/2020 350</td>
</tr>
<tr>
<td></td>
<td>Introduced 2/12/2020 351</td>
</tr>
<tr>
<td>Arizona</td>
<td>Allowing Native American tribes exclusive</td>
</tr>
<tr>
<td></td>
<td>6.75% SB 1158</td>
</tr>
<tr>
<td></td>
<td>Introduced 01/23/2019; died in chamber 355</td>
</tr>
</tbody>
</table>

---

341 Id. § 311(a)(1).
342 Id. § 315(a)(2).
344 Martin Austermuhle, Sports Betting Is Coming to D.C., and Here’s Everything You Need to Know About It, WAMU (Dec. 18, 2018), https://wamu.org/story/18/12/18/everything-you-need-to-know-about-sports-bettings-likely-arrival-to-d-c [https://perma.cc/PUA8-GZ7A](projection based over a four-year period).
346 Id. § 16(a).
354 Id. § 42-5073(B)(7).
<table>
<thead>
<tr>
<th>State</th>
<th>Rights to Sports Gambling</th>
<th>Proposed Action</th>
<th>Introduced Date/Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>N/A</td>
<td>Proposing a constitutional amendment that would remove a constitutional prohibition on certain types of gambling</td>
<td>Introduced 06/27/2019</td>
</tr>
<tr>
<td>Connecticut</td>
<td>N/A</td>
<td>N/A</td>
<td>Introduced 03/07/2019; died in committee</td>
</tr>
<tr>
<td>Florida</td>
<td>N/A</td>
<td>9.89%</td>
<td>Introduced 01/14/2020</td>
</tr>
<tr>
<td>Georgia</td>
<td>N/A</td>
<td>15%</td>
<td>Introduced 01/14/2020</td>
</tr>
<tr>
<td>Hawaii</td>
<td>N/A</td>
<td>10%</td>
<td>Introduced 01/24/2019; died in committee</td>
</tr>
</tbody>
</table>

353. Id. § 5-1201(A).
357. Id.
359. Id. § 1 (g)(7)(i).
362. Fla. S.B. 972, § 547.009.
366. Id. § 57-27-133(a).
369. Id. § 1.
<table>
<thead>
<tr>
<th>State</th>
<th>Contains short provision title only</th>
<th>Tax Rate</th>
<th>Bill/Proposed Legislation</th>
<th>Date Introduced/Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illinois</td>
<td>N/A</td>
<td>6.75%</td>
<td>HB 3315</td>
<td>Introduced 02/15/2019</td>
</tr>
<tr>
<td>Iowa</td>
<td>N/A</td>
<td>6.75%</td>
<td>SSB 1080</td>
<td>Subcommittee Meeting 02/06/2019</td>
</tr>
<tr>
<td>Kansas</td>
<td>Integrity fee on aggregate amount wagered, Tax</td>
<td>0.25%, 6.75%</td>
<td>HB 2068</td>
<td>Introduced 01/24/2019</td>
</tr>
<tr>
<td>Kentucky</td>
<td>Placed at licensed track/ professional sports venue, Placed online, via smart phone, or other off-site technology</td>
<td>10.25%, 14.25%</td>
<td>HB 175</td>
<td>Introduced 02/05/2019</td>
</tr>
<tr>
<td>Louisiana</td>
<td>N/A</td>
<td>13%</td>
<td>HB 587</td>
<td>Introduced 04/15/2019; died in chamber</td>
</tr>
</tbody>
</table>

372 Id. § 1.  
375 Id. § 13(4).  
378 Id. § 6(a).  
379 Id. § 16(H)(19)(i).  
382 Id. § 19(2)(a).  
383 Id. § 19(2)(b).  
386 Id. § 610(A).  
<table>
<thead>
<tr>
<th>State</th>
<th>Bill Number</th>
<th>Fee Amount</th>
<th>Rate</th>
<th>Introduced Date, Fate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maine</td>
<td>N/A</td>
<td>25%&lt;sup&gt;389&lt;/sup&gt;</td>
<td></td>
<td>LD 1348, Introduced 03/21/2019; failed 05/14/2019&lt;sup&gt;390&lt;/sup&gt;</td>
</tr>
<tr>
<td>Maryland</td>
<td>License fee $300,000&lt;sup&gt;392&lt;/sup&gt;, Annual renewal $50,000&lt;sup&gt;393&lt;/sup&gt;</td>
<td>20%&lt;sup&gt;394&lt;/sup&gt;</td>
<td>HB 1132</td>
<td>Introduced 02/08/2019&lt;sup&gt;395&lt;/sup&gt;</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>N/A</td>
<td>12.50%&lt;sup&gt;397&lt;/sup&gt;</td>
<td></td>
<td>S 201</td>
</tr>
<tr>
<td>Minnesota</td>
<td>N/A</td>
<td>6.75%&lt;sup&gt;400&lt;/sup&gt;</td>
<td></td>
<td>SF 1894</td>
</tr>
<tr>
<td>Missouri</td>
<td>N/A</td>
<td>6.75%&lt;sup&gt;403&lt;/sup&gt;</td>
<td></td>
<td>SB 222</td>
</tr>
<tr>
<td>Montana</td>
<td>N/A</td>
<td>8.50%&lt;sup&gt;406&lt;/sup&gt;</td>
<td></td>
<td>SB 330</td>
</tr>
<tr>
<td>Nebraska</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td>LB 990</td>
</tr>
<tr>
<td>New Mexico</td>
<td>Sports betting only permitted at two tribal-</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<sup>389</sup> Id. § 1085(1).  
<sup>392</sup> Id. § 9-121(C)(3).  
<sup>393</sup> Id. § 9-121(C)(5).  
<sup>394</sup> Id. § 9-121(D)(1).  
<sup>395</sup> See generally id.  
<sup>397</sup> Id. § 2(g).  
<sup>399</sup> S.F. 1894, 91st Leg., Reg. Sess. (Minn. 2019).  
<sup>400</sup> Id. § 297J.05(1).  
<sup>401</sup> See generally id.  
<sup>403</sup> Id. § 313.1020(1).  
<sup>406</sup> Id. § 14(1).  
<sup>408</sup> L.B. 990, 106th Leg., 2d Sess. (Neb. 2020).  
<table>
<thead>
<tr>
<th>State</th>
<th>Type</th>
<th>Wagering</th>
<th>Bill</th>
<th>Introduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York</td>
<td>On-site, Mobile wagering</td>
<td>8.50%</td>
<td>SB 17D</td>
<td>01/09/2019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Carolina</td>
<td>Fantasy contests; initial registration fee</td>
<td>10% of previous year’s net revenue</td>
<td>HB 929</td>
<td>04/16/2019</td>
</tr>
<tr>
<td>North Dakota</td>
<td>Gross proceeds ≤ $1,500,000</td>
<td>1%</td>
<td>HB 1254</td>
<td>01/03/2019; engrossed 03/25/2019</td>
</tr>
<tr>
<td></td>
<td>Gross proceeds &gt; $1,500,000</td>
<td>2.50%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ohio</td>
<td>N/A</td>
<td>10%</td>
<td>HB 194</td>
<td>04/09/2019</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SB 111</td>
<td>03/14/2019</td>
</tr>
</tbody>
</table>

---

412 *Id.* § 1367(8).
413 *Id.*
414 *See generally id.*
416 *Id.* § 18E-603(b)(3).
419 *Id.* § 53-06. 1-12(1)(a).
420 *Id.* § 53-06. 1-12(1)(b).
423 Ohio H.B. 194, § 5753.02.
<table>
<thead>
<tr>
<th>State</th>
<th>Proposed Amendment</th>
<th>Revenue Limits</th>
<th>Tax Rate</th>
<th>Legislative Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Carolina</td>
<td>Proposing a constitutional amendment to allow the legalization of sports betting 428</td>
<td>N/A</td>
<td>6.25%</td>
<td>S.J. Res. 57</td>
</tr>
<tr>
<td>South Dakota</td>
<td>Proposing a state constitutional amendment granting the legislature power to authorize sports betting only in Deadwood, South Dakota 431</td>
<td>N/A</td>
<td>S.J. Res. 501</td>
<td>Enrolled 02/11/2020 432</td>
</tr>
<tr>
<td>Texas</td>
<td>Per wager</td>
<td>≤ $50,000 revenue / $50,000 – $134,000</td>
<td>6.25%</td>
<td>HB 1275</td>
</tr>
<tr>
<td>Vermont</td>
<td></td>
<td></td>
<td>3.50% – 4.50% + $1750</td>
<td>H0484 Introduced 02/28/2019 440</td>
</tr>
</tbody>
</table>

424 Ohio S.B. 111, § 5753.021.
428 Id. at 1.
429 See generally id.
431 Id.
434 Id. at § 2005.108(a).
437 Id. § 1310(a)(1).
438 Id. § 1310(a)(2).
2. Taxing State Gambling Revenue—An Overview

Since the Murphy decision emerged, eight states have taken in almost $8 billion in combined sports wagers.\textsuperscript{452} Legislators across the country have introduced more than 100 bills that could legalize sports betting in the majority of states.\textsuperscript{453} Implementing sound regulations and appropriate tax rates are of significant consideration.\textsuperscript{454} Tapping into this new revenue source could prove

\begin{tabular}{|l|l|l|l|l|}
\hline
State & Revenue & Tax Rate & Bill & Status  \\
\hline
Virginia\textsuperscript{441} & N/A & 15\%\textsuperscript{442} & HB 1638 & Introduced 11/20/2018; died in chamber\textsuperscript{443}  \\
\hline
Washington\textsuperscript{444} & Legalizing sports gambling in tribal gaming facilities only\textsuperscript{445} & N/A & HB 1975 & Introduced 02/08/2019\textsuperscript{446}  \\
\hline
Wyoming\textsuperscript{447} & Online only\textsuperscript{448} & 16\%\textsuperscript{449} & HB 225 & Introduced 02/14/2020\textsuperscript{450}  \\
\hline
\textbf{No Current Legislation Pending}\textsuperscript{451} & & & &  \\
\hline
Idaho & & & &  \\
Oklahoma & & & &  \\
Utah & & & &  \\
Wisconsin & & & &  \\
\hline
\end{tabular}

\textsuperscript{439}Id. \textsuperscript{\S} 1310(a)(3).
\textsuperscript{442}Id. \textsuperscript{\S} 58.1-4036(A).
\textsuperscript{445}Id. \textsuperscript{\S} 2(1).
\textsuperscript{446}See generally id.
\textsuperscript{448}Id.
\textsuperscript{449}Id. \textsuperscript{\S} 9-23-104.
\textsuperscript{451}As of the date of this publication.
\textsuperscript{454}Holden, supra note 3, at 375 (discussing the taxation of sports gambling); see also Jennifer Roberts & Greg Gemignani, Who Wore It Better? Federal v. State Government
integral to states’ future budget plans; however, setting tax rates too high could spur illegal gambling and stifle revenues. A To date, the rates imposed in those states that have legalized sports betting vary widely.

Currently, there is no direct federal oversight of sports betting. States can set tax rates without an established cap. Such liberties have resulted in a great variance of rates. Nevada boasts the lowest tax rate at 6.75%, while Rhode Island imposes the highest rate of 51%. Both Rhode Island and Delaware implement revenue sharing programs that require the division of gaming revenue between the state and supporting operators.

New Jersey differentiates tax rates based on the type of wager made. Land-based wagers are taxed at the lowest state gambling rate of 8.5%, while online casino bets are taxed at 13%, and online racetrack wagers at 14.25%. Pennsylvania sports betting is subject to a 34% state tax on land-based wagers and a 2% local tax on sports betting revenue. In Mississippi, where only in-person betting is permitted, the tax includes an 8% state and 4% local tax. West Virginia imposes a 10% tax on both in-person and online wagers. In January 2019, the Washington, D.C. Council gave final approval to fast-track sports betting in the Nation’s Capital. The law, which went into effect on May 3, 2019, imposes a 10% tax on revenue derived from gross sports wagering.

Regulation of Sports Betting, 9 U. NEV. L.V. GAMING L.J. 77, 78–79 (2019) (noting that important issues to consider during legislative deliberations of legalized sports wagering include regulatory committees and tax rates).

455 AUXIER, supra note 267, at 6.
456 See Roberts & Gemignani, supra note 454, at 88.
457 See id.
460 See A UXIER, supra note 267, at 5–6 (discussing sports gaming taxation in the state).
464 S.B. 415, 2018 Leg., Reg. Sess. (W. Va. 2018); see also Rodenberg, supra note 266.
To measure these laws’ impact, it is appropriate to compare projected revenue to actual revenue collected.\textsuperscript{467} Nevada, although the longest running legalized sports betting jurisdiction in the nation, does not publish projected sports betting revenues.\textsuperscript{468} Delaware and New Jersey are the only two states meeting early revenue forecasts.\textsuperscript{469} Pennsylvania—projected to earn $100 million in its first full year of implementation—generated only $2.5 million in its first quarter.\textsuperscript{470} West Virginia projected earnings of $5.5 million in its first year but collected just $862,000 by the third quarter.\textsuperscript{471} Mississippi is expected to capture approximately half of its $5 million projected income during its fiscal year.\textsuperscript{472} Rhode Island, boasting the highest tax rate, fell far short of its projected $23.5 million in revenue.\textsuperscript{473} With three months left in its fiscal year, the state earned just $925,000.\textsuperscript{474}

While actual revenues are lower than first-year projections for most states, the American Gaming Association (AGA) found that of the $7.9 billion wagered across the country since PASPA was overturned, $3 billion derived from new


\textsuperscript{468}AUXIER, supra note 267, at 9.


\textsuperscript{471}AUXIER, supra note 267, at 9 (listing the $5.5 million projection in year one of operations); see also Jennifer McDermott & Geoff Mulvihill, \textit{AP: Most States’ Sports Betting Revenue Misses Estimates}, ASSOCIATED PRESS (Apr. 2, 2019), https://apnews.com/21f9833e917948da36422bb286541b4 [https://perma.cc/FJR3-M2GV] (documenting West Virginia’s partial year tax revenue and noting that a contract dispute led to a shutdown of two sportsbooks and West Virginia’s only betting app at the onset of the NCAA March Madness tournament, one of the largest sports gambling events each year).

\textsuperscript{472}McDermott & Mulvihill, supra note 471.


\textsuperscript{474}Anderson, supra note 473.
markets.\textsuperscript{475} Among the states that have legalized sports betting, $55 million has been raised since May 2018.\textsuperscript{476} Although accounting for only 0.5\% of the state’s entire budget, Nevada’s sports betting revenue set a record high of $44 million in 2018.\textsuperscript{477}

By the close of 2020, sports gambling is expected to be legal in thirty states.\textsuperscript{478} Projected gaming revenue from these states is high.\textsuperscript{479} In New York, a mature sports betting land-based and mobile market is expected to produce over $1 billion in annual revenue.\textsuperscript{480} In Indiana, the annual revenue projection is $12 million.\textsuperscript{481} The estimated annual sports betting revenue range in Arizona is in between $20 and $33 million,\textsuperscript{482} while Montana approximates generating $3.7 million in the first full year of operation.\textsuperscript{483}

Actual sports betting revenue, though variable, will have direct and indirect economic impacts.\textsuperscript{484} Employment opportunities will increase, labor income in the form of wages, salaries, and tips will grow, and the combined fiscal impact consisting of state, local, and federal taxes could total $8.4 billion.\textsuperscript{485}


\textsuperscript{476} \textit{Id.}


\textsuperscript{479} See, e.g., AUXIER, supra note 267, at 9. Gaming revenue is defined as “the money the sportsbook—a term for where bets are placed, whether in a casino, racetrack, bar, or online—collects from wagers after paying out successful bets but before it pays its expenses and taxes.” \textit{Id.} at 3.


\textsuperscript{481} Tom Davies, \textit{Indiana Sports Betting, New Casinos Won’t Mean Windfall}, ASSOCIATED PRESS (Apr. 9, 2019), https://apnews.com/44b760668d942e3a2a9526c45ac75a [https://perma.cc/26UH-G73F].


\textsuperscript{484} OXFORD ECON., supra note 273, at 4–5.

\textsuperscript{485} \textit{Id.}
Establishing revenue-maximizing tax rates that the market can bear is key to securing the benefits of this new revenue stream.\textsuperscript{486}

Whether actual revenue meets, exceeds, or falls short of projections, there exists a strong historical connection between gambling and education.\textsuperscript{487} As previously noted, of the states with legalized sports betting in place, only Washington, D.C., Nevada, and Rhode Island direct some portion of earnings for educational purposes.\textsuperscript{488} The remaining jurisdictions direct sports gambling revenue to the general funds, without specifically earmarking any of such revenue for educational purposes.\textsuperscript{489} However, many of these states provide no restrictions on college sports betting. Specifically, there are no constraints for wagers made on college sporting events in Mississippi, Nevada, New Mexico, Pennsylvania, or West Virginia.\textsuperscript{490} Delaware permits general betting on college sports, but disallows bets made on in-state teams.\textsuperscript{491} Rhode Island bans any form of college sports betting on in-state college teams, while New Jersey allows wagers made on multi-site college tournaments played outside the state.\textsuperscript{492} Washington, D.C.’s legislation does not mention college sports betting at all.\textsuperscript{493}

If legalized wagering in any given jurisdiction includes college sporting events, perhaps colleges and universities are justified in receiving a piece of the pie. Well-crafted sports betting legislation may provide opportunities for increased funding to higher education at levels not seen in decades.\textsuperscript{494} The remainder of this Article proposes that the emerging sports betting arena is well-situated to redirect a portion of state-earned gambling tax revenue to one of the biggest drivers of the market—higher education.


\textsuperscript{487}See Kirk, supra note 18.

\textsuperscript{488}Dorson, supra note 13.

\textsuperscript{489}Id.

\textsuperscript{490}Ben Nuckols, NCAA Can’t Keep Tournament Games Away from Legal Gambling, ASSOCIATED PRESS (Mar. 19, 2019), https://www.apnews.com/66e15b3a43ef49619e57467ceeda0b8c [https://perma.cc/B699-TQPJ].


\textsuperscript{492}Nuckols, supra note 490.

\textsuperscript{493}Id.

\textsuperscript{494}The Oxford Economics group estimated that a 0.25% tax on the total amount wagered (the same amount as the federal gambling excise tax) under a base tax rate model of ten percent and a regulatory model that made sports gambling widely available (a measure they call convenient) could generate $718 million dollars annually. See OXFORD ECON., supra note 273, at 5.
V. DIVIDING THE SPORTS BETTING PIE

Perhaps one of the most ubiquitous aspects of legalizing sports betting at the state level has been the search for a piece of the sports betting pie. The economics of a sports betting transaction would suggest that there is a fair-sized pie for everyone to receive a reasonably sized slice. For instance, if a sportsbook sees a total amount of $500 million wagered in a month, and retains five percent, the sportsbook would hold $25 million in that month, out of which the company must pay their operating expenses and taxes, leaving a small margin of profit for the company. Sportsbook profitability is essential to creating a viable regulated market. In order to create room for sportsbooks to function, operating taxes, licensing fees, and other expenses must be kept reasonable.

One aspect that has drawn significant criticism has been a demand by certain professional sports leagues for a royalty of sorts, which they termed an “integrity fee.” The “integrity fee,” or “royalty,” is effectively a private tax created by the NBA and MLB to extract a portion of gambling operator revenue for themselves. Professional sports leagues lack a viable basis for claiming entitlement to gambling fees because they are private, profit-maximizing organizations that receive substantial public subsidies. In contrast, most collegiate sports programs run at a loss, relying heavily on allocated and generated revenue from the institutions within which they are housed and public charitable contributions. As such, finding an optimal approach that would allow colleges and universities—which internally fund the very sports now enticing legalized wagering across the United States—to recoup some of the

497 A five percent “hold” is fairly typical of a sportsbook. See Jill R. Dorson, MS Sportsbooks Take in $7.7M in Handle in August and Much Bigger Numbers Ahead, SPORTSHANDLE (Sept. 20, 2018), https://sportshandle.com/ms-sportsbooks-take-in-6-2m-in-handle-but-not-all-were-open/ [https://perma.cc/EC95-DGJS].
499 See Sauer, supra note 495.
500 Id.
502 Holden, supra note 23.
504 Berry, supra note 189, at 210.
revenue generated is imperative. To examine these issues, this Part proceeds as follows: Section A discusses integrity fees, Section B examines integrity issues within college sports, and Section C proposes the implementation of a gambling interchange fee on wagers made on college sports.

A. Integrity Fees

The first appearance of the so-called integrity fee emerged when a bill was introduced in the Indiana legislature on January 9, 2018, roughly five months before the Supreme Court ruled PASPA unconstitutional. The bill required, “A sports wagering operator shall remit to a sports governing body . . . an integrity fee of one percent (1%) of the amount wagered on the sports governing body’s sporting events . . . at least once per calendar quarter.” The one percent fee transitions into an amount closer to twenty percent of a sportsbook’s profit when accounting for a five percent hold. A twenty percent reduction in profit would make Indiana’s market virtually untenable for a sportsbook operator. Indiana Representative Alan Morrison, who introduced the bill, obliged sports league representatives who requested inclusion of the integrity fee language.

Within the month, NBA Executive Vice President Dan Spillane testified in a New York hearing that sports leagues should be compensated by the one percent fee because they are experiencing increased risk associated with legalized wagering. On February 12, 2018, MLB joined the NBA in seeking an integrity fee, making the request via lobbyists in a variety of states considering legalizing sports pre-Murphy. Interest in integrity fees from the sports leagues has continued, though the NBA and MLB later rebranded the phrase as a “royalty,” suggesting that the leagues were entitled to a fee on an intellectual property theory based on the idea that the leagues put on sporting events, which are the subject being wagered upon.

505 Holden & Schuster, supra note 99, at 37.
506 Id.
507 Id.
508 See id.
509 Integrity Fees—What Are They and Why Are They So Controversial?, SPORTS HANDLE, https://sportshandle.com/integrity-fees/ [https://perma.cc/AL5C-6YLC].
510 Id. The sports leagues, including the NBA, have been vague on increased expenses they suggest they will experience with legalized sports betting, but it is primarily thought that the leagues expect to increase the amount they spend on their security divisions. Id.
511 Id.
512 Though largely beyond the scope of this Article, the leagues’ theory that they possess an intellectual property right in the games themselves, or in the scores and newsworthy results associated with the games, is misguided. See C.B.C. Distribution & Mktg., Inc. v. Major League Baseball Advanced Media, 505 F.3d 818, 821 (8th Cir. 2007); see also Nat’l Basketball Ass’n v. Motorola, Inc., 105 F.3d 841, 853 (2d Cir. 1997); Daniels v. FanDuel, Inc., 109 N.E.3d 390, 393 (Ind. 2018); Marc Edelman, Lack of Integrity? Rebutting the Myth
The quest for integrity fees by professional sports leagues continued through the end of 2018, with the amount sought having been reduced to 0.25%. However, the decreased fee appeared no more palatable as no state has thus far passed a bill granting a fee. While a number of states introduced legislation with integrity fees, no state has awarded any sports league a fee. Instead, a number of leagues have shifted their requests to a mandate that sportsbooks use official data, provided by sports league contractors. Such requests have received a generally cool reception from lawmakers, except in Illinois and Tennessee where bills passed requiring that sportsbooks use league-approved data for certain types of wagers. The NBA and MLB have primarily driven the search for compensation from legal sports betting operators, but most professional leagues have spoken on the issue. The NCAA, however, an organization of amateur athletes, has remained relatively silent and steadfast in its opposition to legalized gambling with or without an integrity fee.

B. Integrity of Collegiate Sports?

While the term integrity fee may have been a marketing ploy created by the NBA, colleges may be most vulnerable to gambling-related threats as student-athletes are unpaid and thus potentially more susceptible to bribery. Several institutions have inquired via lobbyists about the possibility of receiving an integrity fee associated with sports betting; but the NCAA has remained opposed to any expansion of sports betting whether legal or illegal. Despite the opposition, NCAA President, Mark Emmert, stated: “Sports wagering is going to have a dramatic impact on everything we do in college sports. It’s going to threaten the integrity of college sports in many ways unless we are willing to


\[\text{Integrity Fees—What Are They and Why Are They So Controversial?, supra note 509.}\]

\[\text{Id.}\]

\[\text{Holden & Schuster, supra note 99, at 72 (arguing that there is no legal or commercial justification requiring the use of official data, so the mandates effectively serve the same purpose for the leagues as an integrity fee); Adam Candee, Allow Us to Volunteer: The Tennessee Sports Betting Bill Is Really Bad, LEGAL SPORTS REP. (May 9, 2019), https://www.legalsportsreport.com/32035/tennessee-sports-betting-official-data/ [https://perma.cc/852M-ZQG6].}\]

\[\text{Integrity Fees—What Are They and Why Are They So Controversial?, supra note 509.}\]


\[\text{There is no indication that a legal regulated market would increase match-fixing incidents. Indeed, regulated markets are conducive to identifying corruption in illegal markets. See Holden & Rodenberg, supra note 11, at 473.}\]

\[\text{Purdum, supra note 517.}\]
act boldly and strongly." While the NCAA has opposed the expansion, it has endorsed the federal regulatory model as a lesser of two evils.

The NCAA is unique in regards to other sports leagues, who have been vying for a share of sports betting revenue. The NCAA is a member organization composed of public and private educational institutions that generally offer athletics on a nonprofit basis, as opposed to the profit-maximization model of the professional sports leagues. The NCAA has long sought to educate athletes about the risks of associating with gamblers. The availability of widespread legal sports wagering may however create new temptations for college athletes, but a legalized and regulated market controlling who can place wagers offers protections not available in illegal markets. As a result, it has been suggested that the NCAA embrace the expansion of legal sports betting and seek viable shields to protect athletes.

Indeed, the NCAA has begun building safeguards, including formalizing relationships to monitor the integrity of NCAA events. In addition, individual colleges and universities are communicating with legislators about sports wagering, with Marshall University and West Virginia University purportedly requesting integrity fee-type payments alongside professional sports leagues in West Virginia. Other schools have taken a different approach. Penn State University and the University of Pittsburgh, as well as the University of Arkansas, have sought restrictions on wagering on in-state schools, a practice that is antithetical to protecting integrity. Forcing bettors to wager on local teams outside of the state is likely to drive those persons to illegal sportsbooks, which do not comply with state regulatory requirements and may be less likely to report irregularities, thus creating a threat to those in-state teams. Most colleges and universities, however, have been largely quiet in seeking

---


523 See id.

524 Id.

525 See id.

526 Id.

527 Rodenberg, supra note 8.

528 Purdum, supra note 517.

529 Rodenberg, supra note 8.

530 Id.

531 Holden, supra note 123. The practice of exempting in-state teams may also violate the dormant commerce clause to the extent that sports betting is found to implicate interstate commerce. Id.
compensation from legal sports betting.\textsuperscript{532} They may well be best positioned of all the sports leagues to do so; by virtue of being predominately state-funded institutions, it may be appropriate that a percentage of wagering revenue be earmarked to these schools in a way that is distinct from the allocation of a tax benefiting private organizations.\textsuperscript{533}

C. Gambling Interchange Fees

As discussed, the unyielding needs of colleges and universities to source funds is a growing problem.\textsuperscript{534} Legal, regulated sports wagering could provide higher education an opportunity to recoup some of the funding cuts resulting from the Great Recession and the TCJA. Sports wagering provides a unique opportunity for states to legislate a funding mechanism geared towards universities without implementing new taxes or diverting funds from other beneficiaries.

A gambling interchange fee (GIF), earmarked to benefit state colleges and universities, would provide institutions with additional revenue to account for expenses tied to monitoring sports-related integrity concerns and supplement budgets that have long been the victims of state funding cuts.\textsuperscript{535} A GIF is a transaction fee charged on every wager made to a sportsbook on a national college sporting event.\textsuperscript{536} In particular, the GIF would be comprised of (1) a set percentage fee attached to all wagers made on college sporting events with a state-registered sportsbook, and (2) an increased fee attached to wagers that may pose a greater risk to the integrity of college athletics. Increased integrity risks include certain types of in-play events that could lend themselves to manipulation (such as a missed penalty kick in a soccer game), without necessarily having a significant impact on the overall outcome of the game.\textsuperscript{537}

For example, suppose the State of Mississippi imposes a two percent GIF on all wagers made at registered Mississippi sportsbooks on NCAA sporting events. In addition, assume Mississippi imposes a 10-cent flat fee on wagers made on certain in-play events, including college football field goal attempts. X places a $100 bet at a Mississippi casino on Ole Miss winning the upcoming

\begin{footnotes}
\item[532] See Rodenberg, \textit{supra} note 8.
\item[533] See generally Holden & Schuster, \textit{supra} note 99.
\item[535] See Frank Kehl, \textit{What Are Interchange Fees for Credit Card Processing?}, MERCHANT MAVERICK, https://www.merchantmaverick.com/what-are-interchange-fees/ [https://perma.cc/ZB9M-7EPG] (last updated Aug. 20, 2019). Interchange fees are common in the banking industry and represent the amount that an issuing bank and credit card company receive each time a card is used by a customer to make a purchase. \textit{Id.}
\item[536] See \textit{id.}
\end{footnotes}
football rivalry game against Mississippi State University (MSU), and a $50 wager on MSU’s kicker missing four attempted field goals during that game. The total GIF charged on $150 of X’s wagers is $3.10.\(^{538}\)

Data shows that in the first three months of 2019, 64% of all bets placed in Mississippi were tied to college sports.\(^{539}\) If Mississippi’s total gambling revenue during that period was $25,200,000, $16,128,000 of that total was linked directly to college sports betting. With a 2% GIF attached to wagers made on collegiate athletic events, Mississippi could have remitted $322,560 to state colleges and universities in a manner legislatively determined within the state. That base amount could have increased based on the number and types of wagers made on in-play events posing greater risks to college sports integrity.

Although a GIF will not likely replace the total amount of losses suffered by institutions, it may supplement some of the financial harms suffered by institutions as a result of the TCJA.\(^{540}\) The following chart depicts projected revenues derived from college sports betting in those states that have legalized sports-betting, accompanied by projected revenues from GIF based on the application of a sample 2% rate.

---

\(^{538}\) The total GIF amount includes a 2% fee added to each of the two wagers made ($100 and $50), as well as an additional 10-cent fee on the $50 wager made due to the high risk involved in sports manipulation of field goal attempts. In addition, the GIF in this example is applied pre-tax.


<table>
<thead>
<tr>
<th>State</th>
<th>Projected Revenue&lt;sup&gt;541&lt;/sup&gt;</th>
<th>50 Percent of Projected Revenue derived from college sports betting&lt;sup&gt;542&lt;/sup&gt;</th>
<th>Projected Revenue from GIF (2%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Jersey</td>
<td>$199,176,986.40</td>
<td>$99,588,493.20</td>
<td>$1,991,769.86</td>
</tr>
<tr>
<td>Nevada</td>
<td>$301,048,000.00</td>
<td>$150,524,000.00</td>
<td>$3,010,480.00</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>$808,400,000.00</td>
<td>$404,200,000.00</td>
<td>$8,084,000.00</td>
</tr>
<tr>
<td>Mississippi</td>
<td>$100,800,000.00</td>
<td>$50,400,000.00</td>
<td>$1,008,000.00</td>
</tr>
<tr>
<td>West Virginia</td>
<td>$23,864,696.60</td>
<td>$11,932,348.00</td>
<td>$238,646.96</td>
</tr>
<tr>
<td>Delaware</td>
<td>$15,997,230.86</td>
<td>$7,998,615.43</td>
<td>$159,972.30</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>$11,500,000.00</td>
<td>$5,750,000.00</td>
<td>$115,000.00</td>
</tr>
<tr>
<td>Arkansas</td>
<td>$31,700,000.00</td>
<td>$15,850,000.00</td>
<td>$317,000.00</td>
</tr>
<tr>
<td>Indiana</td>
<td>$122,763,300.00</td>
<td>$61,381,650.00</td>
<td>$1,227,633.00</td>
</tr>
<tr>
<td>Iowa</td>
<td>$49,128,081.00</td>
<td>$24,564,040.50</td>
<td>$491,280.81</td>
</tr>
<tr>
<td>Michigan</td>
<td>$500,000,000.00</td>
<td>$250,000,000.00</td>
<td>$5,000,000.00</td>
</tr>
<tr>
<td>Colorado</td>
<td>$104,000,000.00</td>
<td>$52,000,000.00</td>
<td>$1,040,000.00</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>$15,000,000.00</td>
<td>$7,500,000.00</td>
<td>$150,000.00</td>
</tr>
<tr>
<td>Tennessee</td>
<td>$29,750,000.00</td>
<td>$14,875,000.00</td>
<td>$297,500.00</td>
</tr>
<tr>
<td>Oregon</td>
<td>-$5,300,000.00</td>
<td>$0</td>
<td>$0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>$23,131,282.93</td>
</tr>
</tbody>
</table>

Implementing a 2% GIF on college sports betting revenue (as opposed to “handle”)<sup>543</sup> on half of all legalized projected gambling revenue nets over $23 million to be distributed back for use in higher education. Although outside the scope of this Article, the success of the GIF ultimately rests on states individually adopting reasonable regulations that accompany the prescribed fee. Equitable regulations must include moderate yet competitive tax rates that

---

<sup>541</sup> For projected revenue amounts, see <i>supra</i> notes 273–451.

<sup>542</sup> There is little published information regarding the average percentage of sports gambling revenue derived from college sports betting as compared to professional sports betting. Documentation supports that in the first quarter of 2019, 64% of Mississippi’s sports wagering revenue came from college sporting events. Dorson, <i>supra</i> note 539. We surmise that this percentage is higher than expected due to the March Madness series taking place during that period. For purposes of this Article, we have reduced that percentage to 50% for all states. Actual projected revenue derived solely from college sports wagering (as compared to professional sports wagering) in any given state could be greater or less than this proposed projection.

<sup>543</sup> “Handle” refers to the total amount wagered, as opposed to the amount the bookmaker wins, which is a fairer means of estimating the financial success of a bookmaker. <i>See</i> Dustin Gouker, <i>Handle Does Not Equal Revenue in Sports Betting or Daily Fantasy Sports, and Why It Matters to Get It Right</i>, LEGAL SPORTS REP., https://www.legalsportsreport.com/15160/revenue-sports-betting-and-dfs/ [https://perma.cc/33M2-64JM] (last updated Aug. 8, 2019).
minimize revenue siphoning across state borders and disincentive wagers made in the untaxed illegal market, as well as statutory language designating GIF revenue for use by public colleges and universities.\footnote{544}{See generally Holden, supra note 5.}

The GIF is not a panacea, but instead, a measure to alleviate losses that higher education institutions have experienced over the past several decades, and even more recently via the TCJA. Introducing a GIF may invite opponents, particularly as bookmakers seek to maximize profits while lobbying for the lowest possible total tax rate, and private, professional sports leagues similarly seek their own piece of the pie. However, neither of those interests meet the substantial purpose that public colleges and universities fulfill.\footnote{545}{For one viewpoint on the purpose of the public university, see Michael T. Benson & Hal R. Boyd, The Public University: Recalling Higher Education’s Democratic Purpose, 2015 NEA HIGHER EDUC. J. 69, http://www.nea.org/assets/img/HE/i-Benson_Boyd__rev.pdf [https://perma.cc/32HY-HXER].}

Gambling has long been tethered to education in the United States. The Supreme Court’s decision in \textit{Murphy} could serve as the catalyst for reinvigorating funding to higher education, and their affiliated college athletic departments.

VI. CONCLUSION

Both the TCJA and \textit{Murphy v. NCAA} stand to impend the viability of U.S. institutions and college athletic departments.\footnote{546}{See Tax Cuts and Jobs Act of 2017, Pub. L. No. 115-97, 131 Stat. 2054; see also Murphy v. Nat’l Collegiate Athletic Ass’n, 138 S. Ct. 1461 (2018).}
The 2018 federal tax overhaul included a direct hit on collegiate funding, resulting in the continued diminishment of public educational financing.\footnote{547}{See supra Part IV.A.}

In addition, college athletic directors identified \textit{Murphy} as posing a significant threat to the integrity of collegiate sports.\footnote{548}{See supra Part II.D.}

While more than thirty states have introduced bills legalizing sports betting within their jurisdictions, only two have earmarked any of their newfound gambling tax revenue for higher education.\footnote{549}{See supra note 13 and accompanying text.}

States’ de minimis interest in using a portion of sports gambling revenue to secure the economic future of public colleges and universities is surprising for two reasons. First, public institutions and their affiliated athletic departments directly finance—through allocated and generated resources—the very sports that are helping drive the legalized sports betting market.\footnote{550}{See supra notes 186–92 and accompanying text.}

Second, there exists a lengthy historical connection between gambling and education in the United States.\footnote{551}{See supra Part III.A–B.}

Based on these premises, we suggest that higher education be entitled to a piece of the revenue pie.
This Article proposes that college administrators embrace sports wagering as a means of increasing the integrity of college sports, and advocates that public institutions are well-positioned to capitalize on a portion of states’ gambling tax revenue. This Article introduces the implementation of a gambling interchange fee attached to wagers made on college sporting events to redirect funding back into higher education. Unlike integrity fees, which the professional sports leagues are currently lobbying for, a gambling interchange fee would fit well within the historic blueprint of lottery and gambling revenue being reinvested into U.S. public education. A gambling interchange fee attached to college sports wagering would financially benefit state institutions and generate new revenue to account for expenses associated with sports-related integrity concerns. Post-Murphy the time is ripe for states to give back to the institutions and athletic departments helping to influence the newly legalized U.S. gambling market.

552 See supra Part V.B.
553 See supra Part V.B.
554 See supra Part V.C.
Impoverished IP

STEPHANIE PLAMONDON BAIR*

Intellectual Property (IP) scholarship is generally concerned with how innovation policy impacts social welfare by providing appropriate incentives for innovation. But lately, the question of who participates in IP creation—with an eye to distributive justice as well as social welfare more broadly—has been getting more attention. Most scholars writing in this vein acknowledge IP’s shortcomings in achieving proportionate participation and representation across socioeconomic, race, and gender lines. But many argue that in spite of these flaws, IP regimes can advance distributive justice by giving the poor and other members of disadvantaged groups opportunities to accumulate wealth and improve their position in society.

Yet the aspiration some hold out for IP as this particular type of tool for distributive justice is, unfortunately, unlikely to be realized, because it overlooks how poverty impacts creative decision-making. A large and growing body of psychological research shows that poverty changes the decision-making of those experiencing it. This Article argues that in fact, poverty makes it very difficult to think and act in ways that bring about the creative advances meaningful IP participation requires. IP is thus inherently limited as a mechanism for escaping poverty.

Poverty’s impact on creative thinking and action also has wide-ranging implications for innovation theory and policy that reach beyond specific demographic groups. This Article explores how the psychology of poverty intersects with IP, and in doing so, makes four main contributions to the literature. First, it calls into question the feasibility of scholarly calls for IP to act as a mechanism for empowering the poor. Second, it offers an additional, novel explanation for why we see lower levels of IP participation among socioeconomically disadvantaged groups. Third, it argues that IP scholars need to start looking beyond incentives in their quest to optimize socially beneficial innovation. While IP’s dominant utilitarian theory posits that IP provides needed incentives to innovate, what the account fails to consider is the

---

*Associate Professor of Law, BYU Law School. Thanks to Clark Asay, Chris Buccafusco, Dan Burk, Carys Craig, Andrew Gilden, Eric Goldman, Camilla Hrdy, Dmitry Karshedt, Mark McKenna, Lisa Ramsey, Norman Spaulding, Rebecca Tushnet, Saurabh Vishnubhat, Stephen Yelderman, participants in the 2019 Harvard/Yale/Stanford Junior Scholars Forum, participants in the University of Utah faculty colloquium, participants in the 2019 Works in Progress in Intellectual Property Conference at the University of Houston Law Center, participants in the Notre Dame Law School faculty colloquium, participants in the BYU Law School faculty workshop, participants in the 2018 Intellectual Property Scholars Conference at Berkeley Law, and participants in the 2018 BioLawLapalooza Conference at Stanford Law for helpful feedback.
possibility that some otherwise willing participants are unable to respond to these incentives. Finally, in offering policy recommendations, it turns IP scholars’ current thinking about IP and distributive justice on its head. While these scholars argue that IP—a mechanism traditionally used to spur innovation—should be used to achieve distributive justice, this Article proposes that policies more directly aimed at attaining distributive justice will not only be more effective, but should also help promote innovation.

**TABLE OF CONTENTS**

I. **INTRODUCTION** ................................................................. 524
II. **IP, INCENTIVES, AND DISTRIBUTIVE JUSTICE** ....................... 529
III. **THE PSYCHOLOGY OF POVERTY AND CREATIVITY** .................. 533
    A. Poverty and Decision-Making Generally ............................... 538
    B. Poverty and Creativity ..................................................... 539
        1. Exploitative Versus Explorative Decision-Making ............. 541
        2. Habit- Versus Goal-Based Behaviors ............................... 543
        3. Conclusions ............................................................. 544
IV. **THEORETICAL IMPLICATIONS** ............................................. 546
    A. IP and Distributive Justice .............................................. 546
    B. Explaining IP’s Observed Inequalities (and What It Means for Incentive Theory) .............................................. 551
    C. Beyond Incentives ......................................................... 558
V. **POLICY IMPLICATIONS** .......................................................... 560
    A. Promoting Innovation by Tackling Distributive Justice ........ 560
        1. Universal Basic Income ................................................. 561
        2. Access to Health Services ............................................. 562
        3. Other Possibilities .................................................... 563
    B. Thinking Beyond Incentives: Other Innovation-Promoting Policy Levers .............................................. 564
VI. **CONCLUSION** ..................................................................... 566

**I. INTRODUCTION**

If asked why we have Intellectual Property (IP) rights, most people will tell you that IP provides needed incentives—to create, to commercialize, to disclose creations—that are otherwise lacking for public goods. But other accounts of

---

IP exist as well. One of the more intriguing conceptions of IP envisions it as a means of promoting distributive justice. While the manners in which IP might do so are manifold—by ensuring broad access to the products of intellectual labor, for example, or by encouraging technologies that have high social value—one strain of the literature focuses on how IP can promote distributive justice by broadening access to IP rights among poor and disadvantaged groups.5

The view that IP rights can and should serve to advance distributive justice by giving disempowered groups a tool to take control of their futures has intuitive appeal. Economic and social inequality are a matter of growing public and political concern, and it is nice to think that IP could help remedy some of inequality’s more troubling aspects—a kind of meritocratic equalizer to catapult talented individuals otherwise held back by society’s inequities up and away from their disadvantaged backgrounds. Certain high-profile cases seem to bolster the claim: we see an Oprah Winfrey or a Jan Koum, each raised in humble circumstances and each now successful with the help of IP regimes, and are tempted to conclude that IP can achieve these outcomes on a larger scale.

But how realistic is this goal? Is it achievable either in theory or in practice? To answer these questions, we need to understand how the conditions of poverty interact with the prerequisites for IP participation. If the two are incompatible, the case for IP as a tool of distributive justice is weakened, because those we hope will benefit from IP will be limited in their ability to do so. Moreover, the precise nature of the interaction will tell us what, if anything, we can do about it. While some incompatibilities might be remedied by changing IP doctrine or procedure, others may be more existential in nature.

One such potential point of incompatibility arises at the behavioral level. Those who study the psychology of poverty know that poverty changes

---

2 See, e.g., ROBERT P. MERGES, JUSTIFYING INTELLECTUAL PROPERTY 36–38 (2011) (describing the Lockean account of IP, according to which intellectual production gives rise to natural rights in the products of intellectual labor); Fisher, supra note 1, at 170–71 (describing the personhood theory of IP, according to which IP rights are granted to protect the personhood with which creators imbue their creations).


5 See id. at 347–52.


7 See Justin Hughes & Robert P. Merges, Copyright and Distributive Justice, 92 NOTRE DAME L. REV. 513, 555 & n.182 (2016) (arguing that Oprah Winfrey’s repeated achievement as one of the wealthiest African Americans in the United States is attributable in part to her participation in the “copyright-related industries”).
decision-making. But the literature has given little attention to how this altered thinking impacts creativity. When several strains of empirical evidence are considered together, however, one thing becomes clear: circumstances common in the lives of the poor impact decision-making in ways that make it more difficult to achieve the creative advances meaningful IP participation requires.

For instance, when people have experienced the world as harsh and unfair and are subject to high levels of environmental stress—conditions disproportionately present in poor populations—they are more likely to employ so-called exploitative decision-making strategies, where they “stick to what they know,” versus explorative decision-making strategies, where they consider options about which they have less information—strategies definitionally required for creative thinking. Another line of study shows that sleep deprivation—a condition empirically linked to low income—leads to the privileging of habit-directed (and therefore, again, definitionally noncreative) behaviors over goal-directed behaviors. And while in some cases adversity may trigger and promote creative thinking, in general, the particular flavor of adversity accompanying poverty seems to suppress, rather than incite, the human drive to create.

This incompatibility between poverty and creative thinking appears to limit the potential for IP to serve as a meaningful tool of distributive justice. If those living in poverty tend to over-rely on decision-making strategies inconsistent with creativity, then, on the whole, they will be less able than their financially better-off counterparts to take advantage of the social and economic benefits IP participation has to offer. This is fundamentally inconsistent with the hope that IP can help remedy social and economic inequality. If anything, we should expect IP to exacerbate existing inequalities, because those already at a financial advantage are those best positioned to participate in IP and reap its benefits.

Moreover, unlike some of the other potential incompatibilities between poverty and IP—for example financial barriers, or doctrines that favor large


9 See infra Part II.


12 See infra Part II.

13 See Lee, supra note 4, at 348–51 (discussing how the United States Patent and Trademark Office (USPTO) seeks to reduce these barriers in various ways).
moneymed actors over poorly financed individual creators— the insight that poverty interferes with creative thinking does indeed pose an existential threat to IP’s viability as a tool of distributive justice. While IP doctrines can be changed, and financial barriers can be addressed by legal aid programs or changes in procedural requirements, the behavioral incompatibility between poverty and IP participation is different in kind. No matter how much we increase access to IP rights or tweak individual IP doctrines to make them fairer to low-income groups, the problem won’t be remedied, because it’s a problem of under-creation caused by poverty rather than a problem that is caused or can be fixed by IP itself.

This may seem like a dire conclusion, but in fact it is quite helpful, because it leads to several fundamental insights about the nature of IP and our current attempts to use it as a tool, not only for distributive justice, but also for achieving optimal levels of innovation.

First, the relationship between poverty and creativity suggests that no matter how noble our aspirations for IP, it is inherently ill-suited as a tool of distributive justice, insofar as that term is used to describe the appropriation of IP by poor populations to enhance their financial and social position. Rather than lament this fact, we can use the insight to direct our efforts at achieving distributive justice along more fruitful paths.

Second, the relationship between poverty and creativity helps explain observed inequalities within the IP system. Though empirical data are sparse, the numbers that do exist suggest that children from low-income groups are much less likely to go on to participate in IP in their lifetimes than children from higher income groups. This is a real problem, not just from a distributive justice perspective, but also for the dominant view that IP is about maximizing socially beneficial innovation. If certain groups are innovating at lower levels, it raises the possibility that we are falling short—either in absolute number or in the types of innovations being produced—of optimality. And though a host of reasons for why those from low-income brackets participate less in IP are easily called to mind, the behavioral literature contributes a plausible and consequential explanation that might otherwise be overlooked in the rush to point to more obvious structural and cultural culprits.

Third, the relationship between poverty and creativity prompts reflection on IP scholarship’s treatment of incentives. In keeping with IP’s dominant

15 See id. at 109–31 (suggesting various changes to judge-made patent doctrines).
16 See Lee, supra note 4, at 348–51 (discussing how fee reductions and pro bono technical assistance to small entities and unrepresented inventors could increase access to the patent system).
17 Alex Bell et al., Who Becomes an Inventor in America? The Importance of Exposure to Innovation, 134 Q.J. ECON. 647, 649 (2019) (finding that high socioeconomic status at birth predicts later probability of obtaining a patent).
narrative, the scholarship is very much focused on questions relating to creative incentives.\(^1\) For example, do current IP doctrines provide the right balance of incentives?\(^2\) Are there other, better, ways beyond IP—like grants, or tax breaks, or even social norms\(^3\)—to provide individuals with these incentives? But what rarely gets mentioned is the possibility that some people—perhaps even large, identifiable groups of people, as here—may be unable to respond to whatever incentives we offer them, no matter how wonderful those incentives might be. IP and innovation scholars might thus begin thinking much more broadly and holistically about what we can do to promote innovation, rather than assuming that getting incentives right is the end of the story.

Along these lines, the relationship between poverty and creativity gives us some clues about how we might begin to do this. For example, if poverty does indeed interfere with creative thinking, then we might expect that interventions designed to combat poverty will have a beneficial effect on creativity and innovation. More broadly, it opens the door to considering nontraditional innovation-promoting policy levers—i.e., those that don’t fit the standard incentive formulation of financial reward as innovation carrot. Especially given recent work that has called into question any simple relationship between financial incentives and innovation output,\(^4\) these nontraditional models might take on a larger role in future conversations about innovation policy levers. Examples of such policy levers might include universal basic income programs, policies designed to increase access to health care, or open-ended artistic or scientific grants that aren’t contingent on completing any particular project. Fortuitously, many of these policy levers would also help realize the distributive justice outcomes some scholars wish to achieve through IP. Thus, rather than using IP as a tool of distributive justice, this Article suggests that we would be better off tackling distributive justice head-on—and that doing so should also help achieve the innovation-promoting aims traditionally relegated to IP.

The Article proceeds in three parts. Part I recounts the traditional incentive formulation of IP, along with more recent calls to justify and conceive of IP systems as valuable mechanisms for achieving distributive justice. To critically evaluate the claim that IP can indeed function as an effective tool of distributive

---


\(^2\) Id. at 67.

\(^3\) See, e.g., Dotan Oliar & Christopher Sprigman, *There’s No Free Laugh (Anymore): The Emergence of Intellectual Property Norms and the Transformation of Stand-Up Comedy*, 94 Va. L. Rev. 1787, 1791, 1832 (2008) (raising the possibility “that social norms can provide incentives to create”).

\(^4\) See, e.g., Glynn Lunney, *Copyright’s Excess: Money and Music in the US Recording Industry* 193 (2018) (concluding from an empirical study that more copyright “did not lead to more and better music,” as the incentives for the innovation paradigm would predict, because of satisficing behaviors on the part of top artists); Eric E. Johnson, *Intellectual Property and the Incentive Fallacy*, 39 Fla. St. U. L. Rev. 623, 678–79 (2012) (arguing that psychology research undermines the traditional incentive account of IP).
IMPOVERISHED IP

justice by giving poor populations opportunities to accumulate wealth, Part II identifies and examines a potential point of incompatibility between poverty and IP participation based on the behavioral literature. The basic insight is that poverty changes decision-making. Part II extends the insight to the realm of creativity and concludes that conditions of poverty negatively impact creative decision-making and action.

Part III explores the theoretical and practical implications of that insight. The first, glaring implication is that IP may not be well-suited at all as a tool of distributive justice, for the simple reason that those we hope will benefit from IP are arguably the least able to do so. Second, the insight about poverty’s impact on creative thinking has explanatory power, as scholars struggle to make sense of empirical data showing reduced IP participation among disadvantaged groups. Third, the relationship between poverty and creativity should prompt scholars to think beyond the traditional incentives-for-innovation model when considering how best to maximize socially beneficial creation. Creators do not live in a theoretical vacuum, and ostensibly extraneous circumstances in their lives could impact, for good or ill, their ability to respond to even the most well-designed innovation incentives. Finally, the relationship between poverty and creativity suggests that policy interventions designed to ease poverty’s burden could also have beneficial effects on innovation. Continuing in this vein, it also suggests that additional policy levers with less-than-direct ties to the act of innovation may nevertheless be quite effective at promoting innovation—in part by reaching groups for whom the traditional incentives model has not proved effective.

II. IP, INCENTIVES, AND DISTRIBUTIVE JUSTICE

The most common theoretical account of IP focuses on economic incentives. According to this conceptualization, IP rights exist to encourage people to create things that might not otherwise come into being.22 Because intellectual creations—like books, software, or a new technology—are easy to copy, it can be difficult for the creator to recoup her investment in the creative process.23 Once released to the public, the price a creator can charge for her creation quickly falls as the market gets flooded with copies by free-riders.24 If the creator cannot recover her investment, let alone improve her financial position, the thinking is that she will choose not to create at all.25

---

22 See Lemley, supra note 1, at 129–30 (explaining the incentive theory of patents).
23 See Fisher, supra note 1, at 169.
25 See Olson, supra note 24, at 183.
IP arguably addresses this dilemma by offering creators time-limited exclusive rights over their creations. Creators can thus charge higher prices for their inventions, allowing inventors to profit and offering the needed financial incentives to create. Though granting IP rights imposes costs on society through the deadweight losses resulting from reduced access and competition, the thought is—if IP is working as it should—that these costs should be outweighed by the social benefits of heightened creation and the economic growth that attends it. Because of deadweight losses, however, the efficiency narrative requires that we weigh the costs and benefits and grant IP rights only when and to the extent necessary to reap the social benefits of creation.

This economic-incentives-based rationale dominates the IP literature; but it is by no means the only purported justification for IP rights. The Lockean account, for example, asserts that creators acquire natural rights in their creations by virtue of the labor they have invested in the creative process. Under this view, IP is simply an appropriate recognition of those rights. And the personality account argues that IP rights are necessary to protect ongoing personality interests that creators retain in their works.

26 Fisher, supra note 1, at 169.
27 Id.
28 See T. Randolph Beard et al., Quantifying the Cost of Substandard Patents: Some Preliminary Evidence, 12 YALE J.L. & TECH. 240, 241 (2010); Olson, supra note 24, at 195.
30 See Olson, supra note 24, at 195.
31 See Fisher, supra note 1, at 168–72.
32 A number of additional utilitarian variations on the traditional incentives-to-create account have also been proposed, including accounts wherein IP incentives are required to encourage creators to disclose or commercialize their inventions. See, e.g., Michael Abramowicz, The Danger of Underdeveloped Patent Prospects, 92 CORNELL L. REV. 1065, 1073–76 (2007) (discussing commercialization theory, according to which patents provide incentives for inventors to commercialize their inventions); Christopher A. Cotropia & Mark A. Lemley, Copying in Patent Law, 87 N.C. L. REV. 1421, 1432 (2009) (same); Jeanne C. Fromer, Patent Disclosure, 94 IOWA L. REV. 539, 548 (2009) (same); Mark A. Lemley, The Myth of the Sole Inventor, 110 MICH. L. REV. 709, 745 (2012) (same); Roberto Mazzoleni & Richard R. Nelson, Economic Theories About the Benefits and Costs of Patents, 32 J. ECON. ISSUES 1031, 1038 (1998) (discussing the utilitarian disclosure theory, according to which patents provide incentives for inventors to disclose their inventions rather than keep them secret).
35 E.g., MERGES, supra note 2, at 68–100; Fromer, supra note 34, at 1753; Justin Hughes, The Philosophy of Intellectual Property, 77 GEO. L.J. 287, 329–30 (1988); Margaret
But perhaps the most interesting alternative account of IP comes from those who have conceptualized IP as a valuable tool for promoting human flourishing and distributive justice. Distributive justice theorists owe a debt to Rawls’s philosophical writings on fairness and social justice, which they often employ to address questions of intellectual production and distribution.

A notable work in this vein is Madhavi Sunder’s 2006 article IP. In it, Sunder laments the dearth of “‘giant-sized’ intellectual property theories capable of accommodating the full range of human values implicit in intellectual production.” She argues for a cultural account of IP that recognizes “not just efficiency, but a number of incommensurable values: from the right to health, to the freedom to create, to democracy, equality, and distributive justice.”

Since IP was published, a number of scholars have responded to the call to consider IP through the lens of distributive justice. And while much of this

Jane Radin, Property and Personhood, 34 STAN. L. REV. 957, 971–78 (1982); Rosenblatt, supra note 24, at 457.

36 In calling distributive justice an “alternative” account I don’t mean to suggest that its proponents eschew economic efficiency as a proper frame through which to conduct IP analyses. In fact, most of the scholars cited here explicitly acknowledge the importance of considering efficiency at some level. Where most depart from traditional accounts is in their desire to also consider a range of other values and interests—including distributive justice—while conducting these analyses.

37 See generally JOHN RAWLS, A THEORY OF JUSTICE (1971).

38 E.g., ROSENBLATT, supra note 1, at 325 (arguing that U.S. patent law “already possesses numerous ‘distributive mechanisms,’” and “sketch[ing] the contours


40 Sunder, supra note 3, at 260.

41 Id. at 313; see also Margaret Chon, Intellectual Property and the Development Divide, 27 CARDozo L. REV. 2821, 2823 (2006) (“[I]ntellectual property should include a substantive equality principle, measuring its welfare-generating outcomes not only by economic growth but also by distributional effects.”); Fisher, supra note 1, at 172 (describing a theoretical approach to IP—what Fisher terms the “social planning” approach—wherein scholars take the view “that property rights in general—and intellectual-property rights in particular—can and should be shaped so as to help foster the achievement of a just and attractive culture”); Molly Shaffer Van Houweling, Distributive Values in Copyright, 83 TEX. L. REV. 1535, 1540 (2005) (highlighting the “distributive aspects” of copyright).

literature argues for limiting or otherwise altering intellectual property regimes in various ways to improve public access to intellectual products produced by the well-funded.\textsuperscript{43} many distributive justice theorists also evince a desire to see “historically disempowered individuals”\textsuperscript{44} who are artists, inventors, and creators wielding IP themselves—as a “tool for recognition and redistribution, development, and human rights.”\textsuperscript{45} In other words, because IP participation confers financial and social benefits, the thought is that socioeconomically disadvantaged populations can (and should) make use of IP to improve their lives.\textsuperscript{46}

In fact, some scholars argue that existing IP regimes already achieve these ends to some extent.\textsuperscript{47} In their article Copyright and Distributive Justice, for example, Justin Hughes and Robert Merges contend that “copyright in its current form is a powerful tool to empower creative individuals”\textsuperscript{48}—especially socioeconomically disadvantaged individuals—economically and socially. To support their argument, they point to a 2009 Forbes listing of the wealthiest African Americans in the United States, all of them “self-made.” Based on this list, they conclude that it is copyright—specifically, the “copyright industries” of “music, film, television, broadcast professional sports, and publishing” that can be thanked “for the accumulation of the most substantial African-American fortunes.”\textsuperscript{49}

of a distributive agenda for domestic patent law” going forward); Rosenblatt, supra note 24, at 458–59 (describing the “distributive justice” theoretical approach to IP analysis).


\textsuperscript{44} Sunder, supra note 3, at 263.

\textsuperscript{45} Id. at 264.

\textsuperscript{46} See, e.g., Lee, supra note 4, at 364 (“[O]btaining patents on their own inventions can empower marginalized and low-income communities, thus achieving more equitable distribution of the fruits of intellectual property protection.”). The interest of some scholars in so-called “traditional knowledge” protection is a related strain of this literature. One possible justification for protecting traditional knowledge is that IP rights in such knowledge can protect indigenous populations and local communities from exploitation and help empower them with property rights of their own. See Justin Hughes, Traditional Knowledge, Cultural Expression, and the Siren’s Call of Property, 49 SAN DIEGO L. REV. 1215, 1256–61 (2012) (describing a distributive justification for traditional knowledge protection).

\textsuperscript{47} See, e.g., Lee, supra note 4, at 347–52 (describing various mechanisms by which the U.S. patent system attempts to facilitate IP participation, including fee reductions and technical assistance for small entities and unrepresented inventors); Shaffer Van Houweling, supra note 41, at 1540–41 (discussing how copyright helps subsidize creators who would otherwise not be able to afford the investments necessary to engage in creative labor).

\textsuperscript{48} Hughes & Merges, supra note 7, at 516.

\textsuperscript{49} Id. at 554.
Of course Hughes, Merges, and others writing in this area acknowledge that IP as presently constituted is not the perfect vehicle for empowering the poor, remedying inequality, and achieving distributive justice.\textsuperscript{50} Nor do they argue that IP should be the only, or even the primary, tool for achieving these ends.\textsuperscript{51} But the desire to see IP as one such tool—and to improve its ability to act in this capacity through various doctrinal and procedural adjustments\textsuperscript{52}—remains.

III. THE PSYCHOLOGY OF POVERTY AND CREATIVITY

The idea that IP can and should serve as a mechanism for promoting distributive justice by giving the poor opportunities to reap the social and economic benefits it confers is both intriguing and attractive. Creation and inventorship enjoy a certain mystique in our culture,\textsuperscript{53} as does the conception of the United States as a meritocracy.\textsuperscript{54} The story of IP as a tool of distributive justice taps into both of these romantic notions. It is tantalizing to envision otherwise disadvantaged populations rising up on the wings of their creative talent, boosted by IP regimes that give them the mechanism for translating this talent into money and status. The stirring language employed by distributive justice IP scholars adds rhetorical depth to this vision. Most people sympathetic to social justice generally, for example, would be moved by the thought of the previously-disempowered “assert[ing] themselves as intellectual property subjects, controlling rights in cultural creations, and reject[ing] earlier categorization as law’s objects.”\textsuperscript{55}

But is this distributive agenda for IP realistic? In its defense, the list cited by Hughes and Merges does suggest that at least some subset of the socioeconomically disadvantaged population—specifically, several African Americans from humble financial backgrounds—has benefitted financially and

\textsuperscript{50} See id. at 552 (“Copyright would be a better tool with some of the policy alternatives we describe later in this Article.”); Lee, supra note 4, at 367 (arguing that “Congress, courts, and agencies—particularly the USPTO—should promote a more robust vision of the patent system’s distributive capabilities”).

\textsuperscript{51} Hughes & Merges, supra note 7, at 552 (“[C]opyright is a meager tool for distributive justice compared to basic social reforms that are possible—principally a significant (multi-decade) strengthening and equalization of K-12 public education. Indeed, we believe that copyright would be a better distributive tool if coupled with a substantially strengthened educational system for minorities.”).

\textsuperscript{52} See id. at 573–75 (discussing ways in which the copyright system could better fulfill a distributive agenda); Lee, supra note 4, at 367–74 (discussing ways in which the patent system could better fulfill a distributive agenda).

\textsuperscript{53} See, e.g., Lemley, supra note 32, at 710 (“Any elementary school student can recite a number of canonical American invention stories.”).


\textsuperscript{55} Sunder, supra note 3, at 275.
socially from their careers in IP-heavy industries.\textsuperscript{56} Take Oprah Winfrey (number 1 on the Forbes list cited by Hughes and Merges),\textsuperscript{57} for example. Born in rural Mississippi where she spent the first six years of her life with her grandmother,\textsuperscript{58} the remainder of Oprah’s childhood was spent being shuttled between her mother’s home in a Milwaukee ghetto and her father’s home in Nashville, Tennessee.\textsuperscript{59} By any account, it was a difficult and impoverished upbringing.\textsuperscript{60} But that all changed when she was offered a job as a news anchor with CBS’s Nashville affiliate\textsuperscript{61}—her entrance into the “copyright industries.” From there, she steadily gained increasing exposure and opportunity, appearing on her own talk show and in a breakthrough role in Steven Spielberg’s 1982 film \textit{The Color Purple},\textsuperscript{62} and eventually maturing into the multimedia and publishing giant the public knows her as today.\textsuperscript{63}

Additional anecdotes can also be brought to mind, and they are not limited to the Forbes list of wealthy African Americans. J.K. Rowling was raised in middle class circumstances but was living in poverty as a single mother and collecting welfare benefits when she completed her first \textit{Harry Potter} novel.\textsuperscript{64}

\begin{flushright}
\begin{thebibliography}{64}
\bibitem{HughesMerges2005} Hughes & Merges, \textit{supra} note 7, at 552–53.
\bibitem{Hughes2005a} Id. at 552.
\bibitem{Mair1994} \textit{GEORGE MAIR, OPRAH WINFREY: THE REAL STORY} 13–14 (1994); Nelson, \textit{supra} note 58.
\bibitem{Morgan2008} \textit{See, e.g.}, Thomas Morgan, \textit{Troubled Girl’s Evolution into an Oscar Nominee}, \textit{N.Y. TIMES} (Mar. 4, 1986), https://www.nytimes.com/1986/03/04/movies/troubled-girl-s-evolution-into-an-oscar-nominee.html [https://perma.cc/KYE4-WXX6] (describing Oprah’s history of being abused and her mother’s threat to “put Miss Winfrey into a juvenile detention home”); Nelson, \textit{supra} note 58 (describing how Oprah had been sexually abused by “male relatives and family friends,” how she ran away from her Milwaukee home as a child, and the moment she “realized [she] was poor”).
\bibitem{Morgan2010} Nelson, \textit{supra} note 58.
\bibitem{Nelson2010} \textit{Id.}
\bibitem{Rowling2008} J.K. Rowling, \textit{The Fringe Benefits of Failure}, \textit{TED} (June 2008), https://www.ted.com/talks/jk_rowling_the_fringe_benefits_of_failure?language=sp [https://perma.cc/4VA9-J2GM] (stating that at the time she wrote the first \textit{Harry Potter} novel she was as “poor as it is possible to be in modern Britain, without being homeless”). \textit{But see} Ian Parker,
And WhatsApp founder and Ukraine native Jan Koum lived as a teenager on government assistance in California with his mother before developing an interest in computer programming. The mobile messaging company he founded is the owner of multiple patents and was eventually sold to Facebook for $19 billion dollars.

But as Hughes and Merges acknowledge, anecdotes are not data. So how can we determine whether we should expect these anecdotes to generalize? In other words, what is clear from these examples is that certain members of disadvantaged populations have pulled themselves out of poverty by generating IP-protected artistic and technical content. But what remains to be determined is whether a similar trajectory is available to a meaningful segment of the tens of millions of poor in the United States who remain. If we are going to treat IP as a serious mechanism for achieving distributive justice—at least in the sense we’re talking about here—then this should be the goal.

One way to approach the problem is to ask how compatible poverty is with IP participation. Take a hypothetical poor citizen possessed of both the

---

Mugglemarch, NEW YORKER (Sept. 24, 2012), https://www.newyorker.com/magazine/2012/10/01/mugglemarch [https://perma.cc/B39M-MTB2] (acknowledging that at the time she wrote her first novel Rowling was a “broke single mother, in poor accommodations, at a time of high unemployment,” but also noting her relative advantages: “[S]he was a middle-class graduate, poised to start a teaching career, who claimed modest state benefits while she finished a novel, which she partly wrote in an upscale café owned by her sister’s husband.”).


67 Olson, supra note 65.

68 See Hughes & Merges, supra note 7, at 555 (“We should also repeat that we are not proposing that copyright has wealth redistributive impact for African Americans as a whole, as would be needed in an argument about wealth distribution to meet Rawls’s Difference Principle. Our argument is that on the Rawlsian question of ‘conditions of fair equality of opportunity,’ copyright is doing much good whereas many other social structures are not.”).

69 What Is the Current Poverty Rate in the United States?, U. C. DAVIS CTR. FOR POVERTY RES. (Oct. 15, 2018), https://poverty.ucdavis.edu/faq/what-current-poverty-rate-united-states [https://perma.cc/PM3V-VLYA] (stating that the official poverty rate in the United States as of 2017 was 12.3%, or approximately 39.7 million people); see also Steven Pressman, New Data Paint an Unpleasant Picture of Poverty in the US, CONVERSATION (Sept. 12, 2018), https://theconversation.com/new-data-paint-an-unpleasant-picture-of-poverty-in-the-us-101069 [https://perma.cc/89JW-PE4K] (arguing that “things look even worse” than the official poverty rate “if we use what many scholars like myself believe is a better poverty measure”). According to Pressman, the poverty rate is probably “two to four percentage points above the official U.S. measure,” which translates into up to 12.9 million additional people living in poverty. Id.
necessary talent and desire to generate IP-protectable content.\(^{70}\) Can we expect that this citizen will actually be able to generate the content and avail herself of the advantages IP confers? Or are there conditions specific to the circumstance of poverty that are incompatible with this goal? If the latter, then the case for IP as a generalizable mechanism for distributive justice is weakened.

The quick answer is, of course, that there are many incompatibilities between IP participation and poverty, some of which easily spring to mind. Inadequate education,\(^ {71}\) reduced opportunities for professional employment that provides an outlet for creative talent,\(^ {72}\) and a scarcity of time\(^ {73}\) and money might all help explain why a poor person who otherwise has the aptitude and desire might not end up creating something that is both IP-protectable and likely to generate meaningful income. And even if she does manage to innovate in this way, barriers such as unfamiliarity with the IP system,\(^ {74}\) inability to pay the

---

\(^{70}\) When I speak of IP-protectable content in this way, I’m referring to the subset of IP-protectable content that also has the potential to generate income and social status for the creator. Because copyright in particular has a very low bar for what is considered protectable and no formal requirements for acquiring protection, copyrighted content that does not have such potential is generated by most everyone on a regular basis, but does not help advance the goal of distributive justice I’m concerned with here.


\(^{72}\) See Marlene Kim, Problems Facing the Working Poor, in BALANCING ACTS: EASING THE BURDENS AND IMPROVING THE OPTIONS FOR WORKING FAMILIES 49, 53 (Eileen Applebaum ed., 2000) (reporting that the working poor are overrepresented in jobs that involve repetitive tasks like “agricultural work[] and machine operat[ion]”).

\(^{73}\) MARK A. RUNCO & STEVEN R. PRITZKER, 1 ENCYCLOPEDIA CREATIVITY 333 (1999) (“For many creative people, time is the most precious of all resources, without which creative work is simply impossible.”).

\(^{74}\) See, e.g., Bell et al., supra note 17, at 647–48 (finding that exposure during childhood to other inventors makes it more likely that a child will himself go on to become an inventor); K.J. Greene, Copyright, Culture & Black Music: A Legacy of Unequal Protection, 21 HASTINGS COMM. & ENT. L.J. 339, 353–54 (1999) (explaining how the 1909 Copyright Act’s complex registration requirements may have prevented black artists historically from getting protection for their work); Sunder, supra note 3, at 273 (“Problems encountered in protecting the knowledge of the poor [may] turn . . . on the poor’s lack of knowledge of their rights . . . .”)
required fees,\textsuperscript{75} or even bias on the part of IP’s gatekeepers\textsuperscript{76} may prevent her from acquiring or enforcing IP rights in her creation.

These incompatibilities almost certainly exist, as scholars advancing a distributive agenda for IP have acknowledged to varying extents.\textsuperscript{77} But the hope is that they can be addressed—in some cases relatively easily.\textsuperscript{78} Peter Lee, for example, argues that reducing financial barriers to entry could be a productive way to broaden access to the U.S. patent system, and points to steps the United States Patent and Trademark Office (USPTO) has already taken in this direction as a hopeful sign.\textsuperscript{79}

But there is another incompatibility between poverty and IP participation, perhaps even more pernicious than the others because to this point it has largely flown under the radar. This incompatibility arises at the behavioral level. A growing body of work is revealing how poverty influences decision-making—the way people go about their lives and the decisions they make every day. And much of what researchers have found is relevant to creativity. In fact, what their findings suggest is that it might be very difficult, in the psychological sense, for a person living in poverty to engage in the creative process.


\textsuperscript{76}See, e.g., Andrew Gilden, \textit{Raw Materials and the Creative Process}, 104 GEO. L.J. 355, 375–82 (2016) (arguing that courts tend to apply copyright’s fair use doctrine in ways that favor rich and successful artists at the expense of less-well-known creators); K.J. Greene, \textit{Intellectual Property at the Intersection of Race and Gender: Lady Sings the Blues}, 16 AM. U. J. GENDER SOC. POL’Y & L. 365, 371–72 (2008) (arguing that doctrines like the idea/expression dichotomy, the fixation standard, and the originality standard all favor white artists at the expense of black artists and in fact may encourage white artists to appropriate and gain protection over black artists’ work); Greene, \textit{supra} note 74, at 375–83 (1999) (same).

\textsuperscript{77}See, e.g., Hughes & Merges, \textit{supra} note 7, at 556–61 (discussing the various barriers disadvantaged populations face in their struggle to participate in IP); Sunder, \textit{supra} note 3, at 273 (“Problems encountered in protecting the knowledge of the poor [may] turn . . . on the poor’s lack of knowledge of their rights.”).

\textsuperscript{78}See Hughes & Merges, \textit{supra} note 7, at 561 (proposing “ways in which the copyright system could further strengthen wealth distribution to authors”).

A. Poverty and Decision-Making Generally

It is becoming increasingly clear that the circumstance of poverty influences decision-making. Some of the most comprehensive behavioral work on the subject comes from cognitive scientist Eldar Shafir. Shafir has found that the conditions common to poverty, including “tight financial challenges, instability of income and expenses, low savings, no insurance, and several other stressors,” lead to deep changes in decision-making processes, affecting “attention, cognitive resources, and ensuing decisions.”\(^{80}\)

Often, the practical result of these changes is a co-opting of scarce attentional resources to deal with pressing, day-to-day needs.\(^{81}\) Consequently, fewer resources are available for other mental tasks.\(^{82}\) Shafir’s work focuses on how this makes it particularly difficult for those living in poverty to make the decisions necessary to get ahead—for example, when all one’s attention is focused on today’s urgent need to pay rent, the high interest rates on the payday loan required to meet that need may be overlooked.\(^{83}\) But it is clear that the cognitive load poverty imposes affects more than just financial decision-making. In one study, for instance, Shafir and colleagues found that inducing thoughts about finances reduced cognitive performance in a range of tasks—like spatial processing and creative problem solving\(^{84}\)—in poor participants,\(^{85}\) while having no such effect on the more well-off.\(^{86}\)

The neuroscience research on poverty and the brain supports Shafir’s findings, and is perhaps even more concerning. Shafir emphasizes that what his research shows is that it is the circumstance of poverty itself—rather than some other explanation like genetic differences or stress—that causes cognitive impairments.\(^{87}\) Thus, he sanguinely points to studies showing that moving to a better neighborhood can positively impact things like college attendance and earnings.\(^{88}\) But what a large neuroscience literature suggests is that when

\(^{80}\) Eldar Shafir, Decisions in Poverty Contexts, 18 CURRENT OPINION PSYCHOL. 131, 131 (2017).

\(^{81}\) Id. at 132.

\(^{82}\) Id. at 132–33.

\(^{83}\) Id. at 133.

\(^{84}\) Anandi Mani et al., Poverty Impedes Cognitive Function, 341 SCIENCE 976, 977 (2013) (describing the tests they used to measure cognitive functioning, including a spatial task designed to measure “cognitive control” and “the ability to guide thought and action in accordance with internal goals,” and Raven’s Progressive Matrices test, designed to measure “fluid intelligence,” the capacity to think logically and solve problems in novel situations, independent of acquired knowledge”).

\(^{85}\) Id. (describing how they categorized subjects as either “rich” or “poor,” with those they categorized as poor “roughly corresponding to those in the lower quartile or third of the U.S. income distribution”).

\(^{86}\) Id. at 977–78.

\(^{87}\) Id. at 976 (“This suggests a causal, not merely correlational, relationship between poverty and mental function.”).

\(^{88}\) Shafir, supra note 80, at 134.
poverty is experienced during childhood in particular, it can lead to lifelong deficits in cognitive functioning.\textsuperscript{89} These effects are thought to be linked to disrupted brain development trajectories. Brain development is a process that takes place throughout childhood and young adulthood, and is largely complete by the time a person reaches her mid-twenties.\textsuperscript{90} Because a young person’s brain is the most malleable at this time,\textsuperscript{91} the conditions of poverty can disrupt developmental trajectories,\textsuperscript{92} presumably leading to cognitive impairments—including deficits in language ability, memory, and goal-directed behaviors.\textsuperscript{93} And because the window of malleability eventually closes, these brain impairments, if present in early adulthood, can last a lifetime.\textsuperscript{94} Even if a person eventually escapes a poverty situation, the changes to brain structure and function may persist. So while some effects of poverty on decision-making, such as those detailed by Shafir and his team, may be the direct result of current exposure to poverty conditions, and thus may resolve when the conditions are ameliorated, others may flow from earlier, childhood exposure to poverty, and thus may be resistant to environmental improvements.

B. Poverty and Creativity

The findings on poverty, decision-making, and brain development are concerning, and, as commentators have pointed out, have a range of implications for various areas of social and political concern, including public


\textsuperscript{90} See, e.g., Sara B. Johnson et al., \textit{Adolescent Maturity and the Brain: The Promise and Pitfalls of Neuroscience Research in Adolescent Health Policy}, 45 \textit{J. Adolescent Health} 216, 216 (2009) (explaining that certain parts of the brain “may not be fully developed until halfway through the third decade of life”).


\textsuperscript{92} See, e.g., Kimberly G. Noble et al., \textit{Socioeconomic Disparities in Neurocognitive Development in the First Two Years of Life}, 57 \textit{Developmental Psychobiology} 535, 535 (2015).

\textsuperscript{93} E.g., Martha J. Farah et al., \textit{Childhood Poverty: Specific Associations with Neurocognitive Development}, 1110 \textit{Brain Res.} 166, 168 (2006).

\textsuperscript{94} E.g., Shafir, \textit{supra} note 80, at 134.
health, criminal law, education, and more. But to this point, no one has explored how these findings might read on the potential for the poor to engage in creative and innovative pursuits.

It is clear, however, that they are relevant to this question. For example, Shafir and colleagues’ finding that financial concerns in the poor reduce cognitive functioning in areas like creative problem solving is directly relevant to the question of how well and how often we might expect those living in poverty to innovate. And the fact that poverty experienced in childhood can lead to lifelong difficulties with language and goal-directed behaviors is also concerning, as many creative pursuits require facility with these very tasks.

One 2008 study, for example, found that brain activity in the prefrontal cortex, an area integral to creativity and problem solving, was significantly different between seven and twelve year olds from low socioeconomic households and those from higher-status households. The differences were so extreme that researchers remarked that the poor children’s brain activity resembled that of adults who had suffered brain damage.

More recent empirical work in psychology and neuroscience also supports the hypothesis that those living in poverty may find it particularly difficult, psychologically speaking, to engage in creative pursuits. Specifically, the poor may be pressed by their circumstances to employ so-called exploitative and habit-based decision-making strategies that make creativity harder to come by.

---

97 Farah, supra note 89, at 434 (explaining how SES disparities and their effect on the brain “have implications for education policy”).
98 Mani et al., supra note 84, at 977.
99 See, e.g., Christina E. Shalley, Effects of Coaction, Expected Evaluation, and Goal Setting on Creativity and Productivity, 38 ACAD. MGMT. J. 483, 483 (1995) (finding that subjects were most creative when they were able to set their own creativity goals).
100 Mark M. Kishiyama et al., Socioeconomic Disparities Affect Prefrontal Function in Children, 21 J. COGNITIVE NEUROSCIENCE 1106, 1106 (2008); see also Robert Sanders, EEGs Show Brain Differences Between Poor and Rich Kids, U.C. BERKELEY NEWS (Dec. 2, 2008), https://www.berkeley.edu/news/media/releases/2008/12/02_cortex.shtml [https://perma.cc/ZS96-58XY] (reporting on the study) (“In a study recently accepted for publication by the Journal of Cognitive Neuroscience, scientists at UC Berkeley’s Helen Wills Neuroscience Institute and the School of Public Health report that normal 9- and 10-year-olds differing only in socioeconomic status have detectable differences in the response of their prefrontal cortex, the part of the brain that is critical for problem solving and creativity.”).
101 See Kishiyama et al., supra note 100, at 1106 (“We found that prefrontal-dependent electrophysiological measures of attention were reduced in LSES compared to high SES (HSES) children in a pattern similar to that observed in patients with lateral prefrontal cortex (PFC) damage.”).
1. Exploitative Versus Explorative Decision-Making

A subject of study in the decision-making literature is how and when people employ exploitative decision-making strategies versus explorative decision-making strategies, two ways of thinking that use distinct brain architectures.102 The first type can be thought of as a “stick-with-what-you-know” approach, while the second involves exploring new options.103 While each has its place—104—for example, it would make little sense to keep exploring new ways to get from your home to work when you have already found the most efficient route—external circumstances in peoples’ lives might sometimes cause the appropriate balance of strategies to get out of whack, leading to suboptimal decision-making.105

A recent study shows that stress, as measured by the level of the stress hormone cortisol, is one of these external circumstances.106 Specifically, when subjects were exposed to a stressor that increased their cortisol, they tended to over-rely on exploitative decision-making, weighing costs and benefits incorrectly and sticking to known strategies even when exploring new options would have led to better outcomes.107 The authors hypothesize that stress contributes to a perception that one’s environment is “harsh and unfair;” the decision-maker might thus conclude that it makes little difference what approach he employs, as the outcome will likely be negative no matter what.108 Better, then, to conserve resources by using the less taxing exploitation strategy than to take a risk on exploration the decision-maker has little hope will pay off.109

Cortisol, the hormone that mediates this inefficient decision-making strategy, has close ties to poverty. Poverty is stressful, and the bodies of those who experience it produce cortisol in response to this stress.110 A number of studies have documented how high cortisol levels go hand in hand with poverty and how this affects brain development and decision-making in the moment and

104 See Laureiro-Martínez et al., supra note 102, at 319.
105 See Harms, supra note 103, at 10035.
106 Lenow et al., supra note 10, at 5681.
107 Id. at 5685–86.
108 See Harms, supra note 103, at 10035 (reviewing the Lenow study).
109 See id.
over time. Though cortisol serves a legitimate physiological function—it contributes, for example, to the “fight or flight” response, and insufficient levels of cortisol can also lead to decision-making impairments—in general, sustained high levels of cortisol, such as those seen in impoverished populations, tend to have negative effects on health and decision-making. Perceiving your environment as “harsh and unfair”—a contributor to stress and heightened cortisol also has ties to individuals’ experience as members of traditionally disadvantaged race and gender groups. Exposure to race-related stress, for example, leads to sustained heightened stress response in black subjects.

These findings can tell us something about poverty and creativity. When considering what kind of decision-making—explorative or exploitative—is involved in creativity, we would expect the former to play a particularly important role. Creativity, by definition, means making something new. It would be impossible to make something new without engaging in explorative decision-making—the kind of decision-making that involves seeking out new information and options. If the economically disadvantaged, due to the

---

111 See, e.g., Clancy Blair et al., Salivary Cortisol Mediates Effects of Poverty and Parenting on Executive Functions in Early Childhood, 82 CHILD DEV. 1970, 1979 (2011) (finding that heightened cortisol resulting from poverty leads to impaired executive functioning in children); Farah et al., supra note 93, at 169 (finding that poverty leads to impaired neurocognitive development and hypothesizing that heightened cortisol plays a mediating role).


113 Christopher Bergland, How Do Various Cortisol Levels Impact Cognitive Functioning?, PSYCHOL. TODAY (June 17, 2015), https://www.psychologytoday.com/us/blog/the-athletes-way/201506/how-do-various-cortisol-levels-impact-cognitive-functioning [https://perma.cc/E5QS-KUFW] (reporting on a study finding that very high and very low levels of cortisol result from unstable family environments and lead to impaired cognitive functioning).

114 See, e.g., Bruce S. McEwen, Central Effects of Stress Hormones in Health and Disease: Understanding the Protective and Damaging Effects of Stress and Stress Mediators, 583 EUR. J. PHARMACOLOGY 174, 181 (2008); Bergland, supra note 113; Klein, supra note 112 (“Too much cortisol can suppress the immune system, increase blood pressure and sugar, decrease libido, produce acne, contribute to obesity and more.”).

115 See Harms, supra note 103, at 10035.


119 See Harms, supra note 103, at 10035.
stressful nature of their environments, over-rely on exploitative decision-making strategies at the expense of exploration, it will be more difficult for them to engage in truly creative behaviors.

2. Habit- Versus Goal-Based Behaviors

Just as people can choose whether to employ exploitative or explorative decision-making strategies in their daily lives, they can also choose at any moment whether they are going to engage in reflexive, habit-based, decision-making or more reflective, goal-directed decision-making. Each strategy is directed by different neural circuits and comes with a unique suite of costs and benefits. In goal-directed (also known as reflective or model-based) decision-making, a person consciously desires a specific outcome and analyzes various ways to get there, ultimately choosing the path he thinks is most likely to lead to the result he wants. In contrast, during habit-based (also known as reflexive or model-free) decision-making, a person relies heavily on past experience to make quick utility calculations and does not invest as much conscious thought in the nuances of present circumstances and how they might affect outcomes. While goal-directed decision-making has the advantage of being flexible and adaptable to the specific situation at hand, it is also cognitively costly. Habit-directed decision-making, on the other hand, is an efficient heuristic, but may lead to more errors because it is less flexible, and more of a “one-size fits all” approach. Like exploration and exploitation, habit-based and goal-based decision-making each have their place in human action, but they must be appropriately selected and balanced for optimal decision-making.

The choice between a habit- or goal-based strategy in any given instance also has implications for creativity. Because creativity is about making something new, it and a habit-based approach—which by definition involves doing what you’ve done before—are at odds. Instead, a goal-directed approach, which involves mental flexibility, deep cognitive engagement, and a
close conscious attention to one’s environment\textsuperscript{129} is more consistent with creative action.\textsuperscript{130}

Enter now the evidence that individuals suffering the stresses of poverty may over-rely on habit-based decision-making approaches. A recent study shows that sleep deprivation privileges habit-based decision-making strategies over goal-based decision-making strategies.\textsuperscript{131} Those subject to sleep deprivation over-rely on habit, using it to make decisions even when a goal-based approach would be more appropriate.\textsuperscript{132} Sleep deprivation, in turn, has well-known empirical ties to both poverty and race. Those with financial struggles, as well as African Americans and Latinos, are much more likely to experience poor sleep quality than whites and those with higher socioeconomic status.\textsuperscript{133} Stress is thought to play an important role in this interaction, in addition to specific life circumstances like employment, education, and health status.\textsuperscript{134}

These findings also have something to say about poverty and creativity. What they tell us is that the poor are the most likely to suffer sleep deprivation,\textsuperscript{135} which in turn makes them more susceptible to an over-reliance on habit-based action\textsuperscript{136}—a decision-making strategy unlikely to lead to creative thought.\textsuperscript{137}

3. Conclusions

When one considers together the various strains of research studying the psychology and neuroscience of poverty, one inference seems increasingly inescapable: poverty harms the brain, and interferes with decision-making processes, in ways that hinder one’s ability to think and act creatively.

This conclusion might seem strange at first. After all, aren’t many artists and innovators poor? We can all easily call to mind examples. And what about the proverbial “starving artist”? Isn’t he a proverb for a reason?

In the absence of empirical research that directly tackles the link between poverty and creativity, it is impossible to answer these questions conclusively.

\textsuperscript{129} See Dolan & Dayan, supra note 121, at 314.
\textsuperscript{130} See Ford, supra note 128, at 1116–19.
\textsuperscript{131} Chen et al., supra note 11, at 11979.
\textsuperscript{132} Id. at 11990 (“[S]leep-deprived individuals overrelied [sic] on habitual learning at the expense of goal-directed learning.”).
\textsuperscript{134} Patel et al., supra note 133, at 4754.
\textsuperscript{135} Id.
\textsuperscript{136} Chen et al., supra note 11, at 11979.
\textsuperscript{137} See Ford, supra note 128, at 1113.
But until then, a couple of possibilities might help explain the disconnect between the conclusions I draw here and popular intuitions.

First, when one visualizes the proverbial starving artist, one often pictures someone who has chosen poverty as the price for pursuing and perfecting their craft.\(^{138}\) Psychologically speaking, this could be a very different situation from the one encountered by someone who was raised in poverty, or who otherwise cannot choose to not be poor. While to the former group, poverty might be psychologically empowering, a statement of life choices and values,\(^ {139}\) to the latter, poverty may do little more than make it harder to think creatively in the ways already discussed. And in any case, as some commentators have pointed out, the myth of the starving artist may be just that—a myth.\(^ {140}\)

But if this latter point is true, then why can we so easily bring to mind examples of artists and innovators who emerged with their creations from humble backgrounds? This could be due in part to our collective biases. The story of the poor innovator or artist is a compelling one; for that reason it might be repeated often and, through the workings of the availability heuristic,\(^ {141}\) be more easily called to mind than the many, many stories of artists and inventors who come from more privileged backgrounds.\(^ {142}\) For example, while the tale of Oprah’s humble upbringing might be widely known and reported on, fewer people may be aware that Taylor Swift is the child of a financial advisor and marketing executive,\(^ {143}\) or that Bill Gates’s father was a prominent lawyer.


\(^{139}\) See id.


\(^{142}\) See, e.g., Hannah Ellis-Petersen, Middle Class People Dominate Arts, Survey Finds, GUARDIAN (Nov. 23, 2015), https://www.theguardian.com/artanddesign/2015/nov/23/middle-class-people-dominate-arts-survey-finds [https://perma.cc/8HS4-JZLQ] [discussing the results of a survey finding that over three-quarters of respondents who worked in creative industries came from a middle class background, and considering the argument that “performing arts schools had become dominated by those from affluent, privately educated backgrounds”).

while his mother served on the board of directors of a financial holding company. Because of these biases, we may become overly sanguine about the ability of those living in poverty to engage in creative and innovative pursuits.

IV. THEORETICAL IMPLICATIONS

A. IP and Distributive Justice

For over a decade, IP scholars have been calling for distributive justice to take its rightful place among the justifications for and goals of IP regimes. But converting IP into an effective tool of distributive justice—in the sense of facilitating IP acquisition and assertion by members of disadvantaged groups—may be far more difficult to attain on a large scale than previously imagined. There are the previously-recognized incompatibilities between poverty and IP acquisition and assertion, of course: incompatibilities like the financial and social barriers inherent in navigating the IP system that scholars tend to be optimistic about policymakers’ ability to address. But a more existential challenge to the aspirations of IP and distributive justice proponents is highlighted by the behavioral incompatibility described here: the finding that poverty severely limits the ability of those experiencing it to think and act in creative ways.

Poverty’s destructive influence on creative thought and action will make it very difficult for those living in poverty’s throes to arrive at the creative advances meaningful IP protection requires. Patent law, for example, demands as a prerequisite to protection that inventions be novel. As Greg Mandel has noted, the essence of the novelty requirement is “remarkably akin” to the novelty psychologists consider an essential element of creativity. Copyright, too, requires that expressive works be “original” to merit protection. As the Supreme Court put it in Feist Publications v. Rural Telephone Service Co., this is the “bedrock principle of copyright” law; any work seeking protection must

145 See Chon, supra note 41, at 2823; Fisher, supra note 1, at 171; Ghosh, supra note 42, at 859; Lee, supra note 4, at 325; Sunder, supra note 3, at 315; Shaffer Van Houweling, supra note 41, at 1540.
146 See, e.g., supra notes 71–79 and accompanying text (highlighting some of the incompatibilities between poverty and meaningful IP protection).
147 See, e.g., Hughes & Merges, supra note 7, at 573–75 (discussing ways in which the copyright system could better fulfill a distributive agenda); Lee, supra note 4, at 365–74 (discussing ways in which the patent system could better fulfill a distributive agenda).
148 See supra Part II.
150 Mandel, supra note 118, at 2002–03.
thus display “at least some minimal degree of creativity.” If creativity is required for IP protection, that which harms one’s ability to be creative will also harm one’s ability to participate in IP. And poverty harms one’s ability to be creative.

It is true that copyright’s originality requirement sets a notoriously low bar for what counts as creativity. Poverty’s destructive influence on creativity might therefore not systematically interfere with people’s ability to meet this minimal standard. But in this regard, it’s important to remember the claim that IP and distributive justice scholars are making. Not just that the poor can and should acquire and assert IP rights, but that the poor can and should acquire and assert IP rights that will help them financially and socially. In order for an IP-protected expressive work to achieve this end, we might expect that the level of creativity required to generate the work would not only meet the originality bar, but substantially exceed it. For example, while poverty might not seriously interfere with a person’s ability to compose a straightforward email message, take a snapshot with his phone, or doodle on the back of an envelope—all copyrightable activities that in fact do acquire copyright protection at the moment of fixation—it might very well interfere with that person’s ability to compose a more substantial piece of writing or music that has the potential to make him money and improve his social status.

This is not to say that no one who is poor can ever be expected to engage in meaningful creative activity. That is clearly not the case; some of our most compelling works of literature, for example, have come from authors who were extremely poor; we have also already considered the cases of Oprah, J.K. Rowling, and Jan Koum. But what it does suggest is that the average person raised or living in poverty will be less able than the average person of means to convert a similar level of motivation and talent into creative output.

If this is true, the argument that IP can help advance distributive goals where other social structures have failed loses much of its power. Indeed, while some few economically disadvantaged individuals may be able to harness IP to their advantage, we would expect IP, on the whole, to exacerbate existing

---

153 Id. at 345.
154 See supra Part III.
157 See, e.g., JACKSON J. BENSON, THE TRUE ADVENTURES OF JOHN STEINBECK, WRITER 142, 147 (The Viking Press 1984) (describing how Steinbeck for a time collected welfare and subsisted on the fish he caught and the vegetables from his wife’s garden); JOHN FORSTER, THE LIFE OF CHARLES DICKENS 27 (Sterling Publ’g Co. 2011) (describing Dickens’s impoverished late childhood).
158 See supra Part III.
159 See Hughes & Merges, supra note 7, at 555 (“Our argument is that on the Rawlsian question of ‘conditions of fair equality of opportunity’ copyright is doing much good whereas many other social structures are not.”).
inequalities, as those of greater means are disproportionately able to participate in it and reap its benefits.

Moreover, the challenge the psychology of poverty presents to the IP and distributive justice narrative is existential in the sense that it, unlike some of the other incompatibilities between poverty and creativity that have been recognized, cannot easily be remedied within IP itself. While we should certainly continue to address IP-based concerns like financial barriers to entry and doctrines that may make it more difficult for the poor to acquire and assert IP rights, this will not fix the incompatibility raised here. That’s because the problem is one of reduced creation by the poor rather than reduced protection for the poor. Once a poor person creates something, we can make sure IP is structured such that this person can gain and assert rights in her creation as easily as possible. But because IP is premised on creation, it is powerless to help the poor person who does not create—even if that person would be creating but for their poverty. It is difficult to see, then, how IP can serve as a meaningful tool of distributive justice in this sense without the aid of outside interventions aimed at reducing poverty.

The behavioral incompatibility between IP and poverty also highlights some of the dangers and unintended consequences of the IP and distributive justice account. While the narrative is a relatively innocuous and even commendable story of empowerment, below its surface lies an assumption: that the poor, generally speaking, can in fact engage in the creative pursuits IP rewards with the same ease as those of greater means and similar levels of talent and motivation. This false underlying assumption is dangerous because it perpetuates the myth of the meritocracy: that we live in a society where a given level of talent, motivation, and hard work will tend to lead to a given level of financial and social success, regardless of the starting point of the person possessed of those qualities. And though this is certainly not the intention of IP and distributive justice scholars, there is a sense in which their narrative puts the onus on the poor to pull themselves up by their own bootstraps (with the help of IP) and make something of their lives, without fully recognizing the psychological disadvantages they face relative to the rest of the population in their attempts to do so.

This type of thinking holds peril for those who happen to live in poverty as well as those who are lucky enough to not currently face that challenge. For those of us with greater means, the danger is obvious—we could begin to think that members of poor populations who do not create are simply not talented, hard-working, or motivated enough to do so, even though this could be far from the truth. This, in turn, could stymie the political will to enact more promising distributive justice reforms. But it is also a particularly harmful mode of

---

161 See Stewart, supra note 54.
162 See supra Part III.A.
163 For a related argument in a different context, see Will Stancil, The Scandal that Reveals the Fiction of America’s Educational Meritocracy, ATLANTIC (Dec. 19, 2018),
thought for those who do currently find themselves in dire financial straits. One study, for example, found that economically disadvantaged youth who buy into the meritocratic narrative of hard work are more likely to act out in their middle school years, as they “begin to blame themselves for problems they can’t control.”164 Because the empowerment mindset can so easily be distorted into a meritocratic mindset, it runs the risk of adding one more burden to the already heavily-weighted loads of the poor—the burden of falsely believing that they themselves are fully responsible for their failure to realize their creative ambitions.

In this regard, the fact that some poor people actually do successfully manage to take advantage of IP to escape poverty can act as a double-edged sword. On the one hand, as Hughes and Merges point out, these success stories can serve as important role models for others who aspire to similarly harness the power of IP in their own lives.165 On the other hand, it could lead to incorrect, and potentially harmful, generalizations. If we see that some have been able to transcend their disadvantaged circumstances with the help of IP, we might wrongly assume that this is a realistic option for others in like circumstances.166

It seems, then, that a particular argument common to law and distributive justice scholars—that the poor can and should harness IP rights to their individual and collective advantage—while superficially compelling, has some significant limitations. Foremost, it fails to recognize the degree to which poverty is incompatible with creativity. Because of this, the poor, as a group, will be less able to benefit from IP than those of greater means. And because the problem is one of reduced creation rather than reduced protection, nothing we do within the IP system can hope to change this. Given this reality, the rhetoric of empowerment the argument employs is particularly dangerous, because it incorrectly assumes the poor can in fact take advantage of IP, and indirectly puts the blame on them when they don’t.

Importantly, the analysis here focuses on two particular types of IP for which creativity is particularly relevant—copyright and patent. There are, of course, other forms of IP, including trademark protection in particular, for which

---

164 Melinda D. Anderson, Why the Myth of Meritocracy Hurts Kids of Color, ATLANTIC (July 27, 2017), https://www.theatlantic.com/education/archive/2017/07/internalizing-the-myth-of-meritocracy/535035/ (arguments in the context of a prep school scandal, where a prep school that had incredible success in placing low-income students in Ivy League colleges was found to have engaged in fraud) (“Miracle fixes can excuse complacency. . . . Success stories suggest that, even among the poor children of color who face pervasive societal burdens, the truly deserving can prevail in the end. When inequality is defeatable, it stops feeling so much like injustice. For that reason, many people recoil at attempts to depict segregation, discrimination, and poverty as an inescapable trap, even though, for millions of children, they have proved exactly that”).

165 Hughes & Merges, supra note 7, at 560 (citing evidence that same-race role models can “increase[e] self-esteem and expand[] career horizons for young people”).

166 See Stancil, supra note 163.
the distributive justice analysis might differ significantly. Trademark protection is available to signifiers, such as words or symbols,¹⁶⁷ that designate the source of a good or service.¹⁶⁸ Owners of trademarks acquire rights by using their marks in commerce to identify their products,¹⁶⁹ and can prohibit uses of their mark (or similar marks) that might confuse consumers about a product’s source¹⁷⁰ or dilute the uniqueness of the owners’ mark in consumers’ minds.¹⁷¹ It’s generally thought that trademarks do not reward creativity;¹⁷² accordingly, there is no requirement for acquiring trademark protection akin to patent’s nonobviousness and novelty inquiries or copyright’s originality threshold. Because of this, the negative influence poverty exerts on creativity may not significantly interfere with the ability of those living in poverty to take advantage of trademark protection and the economic advantages it bestows. Indeed, one important strain of the IP and distributive justice literature focuses on how geographic indications—a form of trademark protection granted to products, like Rocquefort cheese or Idaho potatoes, that come from a particular geographic source—are a particularly useful way to put valuable IP rights in the hands of the poor.¹⁷³

The discussion here also centers on the effects poverty wields on the individual, and does not speak to how poverty may or may not influence a community’s collective ability to innovate. This is significant because another avenue scholars have proposed for putting IP in the hands of the poor is through the protection of traditional knowledge and cultural expression—inventions like traditional music and stories or therapeutic and horticultural technologies that arise in particular communities and cannot be traced to a single

¹⁶⁷ Though a trademark is often a word or symbol, under modern trademark jurisprudence protection has been afforded to signifiers like smells, see In re Clarke, 17 U.S.P.Q.2d (BNA) 1238, 1990 WL 354572, at *2 (T.T.A.B. 1990) (finding a signature scent on yarn to be eligible for trademark protection), and sounds, see In re Gen. Elec. Broad. Co., 199 U.S.P.Q. (BNA) 560, 1978 WL 21247, at *3 (T.T.A.B. 1978) (finding a series of bells broadcast by a radio station on the half hour could potentially be eligible for trademark protection); see also Neal R. Platt, Is a Trademark Owner’s Right to Use Its Mark Protected by the First Amendment?, 11 HOFSTRA L. REV. 1261, 1262–63 (1983) (“A trademark can assume almost any imaginable form, as long as it is applied to the goods and is perceptible to the buying public.”).


¹⁶⁹ Id. (setting forth the “use in commerce” requirement).


¹⁷² See Dastar Corp. v. Twentieth Century Fox Film Corp., 539 U.S. 23, 37 (2003) (stating that trademark law was “not designed to protect originality or creativity”). But see generally Jeanne C. Fromer, The Role of Creativity in Trademark Law, 86 NOTRE DAME L. REV. 1885 (2011) (arguing that trademark law does in fact encourage creativity in various ways).

¹⁷³ Sunder, supra note 3, at 301 (referring to geographic indications “as the poor people’s intellectual property rights”).
individual. Further research would be needed to determine to what extent poverty impacts communities’ ability to produce these forms of knowledge. But insofar as these forms of knowledge have been produced, it is certainly consistent with a distributive agenda, and not contrary to my conclusions here, to grant them some form of IP protection.

Finally, to be clear, the conclusion that the poor may not easily be able to harness IP rights to their individual and collective advantage does not disrupt other important strands of the IP and distributive justice literature. Much of this literature, for instance, is concerned with how IP can advance distributive justice by increasing public access to IP-protected goods and services. Because this goal does not depend on the ability of the poor to create and obtain IP rights over their own creations, nothing discussed here calls into question either the feasibility or the desirability of this endeavor. Nor does it mean that for those living in poverty who do manage to create IP-protectable content, we shouldn’t try to make it as easy as possible, through changes to IP doctrine and procedure others have proposed, for them to acquire and assert IP rights. But we should be aware that these situations may be the exception, rather than the norm.

B. Explaining IP’s Observed Inequalities (and What it Means for Incentive Theory)

The psychological incompatibility between poverty and creativity throws into question the assertion that IP can serve as a meaningful tool of distributive justice by giving the poor opportunities to enhance their social and financial positions. But it may also help explain the reality we actually see on the ground. Because in fact, from the limited data that has been collected to date, it does appear that the poor and other disadvantaged groups are acquiring IP rights at much lower rates than their more advantaged counterparts. And although IP and distributive justice scholars’ argument does not wholly depend on ownership of IP rights—Hughes and Merges argue, for example, that simply participating in IP by working in IP-dominated fields benefits authors economically and socially regardless of whether they end up owning the resulting IP rights—IP ownership is one important indicator of the extent to which particular groups are, in fact, participating in IP.

174 See, e.g., Hughes, supra note 46, at 1216–20.
175 See, e.g., Fisher & Syed, supra note 43, at 583 (proposing a prize regime to improve access to pharmaceuticals in developing countries); Lee, supra note 43, at 925 (highlighting “various ‘accommodation strategies’ for integrating distributive values in an innovation system fundamentally predicated on profit maximization”); Shaver, supra note 43, at 122 (discussing how IP limits access among poorer populations to copyrighted works).
176 E.g., Hughes & Merges, supra note 7, at 561–75; Lee, supra note 4, at 347–52, 363–67.
177 See Bell et al., supra note 17, at 650.
178 Hughes & Merges, supra note 7, at 533–36. But see Greene, supra note 76, at 370–72 (arguing that racial minority artists are taken advantage of in contractual negotiations).
In this regard, the most direct measurement of IP participation among the poor comes from Raj Chetty and colleagues’ recent study on innovation and opportunity in America.\textsuperscript{179} The study finds, among other things, that children from the top 1% of families by income are ten times more likely than children from below-median families to go on to apply for or receive a patent.\textsuperscript{180} While some of this difference can be predicted by mathematic proficiency,\textsuperscript{181} math scores account for only a third of the total innovation gap between rich and poor.\textsuperscript{182} In fact, so much of the difference is unrelated to a talent for math that a poor child with strong math skills is less likely to hold a patent than a rich child with inferior math ability.\textsuperscript{183}

Consistent with this finding, Colleen Chien recently reported that IP rights tend to be increasingly concentrated in the hands of a small and elite group.\textsuperscript{184} Over half of new patents granted in 2016 went to the top one percent of patent owners, up from just under forty percent of patents going to the top one percent in 1986.\textsuperscript{185}

And though the following data do not speak directly to the link between poverty and IP, it appears that other minority groups—many of whom tend to cluster at the low end of the U.S. income distribution—\textsuperscript{186} are also underrepresented in the U.S. patent and copyright systems. In their demographic evaluation of patent owners, for instance, Chetty and colleagues found that the white children in their sample were over three times more likely to go on to apply for or get a patent than the black children, and eight times more likely to

\begin{footnotes}
\footnote{179} See generally Bell et al., supra note 17 (explaining that “there are many ‘lost Einsteins’—individuals who would have had highly impactful inventions had they been exposed to innovation in childhood—especially among women, minorities, and children from low-income families”).

\footnote{180} Id. at 649.

\footnote{181} Id. at 650 (finding that third grade math test scores help predict later patenting behavior).

\footnote{182} Id. The predictive power of math scores grows over time; by the eighth grade, these scores predict about half of the innovation gap between rich and poor children. Id. But as the study’s authors point out, this is also consistent with the hypothesis that test scores for poor children worsen over time, suggesting that environmental factors may be causing the poorer children to fall further behind their richer counterparts as time goes on. Id.

\footnote{183} See Bell et al., supra note 17, at 672–73 fig.IV (showing that there are fewer future inventors in a group of low-income children with a 0.5 standardized test score than there are in the group of high income children with a -0.5 standardized test score).


\footnote{185} Id.

\footnote{186} See, e.g., \textit{African American Income, BLACK DEMOGRAPHICS}, https://blackdemo graphics.com/households/african-american-income/ [https://perma.cc/PUM2-YRQ3] (using 2016 U.S. Census Bureau data to report that 46% of African Americans in the United States have annual household incomes below $35,000); Pressman, supra note 69 (reporting that “minorities had higher poverty rates than non-Hispanic whites, mainly because . . . minorities receive lower wages on average than whites”).
\end{footnotes}
do so than the Hispanic children in their sample. On the copyright front, Bob Brauneis and Dotan Oliar found that white copyright holders are overrepresented in copyright registrations, while Hispanic, Asian and Pacific Islanders, American Indians, and individuals of multiple races are underrepresented. Women are also underrepresented compared to men in both patent and copyright ownership.

These findings are concerning, and not only from a distributive justice perspective. The dominant view in the IP literature is that IP is about optimizing socially beneficial innovation. Though it is difficult to know what the “optimal” level of innovation is and where we currently reside in relation to that goal, the fact that certain identifiable groups seem to be innovating at lower levels than others—at least as measured by IP ownership—raises the possibility that we are falling short of optimality, in one of two ways.

First, we could be falling short of optimality in terms of sheer numbers of innovations. For example, if we assume that IP is doing its job well with respect to well-off white males—the best represented group in patent and

---

187 See Bell et al., supra note 17, at 666.
189 Id. at 60–61. Hispanics produced only 45% of the copyrighted works they would have if their participation in the copyright system mirrored their representation in the general population. Id. at 60. The same figure was 83% for Asian and Pacific Islanders, 77% for American Indians and Alaska natives, and 62% for people of multiple races. Id. at 61. For comparison, whites produced 116% of the copyrighted works expected based on their proportion of the population. Id. Unlike with patents, black owners are also overrepresented in the copyright system, producing 120% of the copyrighted works expected based on their proportion of the population. Id.
190 Kyle Jensen et al., Gender Differences in Obtaining and Maintaining Patent Rights, 36 NATURE BIOTECHNOLOGY 307, 307 (2018) [hereinafter Jensen et al., Gender Differences in Obtaining and Maintaining Patent Rights]; see also Bell et al., supra note 17, at 668 (finding that, of a cohort born in 1980, only eighteen percent of the inventors—defined as those who applied for or received a patent in their lifetimes—in the sample were women); Brauneis & Oliar, supra note 188, at 73 (finding that men represent two-thirds of copyright owners, despite comprising only half of the population and just over half of the labor force); Dan L. Burk, Diversity Levers, 23 DUKE J. GENDER L. & POL’Y 25, 31 (2015) (summarizing various empirical results and concluding that “women are at every level pervasively absent from the patent system”); Kyle Jensen et al., Why Do Women Inventors Win Fewer Patents?, YALE INSIGHTS (Apr. 9, 2018), https://insights.som.yale.edu/insights/why-do-women-inventors-win-fewer-patents [https://perma.cc/R4CQ-D3DB] (finding that only ten percent of inventors listed on patents are women, although women represent half the U.S. population and earn half of the science and engineering PhDs).
191 IP participation is an admittedly imperfect proxy for innovation. See, e.g., Bell et al., supra note 17, at 649 n.2 (citing Griliches and discussing the limitations of using patents as a proxy for innovation, including the fact that not all inventions are patented). See generally Zvi Griliches, Patent Statistics as Economic Indicators: A Survey, 28 J. ECON. LITERATURE 1661 (1990) (discussing in detail the use of patents as a proxy for innovation).
copyright—by encouraging them to innovate at optimal levels, then the lower participation rates among other groups might signal that IP is not doing its job so well with these populations, resulting in sub-optimal levels of innovation overall. Of course, another possibility is that IP is actually over-incentivizing wealthy white males. If this is the case, this surplus innovation may compensate for the deficit of IP-protected innovation from minority groups, resulting in optimal levels of innovation overall.

But even if this latter scenario is indeed what’s going on, we may be falling short of optimality on another metric—the types of innovations being produced. Presumably the optimal slate of innovations is one that includes a variety of artistic and scientific works in different media and from different fields of endeavor. Brauneis and Oliar, however (for example), find that different demographic groups tend to register different types of copyrightable works. Registration is only a proxy for actual creation, and some of these differences might reflect barriers to entry for certain groups rather than differences in interest. But we might still hypothesize that relying disproportionately on particular demographic groups for our innovation may result in insufficiently

---

192 With one exception—according to Brauneis and Oliar—black authors are overrepresented in the copyright system, and slightly more so than white authors (while white authors accounted for 74% of the registrations in 2010, though they represented only 63.7% of the population that same year—a ratio of 1.16, black authors accounted for 15% of registrations in 2010 though they comprised only 12.6% of the population—a ratio of 1.19). Brauneis & Oliar, supra note 188, at 59–62. The participation of African Americans in the copyright system has sparked some scholarly interest. While Brauneis and Oliar’s number suggests that African Americans, as a group, are reaping the advantages of copyright, some scholars have pointed out that this is a relatively new development. See, e.g., Greene, supra note 76, at 370 (“[F]or a long period of U.S. history, the work of black blues artists was essentially dedicated to the public domain.”). K.J. Greene has also argued that even when black artists are granted copyrights in their works, they might not receive adequate compensation due to copyright divestment and inequitable contracts that strip them of the lion’s share of their earnings. Id. On the other hand, Justin Hughes and Robert Merges have recently compiled empirical data suggesting that the copyright system is responsible for creating the wealth of the richest African Americans. Hughes & Merges, supra note 7, at 549. Based on these findings, they argue that “copyright in its current form is a powerful tool to empower creative individuals [including African Americans and other minorities] economically.” Id. at 516.

193 See LUNNEY, supra note 21, at 14 (explaining that under the incentive theory of IP, “suboptimal production does not refer to too few [new innovations] in some general ‘more is better’ sense, but to the precise economic relationship wherein ‘the (fully internalized) marginal social value of an additional [innovation] precisely equals its marginal social cost’”). Lunney explores an economic model according to which the market, even in the absence of copyright protection, leads to the over-production of musical works. See id. at 37.

194 Brauneis & Oliar, supra note 188, at 57–72, 75–78 (examining trends in types of registered works by race and gender).

195 See, e.g., Bell et al., supra note 17, at 649 n.2 (citing Griliches and discussing the limitations of using patents as a proxy for innovation, including the fact that not all inventions are patented). See generally Griliches, supra note 191 (discussing in detail the use of patents as a proxy for innovation).
varied output—both in terms of the types of works being created, and in the content of those works, as the substance of innovative works generally reflects the individual backgrounds and experiences of their creators.\textsuperscript{196}

If the lower rates of IP ownership in poor and minority populations are indeed a signal of sub-optimality in either number or variety of innovations, then specific social harms follow. On the numbers side, if we are failing to achieve the optimal level of innovation, society is missing out on the collective benefits these creations would have brought—benefits like access to new products, books, songs, and visual arts, technological advances, and the new jobs and growth that are spurred by them.\textsuperscript{197} The Chetty study estimates that if “women, minorities, and children from lower-income families were to invent at the same rate as white men from high-income (top-quintile) families, the total number of inventors in the economy would quadruple.”\textsuperscript{198} If the contributions of these groups are necessary to achieve optimal levels of innovation, this is a significant societal loss. And this is to say nothing of the personal economic losses involved, which will be borne primarily by the economically disadvantaged.\textsuperscript{199}

On the variety side, a sub-optimal assortment of innovations deprives society of the full range of human ingenuity. And there may be less obvious harms as well. Exposure to multicultural literature, for example, reduces prejudice and helps children develop empathy for members of races other than

\textsuperscript{196} See Brauneis & Oliar, \textit{supra} note 188, at 92 (“We believe that people bring something from themselves into their creativity, and that the authorship scene would integrate more insights, cater to more tastes, and generally be better and more interesting if a broader variety of people were involved in cultural production . . . .”).

\textsuperscript{197} See, \textit{e.g.}, Burk, \textit{supra} note 190, at 33 (citing Jennifer Hunt et al., \textit{Why Don’t Women Patent?} (Nat’l Bureau of Econ. Research, Working Paper No. 17888, 2012) for the proposition that the innovation gap between men and women alone represents a per capita GDP loss of 2.7%).

\textsuperscript{198} Bell et al., \textit{supra} note 17, at 653.

\textsuperscript{199} The burden will be borne primarily by disadvantaged groups because, as IP and distributive justice scholars have recognized, IP rights often translate into valuable financial gains for their owners. Perhaps somewhat surprisingly, empirical evidence suggests that this is true even for patents and copyrights owned by individual inventors and artists rather than by large companies. See Michael S. Kramer, \textit{Valuation and Assessment of Patents and Patent Portfolios through Analytical Techniques}, 6 J. MARSHALL REV. INTELL. PROP. L. 463, 485–86 (2007) (valuing issued patents using an analytical valuation model and finding, unexpectedly, that unassigned patents—i.e., those owned by their inventors and not assigned to an employer or company—were the fourth-most valuable category of patents analyzed); Hughes & Merges, \textit{supra} note 7, at 529 (arguing that that “copyright [has] a positive impact on the income of individual citizens,” via direct returns from individually owned copyrights or, in some case, through increased income as a reward for producing copyrighted works-for-hire for employers). If the economically disadvantaged are not acquiring or contributing to IP rights at the same rate as their peers, this puts them—as a group and as individuals—at an even greater financial disadvantage. This is particularly true given that IP rights are, in effect, a transfer of public wealth to private individuals. If, as the empirical evidence suggests, this transfer disproportionately benefits people who are already wealthy, it serves to compound existing economic inequalities.
their own—\footnote{Timothy Coon, How Does Exposure to Multicultural Literature Benefit Children’s Thought Processes About Race? (Aug. 2012) (unpublished M.S. thesis, St. John Fisher College) (on file with Fisher Digital Publications, St. John Fisher College).} if such literature is produced at sub-optimal levels, we are forgoing wide-ranging benefits we might have otherwise reaped. Societal harms aside, the individuals who would be creating but-for their poverty are also harmed in personal ways that go beyond the financial.\footnote{One of the non-economic rationales for granting IP rights, for example, proposes that people create to express themselves as human beings and therefore have a personality-based interest in their creations. \textit{E.g.}, Fromer, \textit{supra} note 34, at 1753; Hughes, \textit{supra} note 35, at 330; MERGES, \textit{supra} note 2, at 68–100; Radin, \textit{supra} note 35, at 971–78. If people who could and would be expressing themselves creatively but-for their financial status are not doing so, this arguably impoverishes their lives in ways that lowered financial gains and social status don’t fully capture.}

This discussion of optimality assumes that IP ownership is a reasonable proxy for innovation more broadly,\footnote{See, \textit{e.g.}, Bell et al., \textit{supra} note 17, at 649 n.2 (citing Griliches and discussing the limitations of using patents as a proxy for innovation, including the fact that not all inventions are patented). See \textit{generally} Griliches, \textit{supra} note 191 (discussing in detail the use of patents as a proxy for innovation).} so that we can indeed conclude from the data cited that members of poor and other minority groups are innovating at lower levels than their more advantaged counterparts. Though I believe this conclusion is essentially correct, there is certainly more to the story. Specifically, the data cited might underestimate the degree to which the poor are innovating, for a number of reasons. The poor may be less likely to seek out IP rights in their creations due to financial or structural barriers. They may be denied IP rights at higher levels due to bias, or because IP doctrines and

\footnote{See \textit{generally} Griliches, \textit{supra} note 191 (discussing in detail the use of patents as a proxy for innovation).}

\footnote{Those at the bottom of the socioeconomic scale may lack personal familiarity with the IP system, and they are also probably less likely to have people in their social networks with this familiarity. \textit{See} Bell et al., \textit{supra} note 17, at 651 (finding that exposure during childhood to other inventors makes it more likely that a child will himself go on to become an inventor); Greene, \textit{supra} note 74, at 353–54 (explaining how the 1909 Copyright Act’s complex registration requirements may have prevented black artists historically from getting protection for their work); Sunder, \textit{supra} note 3, at 273 (“[P]roblems encountered in protecting the knowledge of the poor [may] turn . . . on the poor’s lack of knowledge of their rights . . . .”).}

\footnote{See, \textit{e.g.}, Jensen et al., \textit{Gender Differences in Obtaining and Maintaining Patent Rights}, \textit{supra} note 190, at 309 (finding that women inventors with names that made it difficult to identify their gender were more likely to have a patent application accepted as compared to women with names that easily identified them as female).}
concepts entrench certain assumptions about how innovation should proceed. Or they may disproportionately choose to innovate in areas that are not IP-protectable.

Nevertheless, the conclusion we can draw from the data—that there is an innovation gap, or at least an IP gap, between rich and poor—is consistent with the psychological incompatibility between poverty and creativity. Indeed, the psychological profile of poverty provides an additional, novel explanation for these data: it tells us that it is not just bias, or financial barriers, or innovation outside IP, that is causing us to see the disparities we do. Something else is going on, and that something may be happening at the level of creative decision-making. This is significant because a full understanding of the IP/innovation gap can help us decide what will work—and what won’t—as we attempt to address it.

206 Dan L. Burk, Do Patents Have Gender?, 19 AM. U. J. GENDER SOC. POL’Y & L. 881, 889–903 (2011) (discussing the ways in which intellectual property doctrines might unintentionally perpetuate common gender biases); see also Burk, supra note 190, at 30–31 (arguing that stereotypical “rational” and “analytical” male approaches to problem solving, as opposed to more stereotypically female “emotive” or “intuitive” problem solving methods, “are more amenable to satisfaction of the teaching and disclosure requirements of patent law as currently formulated”).

207 For example, Peter Lee writes of “social innovation”: “novel creations that serve social needs,” but fall outside the bounds of patent protection. Peter Lee, Social Innovation, 92 WASH. U. L. REV. 1, 4 (2014). It is possible that social innovation disproportionately attracts members of certain demographic groups. According to Lee, for example, social innovation tends to be pluralistic and “arise collaboratively from communities.” Id. at 28. There is some evidence that collaboration is a particular strength of women, on average, as compared to men. See, e.g., Alice H. Eagly & Mary C. Johannesen-Schmidt, The Leadership Styles of Women and Men, 57 J. SOC. ISSUES 781, 782, 790 (2001) (discussing previous research suggesting that women have more collaborative leadership styles than men, but suggesting that this could be due in part to social expectations and internalized gender stereotypes about how women should behave in the workplace); Public Release, Univ. of Toronto, Women More Collaborative in Workteams: Study (Apr. 19, 2005), https://www.eurekalert.org/pub_releases/2005-04/uot-tms041905.php [https://perma.cc/XL7U-D5FX] (reporting on research by Jennifer Berdhal finding that teams with woman leaders tended to become more egalitarian over time, while those with male leaders retained a hierarchical structure with the leader at the center. Interestingly, the more egalitarian teams also performed better as judged by outsiders). And some of the social innovations Lee describes, like microfinance, are explicitly designed to help poor and minority communities. Lee, supra note 207, at 20 (“[T]he efforts of [microfinance institutions] are explicitly distributive; [one such entity] targets the poor and overwhelmingly lends to women, who comprise ninety-seven percent of borrowers.”). These fields may be particularly attractive to members of these communities who are looking for an outlet for their creative talents. See Sunder, supra note 3, at 290 (discussing how the DevNat licensing regime, originally formulated as a way for creators in the developed world to provide access to their creations in the developing world, has also been embraced by creators within the developing world as a way to support their own communities).
C. Beyond Incentives

Whether you subscribe to the dominant view that IP should focus primarily on the goal of maximizing socially beneficial innovation, or you believe that IP should also be cognizant of distributive concerns, the evidence that poverty interferes with creative decision-making should change the way you think about how to best achieve each of these goals.

As explored in the last section, one important takeaway for the efficiency camp is that poverty may be contributing to a suboptimal supply (in amount or variety) of innovation. This has implications for how IP scholars tend to think about innovation and how best to promote it.

Consistent with the utilitarian narrative that we need incentives designed to combat intellectual goods’ free-rider problem to encourage potential innovators to create, much of the current IP and innovation scholarship focuses on these incentives. For example, scholars often ask whether current IP doctrines provide the appropriate level of incentives, wherein the benefits reaped from more innovation outweigh the deadweight losses IP incurs. Or whether there are other, better, ways to provide individuals with these incentives—like grants, or prizes, or tax breaks, or social norms. The underlying, unstated

208 See supra Part I.
209 See, e.g., LUNNEY, supra note 21 (arguing that the copyright system currently offers excessive levels of protection to musical works, with attendant costs that outweigh the benefits of increased number and quality of works); Mark A. Lemley, Faith-Based Intellectual Property, 62 UCLA L. REV. 1328, 1338–44 (2015) (arguing that we should grant IP rights only to the extent necessary to optimize innovation); Olson, supra note 24 (analyzing patentable subject matter doctrine in light of the utilitarian rationale for patents).
212 See generally Daniel J. Hemel & Lisa Larimore Ouellette, Beyond the Patent-Prizes Debate, 92 TEX. L. REV. 303 (2013) (arguing that tax incentives may be a superior incentive mechanism for the production of intellectual products); Camilla A. Hrdy, Patent Nationally, Innovate Locally, 31 BERKELEY TECH. L.J. 1301 (2016) (arguing that state-funded tax and other incentives might prove to be a superior mechanism for funding innovation).
213 See, e.g., Kate Darling & Aaron Perzanowski, Introduction to Creativity Without Law: Challenging the Assumptions of Intellectual Property 1, 2 (Kate Darling & Aaron Perzanowski eds., 2017) (arguing that IP law has “historically disregarded non-legal regulatory tools that enable more granular, and potentially more effective, management of creative incentives”); Oliar & Sprigman, supra note 20, at 1832 (“None of the foundational theoretical studies (as distinguished from recent studies in IP law that focus on particular creative communities) meaningfully acknowledges the possibility that social norms can provide incentives to create.”); Kal Raustiala & Christopher Jon Sprigman, When Are IP
assumption is that if we get incentives right, the optimal level of innovation should follow.\textsuperscript{214}

What often gets overlooked in the discussion of incentives, however, is the possibility that there are potential innovators—those with the necessary talent and motivation to create—whose circumstances may prevent them from responding to optimal incentives in the expected way, or even at all. These circumstances may be embedded in their innovation environments, which may give rise either to “innovation dilemmas” of various kinds\textsuperscript{215} or counterproductive social norms that lead to sub-optimal decision-making about which avenues of research to pursue.\textsuperscript{216} Or, like here, they may be broader life circumstances that make it difficult for potential innovators to think creatively at all. Whatever the precise situation, the point is that creators are people living in complex environments, and ostensibly extraneous social and psychological forces in their lives will impact their ability to respond to even the most well-designed innovation incentives.

What should innovation scholars do about this? First, to the extent they can identify these social and psychological forces, they should. This will help them determine whether they can actually expect individuals to respond to the innovation incentives provided. In doing so, they may find large, identifiable groups of people—in this case, the poor—who may not be responding to innovation incentives as expected for reasons that have little to do with the quality of the incentives themselves.

Second, once scholars have identified forces that may be interfering with the ability to respond to innovation incentives, they can ask whether they can do anything to address and counter these forces, and whether they can expect these interventions to promote innovation in socially beneficial ways. This line of inquiry should prompt innovation scholars to move beyond the incentives-for-innovation paradigm and begin thinking much more broadly about innovation and how to promote it.

It is an especially timely moment to do so, given that scholars have begun to question whether innovation indeed flows from incentive in the simple manner so often presumed. Despite the prevalence of the canonical incentive rights necessary?: evidence from innovation in IP’s negative space. In 1 Research Handbook on the Economics of Intellectual Property Law: Theory 309, 314 (Ben Depoorter & Peter Menell eds., 2019) (“Many negative space studies have documented the powerful role social norms play in stimulating innovation and constraining appropriation.”). See Ouellette, supra note 18, at 66–67 (framing the question for innovation scholars as whether “patents provide a net innovation incentive,” or whether, alternatively, “other . . . incentives [are] superior”).

\textsuperscript{215} Brett M. Frischmann et al., Conclusion to Governing Knowledge Commons 469, 471–72 (Brett M. Frischmann et al. eds., 2014) (identifying a number of “innovation dilemmas” that confront potential innovators, including infrastructure problems and coordination challenges among research groups).

\textsuperscript{216} Stephanie Plamondon Bair & Laura G. Pedraza-Fariña, Anti-Innovation Norms, 112 NW. L. REV. 1069, 1091–95 (2018) (describing how informal rules growing from social forces and psychological biases that may lead creators down suboptimal innovative paths).
story in IP, the empirical evidence for this account is, at best, inconclusive. For example, in a recent empirical study of copyright’s effect on the music industry, Glynn Lunney concluded that more copyright revenue “did not lead to more and better music,” as the incentives for innovation paradigm would predict. In fact, it led to less high quality music output. Lunney explains this result by satisficing behaviors on the part of top artists: as the lion’s share of increased copyright revenue was funneled to a few top artists, these artists presumably felt little need to keep up a frenetic pace of music output, and their productivity dropped. Lunney’s conclusion that the psychological phenomenon of satisficing adds complexity to the incentives-for-innovation paradigm is consistent with a growing realization that social and behavioral realities may disrupt this paradigm in various ways.

Given that a simple relationship between financial incentives and innovation output has been called into question—and indeed, that incentives may not work at all for certain demographic groups—perhaps it is time to stop focusing exclusively on incentives, and to consider other, non-traditional innovation-promoting policy levers. In doing so, scholars might discover mechanisms that are superior in various ways to incentives. Indeed, some of these levers may be better able than IP or other financial incentives to achieve both distributive justice and innovative efficiency goals.

V. POLICY IMPLICATIONS

A. Promoting Innovation by Tackling Distributive Justice

The relationship between poverty and creative thinking throws into question claims that IP can serve as a meaningful mechanism for distributive justice by giving the poor a tool to improve their economic and social situations. It also helps explain why we see lowered IP participation among the poor, and sounds a warning note to efficiency scholars who assume that getting innovation incentives right is the sole mandate of a robust innovation policy.

Considering these insights together, an interesting possibility presents itself—one that turns IP scholars’ current thinking about IP and distributive justice on its head. Rather than using IP as a mechanism for promoting

217 See, e.g., Lemley, supra note 209, at 1335 (“[W]e have gone out, collected the evidence, and found that it is far from clear that IP is doing the world more good than harm.”).
218 LUNNEY, supra note 21, at 193.
219 Id.
220 Id. (“[A]s revenues increased, earnings for top artists rose sharply; as they did, our top artists started producing fewer songs.”).
221 See, e.g., Bair, supra note 34, at 314–16 (discussing psychology-based challenges to the incentive theory of patents).
222 See supra Part III.A.
223 See supra Part III.B.
224 See supra Part III.C.
empowerment and distributive justice in poor populations—a strategy that likely will not work on a large scale—policies more directly aimed at tackling distributive justice should not only be more successful at attaining this goal, but should also help promote innovation by increasing creative participation among currently underrepresented groups.

The logic of this proposition is as follows. If the circumstance of poverty does indeed interfere with creative thinking, then there are some number of talented, motivated individuals currently living in poverty who but-for their poverty would be innovating. Policies designed to tackle poverty, if successful, should have positive spillover effects on innovation as those talented, motivated individuals—now freed from the cognitive demands poverty imposes—begin to put their creative talent and motivation to good use. Indeed, as Chetty estimates, these positive spillover effects might be quite substantial.225

What might these interventions look like? To achieve the positive spillover effects on innovation, distributive justice measures that directly tackle the creativity-disrupting cognitive burdens poverty inflicts would seem to be the most promising.

1. Universal Basic Income

One such intervention that has received a lot of attention lately is the universal basic income (UBI).226 The basic idea behind a UBI is simple: the government provides each of its citizens with some minimum level of cash income—no strings attached.227 The manner in which the government might do this could take a number of forms, including through a progressive taxation system, or via simple cash payments.228 The UBI helps achieve a Rawlsian vision of distributive justice by ensuring that all citizens have at least the minimum amount of resources necessary for meaningful participation in

225 Bell et al., supra note 17, at 653 (estimating that if “women, minorities, and children from lower-income families were to invent at the same rate as white men from high-income (top-quintile) families, the total number of inventors in the economy would quadruple”).


228 See id. at 1196–97, 1234–36.
By directly easing the financial burdens of the poor, the UBI should also reduce the cognitive loads imposed by poverty, thereby freeing the minds of its beneficiaries to think in more creative ways.

One question UBI scholars have asked is whether the UBI should be available to children as well as adult citizens. In terms of achieving positive spillover effects on innovation, making the UBI available to children would be ideal, because of the particularly harmful and long-lasting effects of poverty on the creative thinking capacities of children. To the extent policymakers can prevent or mitigate these impacts by directly ameliorating the financial condition of children, they should expect to reap correspondingly larger positive spillovers resulting from increased innovation and creative thinking throughout these children’s lifetimes.

2. Access to Health Services

There are other possibilities as well. Increasing access to affordable basic health services, through a government-funded single payer system or targeted improvements to the current market-based approach, would arguably help achieve distributive justice ends by ensuring that the poor are not disproportionately bearing the costs of health care in the U.S. But it should also have positive effects on innovation, via several mechanisms. First, by easing or removing the financial burden imposed on the poor in exchange for essential health services, health care reform would ease the cognitive load caused by preoccupations about how to pay for these services. Second, by increasing access to affordable health care, the poor should receive better care

---

229 See Linda Sugin, Competitive Philanthropy: Charitable Naming Rights, Inequality, and Social Norms, 79 Ohio St. L.J. 121, 141–42 (2018) (arguing that “Rawls’s theory is consistent with [a UBI], since he demands that individuals have the ability to participate in society, which presumably requires some baseline resources”).

230 See Fleischer & Hemel, supra note 227, at 1253 (noting that “[s]everal prominent proponents of a UBI explicitly exclude children from their proposed basic income schemes,” before arguing that children should also be the direct beneficiaries of any UBI program).

231 See supra Part II.A.


233 See id. at 71–72 (arguing that the U.S. health care system as currently structured, promotes distributive injustice because both “the burden of paying for public goods such as health care for the uninsured, medical education, and pharmaceutical research . . . fall[s] disproportionately on those with less ability to pay,” and “persons with lower incomes [are] compelled to pay, as part of the price of having any health insurance at all, either for coverage designed by and for elite interests or for health care that is consumed disproportionately by the well-to-do”).

234 See supra Part II (discussing how cognitive loads imposed by financial difficulties affect decision-making and impair creative thinking).
than they are currently receiving, which in turn should affect their ability to think creatively. Ill health imposes its own physical and cognitive burdens independent of financial worries, and those who aren’t receiving basic health services may be less able to engage in creative pursuits. Indeed, the fact that poverty privileges non-creative habit-based decision-making over goal-based decision-making could be due in part to reduced health in poor populations.

Third, reduced access to maternal health care services is one of the hypothesized contributors to impaired brain development in poor children. Tackling this aspect of creativity impairment in children should help mitigate the negative effects poverty imposes on a child’s cognitive potential.

UBI and health care reform might also have a somewhat different (but related) salutary effect on innovation. Some subset of the population might be willing to accept a fairly low standard of living if doing so allows them to pursue creative goals. Having access to a guaranteed minimum wage and basic health care services could help make the decision to pursue a creative life easier for those who might otherwise resign themselves to a profession that pays the bills, but does not offer the time or opportunity to innovate. And for those who have already committed to pursuing creative vocations, it could free up additional time to do so. One Australian scholar, for example, estimates that “most artists in Australia spend less than 50% of their time on their creative vocation;” the balance being expended on mundane work to meet their basic needs. Programs that provide for these basic needs would arguably improve the productivity of those in creative professions who spend significant time doing non-creative work in order to make ends meet.

3. Other Possibilities

Distributive justice scholars have proposed a number of additional policy interventions designed to achieve their ends. For example, in their article

---

235 See, e.g., Farah, supra note 89, at 428 (“[S]ocial science research shows that poverty is associated with shorter and less healthy lives.”); Sarah Toy, U.S. One of the World’s Worst on Health Divide Between Rich, Poor, USA TODAY (June 7, 2017), https://www.usatoday.com/story/news/2017/06/07/us-one-worlds-worst-health-divide-between-rich-poor/102583180/ [https://perma.cc/UL6V-MLJX] (citing research finding that “poor Americans reported worse health than rich U.S. residents in significant numbers,” and concluding that a probable cause of this disparity is reduced access to health care among poor populations).

236 Chen et al., supra note 11, at 11979 (finding that sleep deprivation privileges habit-based over goal-based decision-making strategies); Patel et al., supra note 133, at 475 (finding that “health indicators significantly influenced sleep quality most prominently in poor individuals”).

237 Farah, supra note 89, at 431 (citing “prenatal health” as one of the causal factors that may explain the association between poverty and impaired brain development in children).

238 See, e.g., Goins, Story of the Starving Artist, supra note 140 (arguing that it is fear of extreme poverty that drives many would-be artists to “become lawyers instead of actresses, bankers instead of poets, doctors instead of painters”).

239 Chohan, supra note 138.
arguing that copyright protection can serve distributive justice ends, Hughes and Merges acknowledge that other mechanisms may be superior to IP in this respect.\footnote{Hughes & Merges, supra note 7, at 547.} Espousing a Rawlsian vision of distributive justice that requires at least equality of opportunity, if not equality of outcome,\footnote{Id.} they cite to the potential of educational reforms designed to give “all children equal access to pre-school, kindergarten, K-12 education, sports programs, summer camps, music lessons, university, and the like.”\footnote{Id.}

Equal access to education is doubtless an important component of achieving distributive justice. Indeed, uneven educational opportunities between the poor and the rich may help explain both why it is so difficult for poor children to escape poverty\footnote{See supra notes 191–95 and accompanying text.} and why the former group appears to be innovating at lower levels than the latter.\footnote{See supra Part II.A.}

But the psychology of poverty tells us that if we truly want to achieve distributive justice, education reform will not, on its own, suffice. Because poverty interferes with brain development and co-opts scarce cognitive resources,\footnote{See, e.g., Eduardo Porter, Inequality Undermines Democracy, N.Y. TIMES (Mar. 20, 2012), https://www.nytimes.com/2012/03/21/business/economy/tolerance-for-income-gap-may-be-ebbing-economic-scene.html [https://perma.cc/JJ37-7H8W] (citing a study estimating that only six percent of children born to parents in the bottom quintile of income will rise as adults to the top quintile, while forty percent will remain in the bottom quintile).} poor children will not be fully able to take advantage of even the best educational opportunities. To achieve true distributive justice—and to achieve maximum positive spillover effects on innovation—the underlying causes of this cognitive disruption should be addressed. In this respect, interventions that tackle poverty itself, and identified contributors to the physiological and psychological effects of poverty—like lower quality health care—have the most potential to eliminate barriers to distributive justice and increase participation in innovation.

B. Thinking Beyond Incentives: Other Innovation-Promoting Policy Levers

Of note for efficiency scholars is the fact that none of the policy interventions discussed to this point fit the standard incentive formulation where an individual or organization innovates (or shows that it will innovate) and receives some financial reward in return—a patent, a copyright, a prize, a grant, or a tax break. And yet scholars might expect these policies to have significant innovation-promoting effects.

This conclusion opens up an intriguing possibility for innovation scholars: if programs and policies that don’t offer a direct incentive for innovation can
nevertheless have significant innovation-promoting effects, perhaps—
especially given that the incentive story of innovation is more complicated than
is often assumed—they should be thinking about other ways to promote
innovation that similarly don’t fit the traditional incentive-for-innovation
model.

What might some of these policies look like? As a first modest step, we
might consider a more robust use of federal funds to finance artists and scientists
without making these funds contingent on output. Currently, most federal
science funding requires applicants to submit a grant proposal, which is peer-
reviewed and granted based on a number of criteria, including the expected
significance of the work and the potential for completion. Senior researchers
spend much of their time on these grant proposals, time that arguably could be
better spent on more creative tasks. Administering these grant systems is also
costly. One Canadian study estimated, for example, that it cost more to run a
national grant system in 2007 than it would have to simply give every qualified
researcher in the country a baseline grant of $30,000. Finally, a common
observation among researchers is that the level of detail required to secure
funding is such that the work is often substantially complete before the funding
arrives. Scientists thus often use secured funds to finance their next
exploratory project—a reality which is more in line with the concept of a
noncontingent grant based on past productivity than with the current, expensive
model of proving up a project’s worthiness. While some federal funding
opportunities for artists are somewhat more open-ended, the arts too could
perhaps benefit from funding models that are less focused on the traditional
incentive model.

247 See, e.g., Jenny Rohn, Show Me the Money: Is Grant Writing Taking Over Science?, GUARDIAN (Apr. 2, 2013), https://www.theguardian.com/science/occams-corner/2013/apr/02/1 [https://perma.cc/X352-PRVB] (“There is some evidence that having the vast majority of scientists spend the vast majority of their time writing grants instead of doing and thinking science might be a tad inefficient, and not, perhaps, the best way to get science done.”).
248 Richard Gordon & Bryan J. Poulin, Cost of the NSERC Science Grant Peer Review System Exceeds the Cost of Giving Every Qualified Researcher a Baseline Grant, 16 ACCOUNTABILITY RES. 13, 13 (2009).
249 Personal observation of the author.
250 See, e.g., Creative Writing Fellowships: Applicant Eligibility, NAT’L ENDOWMENT FOR ARTS, https://www.arts.gov/grants-individuals/creative-writing-fellowships/applicant-eligibility [https://perma.cc/9ANF-49DK] (describing eligibility criteria for an individual creative writing grant that requires only a demonstration of past productivity).
251 See, e.g., Grants for Organizations, NAT’L ENDOWMENT FOR ARTS, https://www.arts.gov/grants/apply-grant/grants-organizations [https://perma.cc/YW2M-RL46] (listing various grants for organizations and specifying that for organizations, the NEA “fund[s] projects only”).
VI. CONCLUSION

A growing number of scholars think that IP should serve distributive justice goals. While this is an attractive idea, it is not clear that IP is actually a good mechanism for doing so. This Article has argued that, in fact, IP is probably not the best way to achieve distributive justice, at least in the sense of providing the poor with opportunities to accumulate wealth and improve their social status.

The reason lies partly in the psychology of poverty. Because poverty affects decision-making in ways that make creative thinking more difficult, IP seems to be inherently limited as a tool for escaping poverty. At the same time, expecting IP to be such a tool gives rise to its own set of evils, because it incorrectly assumes the poor can in fact take advantage of IP, and indirectly puts the blame on them when they don’t.

These conclusions should be of interest not only to those who think of IP in distributive justice terms, but also to those who subscribe to IP’s more traditional, utilitarian accounts. The psychology of poverty helps us understand why current levels of innovation are suboptimal and highlights a potential failure of the incentive model of innovation production. A solution to this failure is to start thinking beyond the traditional incentives-for-innovation model when considering how best to maximize socially beneficial creation. Fortuitously, doing so should help us achieve distributive justice goals much more efficiently than IP ever could.
Would You Like to Save Your Game?:
Establishing a Legal Framework for Long-Term
Digital Game Preservation

Will White*

Table of Contents

I. Introduction ................................................................. 567
II. A Brief History of Digital Games and Their Formats ....572
   A. Invention .................................................................. 573
   B. Cartridges and the First Home Consoles .................. 574
   C. The Move to Optical Media .................................... 576
   D. The Rise of Digital Distribution ............................... 576
III. The Challenges Facing Game Preservation ............... 578
    A. Technical Challenges and the Need for Preservation .... 578
       1. Media Degradation ........................................... 578
       2. Obsolescence .................................................. 580
    B. Legal Hurdles for Game Preservation ..................... 583
       1. Copyright ....................................................... 583
       2. Source Code and Trade Secret Law ...................... 586
       3. DRM and Server-Side Authentication .................. 587
    C. Current State of Game Preservation ....................... 591
IV. Establishing a Federally Funded Preservation Board .... 591
    A. Film Preservation: A Case Study ............................ 592
       1. Early Film Preservation Efforts (Late 1800s–1988) ..... 594
       2. Establishment of the National Film Preservation Board... 596
    B. Establishing a Federally Funded Preservation Board .... 598
       1. Establishing and Maintaining a Source Code Archive .... 599
       2. Establishing a Collective Stance Regaining DMCA
          Exceptions ......................................................... 602
       3. Establishing Model Agreements and Best Practices ..... 603
V. Conclusion...................................................................... 604

I. Introduction

“Anything not saved will be lost.”

*Managing Editor, Ohio State Law Journal; Juris Doctorate Candidate, The Ohio State University Moritz College of Law, 2020. Special thanks to Darryl Walton for input on trade secret law and to Nick Uguccini from “The Poor Student Hobbyist” for input on some of the technical aspects of this Note (https://thepoorstudenthobbyist.com).
Since their invention in the 1960s, digital games—also known as video games, computer games, or electronic games—have grown from a quirky digital novelty into a diverse artistic medium, an economic juggernaut, and a mainstream cultural touchstone. As of 2018, around 60% of Americans play

---


2 For simplicity’s sake, “digital games,” within the context of this Note, will be used primarily to refer to games developed for home video game consoles. While a good portion of the discussion in this Note can be applied to computer games, arcade games, mobile phone games, and other electronic games (such as the Tiger Electronic games of the ’90s), each of these forms presents its own unique history and preservation challenges. This Note also focuses on the preservation of software, though it should be noted that the preservation of hardware presents its own difficulties and considerations. See, e.g., Adi Robertson, The Last Scan: Inside the Desperate Fight to Keep Old TVs Alive, VERGE (Feb. 6, 2018), https://www.theverge.com/2018/2/6/16973914/tvs-crt-restoration-led-gaming-vintage [https://perma.cc/7DV2-7X6H] (covering the obsolescence and death of CRT televisions and noting how a “game”’s look and feel is often highly dependent on specific hardware setups).

3 See, e.g., Paola Antonelli, Video Games: 14 in the Collection, for Starters, MUSEUM MOD. ART (Nov. 29, 2012), https://www.moma.org/explore/inside_out/2012/11/29/video-games-14-in-the-collection-for-starters/ [https://perma.cc/Y7RW-UPP8] (announcing the admission of digital games to the Museum of Modern Art and proclaiming, “Are video games art? They sure are, but they are also design . . . [The games selected are] outstanding examples of interaction design—a field that MoMA has already explored and collected extensively, and one of the most important and oft-discussed expressions of contemporary design creativity.”). As of 2011, the United States Supreme Court has held that digital games qualify for First Amendment protection, holding that “[l]ike the protected books, plays, and movies that preceded them, video games communicate ideas—and even social messages—through many familiar literary devices (such as characters, dialogue, plot, and music) and through features distinctive to the medium (such as the player’s interaction with the virtual world).”, Brown v. Entm’t Merchs. Ass’n, 564 U.S. 786, 790 (2011).


some form of digital games daily.\(^6\) Despite the persistent boy’s club nature of the industry and culture around digital games,\(^7\) actual consumption of digital games has gradually become more egalitarian, with women representing 45% of self-identified gamers as of 2018.\(^8\) Further, the average age of players continues to get older over time, with the vast majority of digital game players now over the age of eighteen.\(^9\)

Regardless of the ongoing discussions regarding the artistic merits and capabilities of digital games, their cultural and historical relevance is undeniable. In addition to representing how a growing portion of our current culture spends their leisure time, even their most problematic aspects offer

---


\(^7\) See generally Leigh Alexander, All the Women I Know in Video Games Are Tired, OFFWORLD (May 29, 2015), https://boingboing.net/2015/05/29/all-the-women-i-know-in-video.html [https://perma.cc/MS4J-PNS2] (discussing the exhausting lengths that women often must go to in order to feel validated working in or with the digital games industry).

\(^8\) Christina Gough, Distribution of Computer and Video Gamers in the United States from 2006 to 2019, by Gender, STATISTA (July 3, 2019), https://www.statista.com/statistics/232383/gender-split-of-us-computer-and-video-gamers/ [https://perma.cc/WL8U-CH7Z] (showing that this percentage has gradually risen from 38% in 2006); see also DUGGAN, supra note 6, at 2 (finding that “[a] majority of American adults (60%) believe that most people who play video games are men,” but that a “nearly identical share of men and women report ever playing video games (50% of men and 48% of women)”).

unique widows into contemporary understandings of topics such as modern
gender dynamics, race relations, and public perceptions of the military.

Unfortunately, the rapid growth and evolution of this powerful medium has
left a scattered and rapidly decaying history in its wake. The physical media
holding many of the games from the last few decades has literally begun to rot
away, the code needed to ensure long term preservation is often lost or
inaccessible, and several layers of legal barriers stand in the way of looking
for lawful solutions to long-term preservation. As a young and technology-
driven medium, the digital games industry and the conversations around it have
an often crippling short-term memory, and the conversations around digital
games often focus on the innovations of the future rather than the foundations
laid by the past. Many of the foundational games that represent the birth and
early evolution of digital games are disappearing bit by electronic bit, and we
are actively losing the generation of inventors and designers that defined the
earliest days of digital games. Without proactive measures, much of the history


13 See infra Part III.A.1.

14 See infra Part III.A.2.

15 See infra Part III.B.

16 See Gita Jackson, The Vast, Unplayable History of Video Games, OFFWORLD (May 28, 2015), https://boingboing.net/2015/05/28/the-vast-unplayable-history-o.html [https://perma.cc/79AR-RMW2] (quoting author and professor Ian Bogost as saying, “Collectively, we have a short memory, mostly back to the childhoods of whatever generation is currently not fed up with games enough to romanticize it . . . [T]he bigger trends always seem to start from scratch, unaware of what came before, unable to incorporate and build upon it.”).

of digital games stands to be lost forever—potentially within the next few decades.\textsuperscript{18}

In this current atmosphere of retro resurgence,\textsuperscript{19} decades may seem like plenty of time to craft a solution and many older games may still feel readily available to the general public.\textsuperscript{20} But the early days of film preservation presents a harrowing case study of just how quickly those years can pass and how large swaths of our culture and history can be forever lost when industry figureheads and legislators fail to coordinate with preservationists.\textsuperscript{21} The problems facing digital game preservation are wrapped up in a unique web of intellectual property law that may take many years to untangle.\textsuperscript{22} Meanwhile, the broad commercial value (and thus availability) of much of gaming history should not be expected to maintain its current intensity past our collective familiarity with it.\textsuperscript{23} As the generations of people who cherish these early games begin to die off, so too will the games themselves.\textsuperscript{24} We as a society must prioritize preservation while interest is high and the games are still available, or else we risk repeating the failures of early film preservation and losing a vital part of our shared cultural history.


\textsuperscript{21}See infra Part IV.A.

\textsuperscript{22}See Mark Methenitis, \textit{Laws of the Game: Intellectual Property in the Video Game Industry}, in \textit{VIDEO GAME POLICY: PRODUCTION, DISTRIBUTION, AND CONSUMPTION} 11, 11 (Steven Conway & Jennifer deWinter eds., 2016) (“Software is one of the only items in which all three major forms of intellectual property—copyright, trademark, and patent—can be present in one product.”).

\textsuperscript{23}See Christian Candia et al., \textit{The Universal Decay of Collective Memory and Attention}, 3 NATURE HUM. BEHAV. 82, 85 (2019) (finding that cultural products such as music and movies typically remain in the collective “communicative memory” of the public for only five to ten years, while biographies of famous individuals can be expected to last fifteen to thirty years); see also Sean Fenty, \textit{Why Old School Is “Cool”: A Brief Analysis of Classic Video Game Nostalgia}, in \textit{PLAYING THE PAST: HISTORY AND NOSTALGIA IN VIDEO GAMES} 19, 22 (Zach Whalen & Laurie N. Taylor eds., 2008) (referring to people born in the last forty years as the “games generation” and discussing how the unique spatial features of digital games create a powerful nostalgia in players by offering “a past within which players can participate, and a past in which players can move and explore”).

\textsuperscript{24}Candia et al., \textit{supra} note 23, at 87–88 (explaining how “the dynamics of human collective memory follow a universal decay function”).
The goal of this Note is to urge the establishment of an administrative and legal framework that facilitates the long-term preservation of digital games and promotes coordination between game publishers, industry organizations, and public and nonprofit archives. Part II begins with a streamlined history of digital games, with a specific focus on the diversity of storage mediums and the recent transition from physical to digital distribution. Part III discusses the technological and legal hurdles that stand in the way of preservation, the antagonism of major industry organizations towards current preservation efforts, and the current state of digital game preservation. Much of this Part will focus specifically on the importance of source code to long-term preservation and how industry negligence and volatility has led to and will continue to lead to irreparable losses. Part IV urges the legislative establishment of a national game preservation board to coordinate preservation efforts between game publishers, industry organizations, and public and nonprofit archives. A brief history of the early losses of film preservation is presented as a case study of just how quickly the foundations of a new medium can be lost, while the eventual establishment of the National Film Preservation Board presents a model for how to approach digital game preservation. Part VI briefly concludes.

II. A BRIEF HISTORY OF DIGITAL GAMES AND THEIR FORMATS

The history of games has been shaped by their dual existence as an interactive mode of expression driven by human ingenuity and creativity, and their commercial identity as consumer products, driven by tech trends and popular tastes. It is this latter identity as a tech-driven consumer product that has left the medium’s history scattered across a variety of incompatible formats and hardware configurations. Consideration of even the most streamlined history of console games and the transition from physical media to digital distribution underscores the magnitude of this issue and the technological hurdles that must be cleared to ensure long-term preservation. Note that, perhaps ironically, the brevity of such a truncated history will necessarily skip over many of the more esoteric corners of digital games’ history (such as the endearingly

---

25 Even this briefest of histories underscores the sheer number of formats and hardware configurations presented by digital games and the unique challenges of preserving a medium whose history is scattered across these disparate platforms.

26 Steven Conway & Jennifer deWinter, Introduction to VIDEO GAME POLICY: PRODUCTION, DISTRIBUTION, AND CONSUMPTION, supra note 22, at 1, 2 (“[V]ideo games are part of a broader media landscape that exists in the contested space between artistic freedom and economic incentives.”).
bizarre Vectrex home vector graphics system)\textsuperscript{27} that are arguably the most in danger of being lost to time.\textsuperscript{28}

A. Invention

We as human beings are naturally hardwired to play.\textsuperscript{29} Thus, it seems only natural that, given some time, the emerging computer technology of the mid-twentieth century would produce games of some sort.\textsuperscript{30} In fact, it seems it was so natural that it happened in at least three separate and isolated instances.\textsuperscript{31} The earliest known digital game was William Higginbotham’s 1958 Tennis for Two, which played on a scaling oscilloscope and was created as a tech demo so that people would have something to interact with during an open house at the U.S. government’s Brookhaven National Laboratory.\textsuperscript{32} Four years later, in 1962, without any knowledge of Tennis for Two, Steve Russell and a group of MIT students designed the second known digital game, Space Wars, programmed on a stack of paper punch cards—purely for their own amusement rather than any commercial aspirations.\textsuperscript{33} The game became well known across the MIT campus but never reached the broader public consciousness.\textsuperscript{34} It wasn’t until 1972, with Ralph Baer’s invention of the Magnavox Odyssey home console, that digital games first became available to general consumers.\textsuperscript{35} The first reaction to this monumental marriage of technology and play was, as one might

\textsuperscript{27} See Kent, supra note 1, at 230–33; Ben Kuchera, Why I’m Still Hung Up on the Weirdest Console Flop of the 1980s: When You Can’t Separate the Hardware from the Games, Polygon (Aug. 1, 2018), https://www.polygon.com/2018/7/30/17616756/vectrex-backlog-classic-gaming [https://perma.cc/7RUY-7UQF].

\textsuperscript{28} For simplicity’s sake, the history presented below is based on American markets using American release dates. For a more complete picture of the earliest days of digital games, see generally Kent, supra note 1.

\textsuperscript{29} Kerrie Lewis Graham & Gordon M. Burghardt, Current Perspectives on the Biological Study of Play: Signs of Progress, 85 Q. Rev. Biology 393, 400–02 (2010) (finding that “[p]lay is well-developed in primates, rodents, carnivorans, ungulates, elephants, and cetaceans,” but that play behavior may extend as far as insects and other invertebrates).

\textsuperscript{30} See Gita Jackson, Spacewar! Creators Didn’t Know They Were Making History, Kotaku (Dec. 5, 2018), https://kotaku.com/spacewar-creators-didnt-know-they-were-making-history-1830887504 [https://perma.cc/26LU-PL3A] (quoting one early inventor, Martin Graetz, as saying, “There was never any sense in any way that we were creating something new. . . . It was all a lark.”).

\textsuperscript{31} Kent, supra note 1, at 26; Mark Guttenbrunner et al., Keeping the Game Alive: Evaluating Strategies for the Preservation of Console Video Games, 5 Int’l. J. Digital Curation 64, 68 (2010); Jackson, supra note 30.

\textsuperscript{32} Kent, supra note 1, at 18; Guttenbrunner et al., supra note 31, at 68.

\textsuperscript{33} Kent, supra note 1, at 16–21; Guttenbrunner et al., supra note 31, at 68; Jackson, supra note 30.

\textsuperscript{34} Kent, supra note 1, at 20.

\textsuperscript{35} Id. at 21–26.
expect, one of mild confusion, public disinterest, and economic failure. A new medium was born.

B. Cartridges and the First Home Consoles

At their very simplest, programmable cartridges consist of a plastic shell, a printed circuit board, and some form of ROM chip to store the game. Both internal and external designs vary wildly from console to console, and even within the same console, individual game cartridges may vary wildly or even contain extra hardware elements such as graphics chips, save batteries, or sound chips.

The Magnavox Odyssey technically had the first game cartridges, but these were quite different from the cartridges many of us are familiar with today. The Odyssey’s cartridges contained no actual software but “merely complete[d] different circuit paths within the hardware itself to define the rule set for the current game.” Meanwhile, when Atari released the consumer version of Pong to home markets in 1975, its strategy was even simpler: the system only played Pong.

---

36 See Alexander Smith, _1TL200: A Magnavox Odyssey_, THEY CREATE WORLDS, https://videogamehistorian.wordpress.com/2015/11/16/1tl200-a-magnavox-odyssey/ [https://perma.cc/YA3X-QCAN] (finding that the Odyssey’s early poor performance was likely the result of high price and confusing marketing).


38 See, e.g., _The SNES Cartridge, Briefly Explained_, POOR STUDENT HOBBYIST (May 18, 2019), https://thepoorstudenthobbyist.com/2019/05/18/custom-pcb-explanation/ [https://perma.cc/J7N5-CJFP] (explaining that Super Nintendo cartridges alone can be sorted into “six major categories,” which may then contain various “enhancement chips”).


41 See Apparatus for Producing a Plurality of Audio Sound Effects, U.S. Patent No. 4,314,236 (filed Jan. 24, 1979) (registering the patent for the “POKEY” sound chip, used in certain Atari 7800 games).

42 Smith, supra note 36.

43 Id.

44 KENT, supra note 1, at 80–81.
It wasn’t until 1976 that the obscure Fairchild Channel F system introduced what we now recognize as programmable cartridges.\(^{45}\) This same technology soon became ubiquitous across home consoles and was adopted by the Atari 2600 in 1977, the Mattel Intellivision in 1980, and Coleco’s Colecovision in 1982.\(^{46}\) This same technology was used by Nintendo and Sega throughout the late eighties and early nineties\(^{47}\) with the Nintendo Entertainment System (1985),\(^{48}\) the Sega Master System (1987),\(^{49}\) the Sega Genesis (1989),\(^{50}\) the Super Nintendo Entertainment System (1991),\(^{51}\) and the Nintendo 64 (1996).\(^{52}\)

Despite the fact that optical media has almost fully supplanted cartridges as the preferred storage medium for modern home consoles,\(^{53}\) cartridges still remain in use to this day.\(^{54}\) And while cartridges reigned supreme after the release of the Channel F, they were not the only storage medium used during this era. Certain consoles and computers also stored games on similar but much more fragile media such as floppy disks,\(^{55}\) cassettes,\(^{56}\) and even VHS tapes.\(^{57}\)

---

\(^{45}\) Guttenbrunner et al., supra note 31, at 69.

\(^{46}\) Id.

\(^{47}\) This was following a brief market crash in 1983 that eliminated most of the industry and paved the way for Nintendo and Sega to rise in prominence. See Kent, supra note 1, at 234–40; N. R. Kleinfield, Video Games Industry Comes Down to Earth, N.Y. TIMES (Oct. 17, 1983), https://www.nytimes.com/1983/10/17/business/video-games-industry-comes-down-to-earth.html [https://perma.cc/88VW-7L8S].

\(^{48}\) Kent, supra note 1, at 296–97.

\(^{49}\) Guttenbrunner et al., supra note 31, at 70.

\(^{50}\) Kent, supra note 1, at 404.

\(^{51}\) Id. at 431.

\(^{52}\) Id. at 537.

\(^{53}\) See infra Part I.L.C.

\(^{54}\) The 2017 Nintendo Switch is the most recent example of a home console that uses cartridges as a storage medium despite the greater cost of production. Wesley Yin-Poole, Why Nintendo Switch Games Are Ending Up More Expensive, EUROGAMER (Mar. 13, 2017), https://www.eurogamer.net/articles/2017-03-10-why-nintendo-switch-games-are-ending-up-more-expensive [https://perma.cc/P9VS-3XRW] (reporting that the cost of producing Nintendo Switch cartridges might be causing Switch games to cost more than those on consoles that use Blu-ray discs).


C. The Move to Optical Media

Beginning in the early nineties, the lower price and higher capacity of compact discs (CDs) caused many console manufacturers to switch to optical media.\textsuperscript{58} In 1992, Sega released the Sega CD add-on for the Sega Genesis, a peripheral which allowed the Genesis to play the first CD-ROM games for American home consoles.\textsuperscript{59} The oft-forgotten 3DO was the first American console to play CD-ROMs out of the box in 1993, but the true shift toward optical media didn't become a market force in digital games until the release of the Sony PlayStation in 1994.\textsuperscript{60} Optical media has been the primary storage medium for home consoles ever since, with most manufacturers moving to DVDs in the early 2000s and, later, Blu-rays discs, starting with the PlayStation 3 in 2006.\textsuperscript{61}

D. The Rise of Digital Distribution

The earliest example of digital distribution for home consoles in the United States can be traced all the way back to the Atari 2600 and the GameLine download service launched in 1983.\textsuperscript{62} Using a special cartridge with a built-in modem, users could download Atari 2600 games over their phonelines.\textsuperscript{63} Unfortunately for GameLine, the launch of the service came near the end of the 2600's life and coincided with the 1983 collapse of the video game market, leading to the service’s quick death and relative obscurity.\textsuperscript{64} Over a decade later, Sega would launch the similarly ill-fated Sega Channel, which downloaded games to the Genesis through a local cable provider.\textsuperscript{65}

Despite these inauspicious beginnings, digital distribution has come to be a major force in the market,\textsuperscript{66} and its proliferation has strong implications for game preservation. Each of the “big three” modern home console

\textsuperscript{58} Guttenbrunner et al., supra note 31, at 70. “Optical media” includes storage mediums such as CDs, DVDs, and Blu-ray discs. Id. at 73.

\textsuperscript{59} KENT, supra note 1, at 451.

\textsuperscript{60} Guttenbrunner et al., supra note 31, at 70–71. A newcomer to the market, Sony controlled 47% of the console market by August 1997. KENT, supra note 1, at 558.

\textsuperscript{61} Guttenbrunner et al., supra note 31, at 71–73.


\textsuperscript{63} Id.

\textsuperscript{64} Id.


\textsuperscript{66} See ESA 2018 REPORT, supra note 4, at 11 (showing that 79% of digital game sales in 2017 were in a digital rather than physical format). Note, however, that this statistic includes full games, subscriptions, add-on content, mobile apps, and social network games.
manufacturers—Nintendo, Sony, and Microsoft—offer robust digital sales platforms and subscription services, and an increasing number of games are receiving “digital-only” releases.

In this new digital-only environment, many game developers have shifted from a “games as product” approach to a “games as service” approach. Unlike with the “games as product” approach, where a purchaser has access to the entire game out of the box and never needs to worry about whether they have an internet connection, the games as service model relies on web-based infrastructure so that games can be “continuously updated.” Under the games as service model, the content of a single game can drastically morph and shift over time, and access to all or a portion of the game may rely on a steady internet connection—but once the servers providing that content are taken down, all of that content essentially disappears. A popular element of this approach includes “server-side authentication,” which requires a game to connect to the developer’s server in order to run. Once the authenticating server is taken down, the game can be rendered unplayable or certain portions of the game may become unavailable.

Even if a buyer opts for a physical copy of a single-player game with no meaningful network features, the games as service model means that many of

---


68 In response, this trend of digital-only releases has led to a paradoxical cottage industry of small print publishers producing limited physical releases of many of the more popular digital-only digital games. See, e.g., About SRG, SPECIAL RES. GAMES, https://specialreservegames.com/pages/about-us [https://perma.cc/64LB-GPBA]; About Us, LIMITED RUN GAMES, https://limitedrungames.com/pages/about-us [https://perma.cc/UZB3-T2ZH].

69 Chris Kerr, Games as a Service Drives Huge Market Value Spike for EA, Activision, GAMASUTRA (Oct. 19, 2018), https://www.gamasutra.com/view/news/328999/Games_as_a_service_drives_huge_market_value_spike_for_EA_Activision.php [https://perma.cc/CCA3-352B] (showing that this approach to games as service “has been fueling growth over the past six years” with major publishers).


71 See id.; Kyle Orland, Game Industry Pushes Back Against Efforts to Restore Gameplay Servers, ARS TECHNICA (Feb. 21, 2018), https://arstechnica.com/gaming/2018/02/preservation-or-theft-historians-publishers-argue-over-dead-game-servers/ [https://perma.cc/M78X-5R23] (covering the industry’s resistance to consumer created workarounds for such servers, even after the official servers have been taken down).

72 DIGITAL RIGHTS MANAGEMENT: THE LIBRARIAN’S GUIDE 54 (Catherine A. Lemmer & Carla P. Wale eds., 2016); Orland, supra note 71.

73 Orland, supra note 71.
the games on the shelves today are unfinished products. As of this generation, even physical copies of major releases are frequently incomplete and require “Day-One patches” to finish features or fix major bugs. Some games may even lack major story elements which need to be patched in. Any archive that simply preserves a physical copy of a modern game—or the code copied from that physical copy—may only hold a fragment of the complete game. If and when the online services providing the necessary patches and downloadable content disappear, those parts of the game will no longer be accessible, and the copies held in archives will be forever be incomplete or even inoperable.

III. THE CHALLENGES FACING GAME PRESERVATION

The challenges facing digital game preservation can be sorted into two major categories. The first are the technical issues necessitating preservation in the first place, stemming largely from media degradation and obsolescence. These are the harsh physical truths that, without intervention, can eventually render games unplayable. The second category stems from the legal impediments that prevent or complicate preservation, often stemming from industry stances and practices either directly or indirectly adverse to preservation efforts.

A. Technical Challenges and the Need for Preservation

1. Media Degradation

Time will take its toll on all things, and the media used to store the code of digital games is no exception. For the most part, digital games—especially older games—were not designed to last forever. Game cartridges are subject to corrosion from moisture and battery acid, while optical media can be rendered unplayable.

74 See Rami Ismail, Why ‘Day-One Patches’ Are So Common, KOTAKU (Aug. 8, 2016), https://kotaku.com/why-day-one-patches-are-so-common-1784967193 [https://perma.cc/UHQ7-XASC] (presenting one game developer’s insight into the many reasons why modern games go to print before they are technically finished); see also, e.g., Owen S. Good, No Man’s Sky’s Day-One Patch, Available Now, Dramatically Changes the Entire Game (Update), POLYGON (Aug. 7, 2016), https://wwwpolygon.com/2016/8/7/12397380/no-mans-sky-day-one-patch [https://perma.cc/5HUF-9F66] (describing how, immediately upon release, the intensely hyped No Man’s Sky required an update that overhauled nearly every aspect of the game).

75 See, e.g., Jason Schreier, Kingdom Hearts III’s Epilogue Will Be Patched in After Launch, KOTAKU (Jan. 11, 2019), https://kotaku.com/kingdom-hearts-iiis-epilogue-will-be-patched-in-after-l-1831669786 [https://perma.cc/P6U3-CLR6] (reporting that several parts of Kingdom Heart’s III’s story were to be patched in after release).

76 See supra Part II.

77 See supra Part II.B for a brief description and history of programmable cartridges.

78 See supra Part II.C for a brief description and history of games on optical media.
unreadable by pits, scratches, and chemical damage from inks, adhesives, and other materials. But beyond just simple wear and tear, different storage mediums have different lifespans, and each has an expiration date, after which the data stored on them becomes lost or unreadable.

Games stored on cartridges face what is referred to as “bit rot.” Over time, the individual bits of data stored on the ROM chips that hold the game’s code can “flip,” resulting in the “gradual and natural decay” of the data over time, which can then render the game unplayable. Most “masked ROM” cartridges are relatively safe from bit rot, and while the general lifespan of masked ROM chips is unknown, it is expected that many will outlive the copyrights preventing their legal duplication. On the other hand, cartridges using “EPROM” (Erasable Programmable ROM) only have an expected lifespan of twenty-five years, with many of the earliest EPROM cartridges already showing signs of bit rot.

Optical media faces its own version of bit rot in the form of “disc rot.” Disc rot can come in several forms, including the appearance of pin-hole specks in the data layer of the disc, “edge-rot,” and “bronzing”—each of which can quickly render a disc unreadable. While the prevalence of disc rot varies depending on manufacturing techniques and environmental factors, older discs

---

79 Devin Monnens, Losing Digital Game History, Bit by Bit, in BEFORE IT'S TOO LATE, supra note 18, at 142.
80 For a general ranking of how long different media formats are expected to last before becoming unreadable, see Media Stability Rankings, MUSEUM OBSCOLETE MEDIA, https://obsoletemedia.org/media-preservation/media-stability-ratings/ [https://perma.cc/GNW6-9ANB].
82 Id.
83 Id. at 142.
84 Id.
85 Id. at 142–43. A rare smattering of arcade games also has a security feature which places some code on battery-powered SRAM, causing the game to “commit suicide” once the battery eventually dies. Id.
86 Id. at 142–43.
88 Smith, supra note 87.
seem to be especially at risk, with Sega CD (1992), Sega Saturn (1995), and even Sega Dreamcast (1999) games already beginning to show signs of rot.89

2. Obsolescence

As technology moves ever forward, digital game preservation also faces the threats and complications of technological obsolescence.90 Even without the threat of media degradation, a game’s data can rapidly become unplayable in the latest hardware and software environments.91 Meanwhile, “the high costs of maintenance and inevitable failure of computer components” makes the long-term preservation of most old hardware unfeasible.92 A long-term preservation strategy must consider the challenges posed not just by the hardware and software environments of the next few years, but the hardware and software environments decades or even centuries down the line.

To combat obsolescence, many archivists argue that source code is vital for long-term preservation.93 For preservation purposes, the source code of a digital game is the “equivalent of the original camera negative of a film”94 or a digital “manuscript.”95 The source code contains the raw assets of the game itself and

89 Id.
90 Monnens, supra note 79, at 146.
91 Id. at 143 (finding that “[d]igital games are particularly susceptible to media obsolescence” due to the rapid evolution of technology and general lack of backwards compatibility between home consoles); see also Matt Bertz, The Digital Archaeologists: How GOG.com Rescues Games from the Dustbin of History, GAMEINFORMER (Mar. 7, 2019), https://www.gameinformer.com/2019/03/07/the-digital-archaeologists-how-gog-com-rescues-games-from-the-dustbin-of-history [https://perma.cc/L5CH-34HN] (reporting on Good Old Games (GOG), a company that focuses on the preservation and sale of computer games from the ‘80s and ‘90s; GOG employs a team of experts to update each game to run on new computers and claims that “for the more complicated titles, the [modernization] process can take up to six months”).
92 Monnens, supra note 79, at 144.
93 Luca Taborelli et al., Video Games You Will Never Play 44 (2016) (quoting Jeremy Thackray, assistant curator at the Centre for Computing History, in an interview: “It’s one thing to collect a retail copy of a game, but for long term preservation the source code is vital. Big museums are starting to latch on to this: the Museum of Modern Art in New York has made it a priority to collect source code for their design collections, for instance.”).
can facilitate adaptation from one software environment to the next without necessarily needing to replicate the original hardware environment.\footnote{96See Trevor Owens, Duke’s Legacy: Video Game Source Disc Preservation at the Library of Congress, LIBR. CONGRESS (Aug. 6, 2014) (reporting on the accidental submission of the source code of an unreleased game to the Library of Congress and the process taken to access the game’s data); see also Alexandra J. Horne, Comment, Shared Rights to Source Code: The Copyright Dilemma, 32 SANTA CLARA L. REV. 497, 505 (1992) (“A computer software program is developed by a human being . . . using source code. Although source code is understandable to a human, it is not intelligible to a computer. Therefore, in a simplified case, a two-step process is used to translate the source code into a machine-readable, binary language of ones and zeros.”).}

Without source code, the only other option for preservation in new hardware and software environments is to create “emulators” to replicate the original systems.\footnote{97See Sho Kikugawa, How Emulators Work, PC GAMER (Feb. 5, 2016), https://www.pcgamer.com/how-emulators-work/ [https://perma.cc/A367-Q5S9] (explaining the basic technical aspects of emulators).} Emulators are pieces of software which attempt to mimic the physical hardware of a specific console or computer system.\footnote{98Michael Wahba, The Bits and Bytes of Video Game Preservation, SCHOLARLY GAMERS (Nov. 9, 2018), https://www.scholarlygamers.com/feature/2018/11/09/the-bits-and-bytes-of-video-game-preservation/ [https://perma.cc/6ECD-WTE5].} Essentially, they are programs pretending to be physical devices.\footnote{99Id.} Building effective emulators is difficult and even when they manage to run a game, there can often be severe technical issues.\footnote{100See Charlotte Thai, How to Give Cartridge-Based Video Game Data an Extra Life, HOW THEY GOT GAME (Oct. 24, 2013), http://web.stanford.edu/group/htgg/cgi-bin/drupal/?q=node/1179 [https://perma.cc/BZT5-YSHQ] (presenting some of the technical issues faced when attempting to archive and emulate cartridge-based games from the Cabrinety Collection); see also Wahba, supra note 98 (“[A]s consoles get more and more complex, developing emulators similarly requires much more work.”).} The development process is further complicated by the fact that “the exact specifications of console video game systems and development documentation for game developers are usually confidential,” meaning that emulator developers are often flying in the dark.\footnote{101Guttenbrunner et al., supra note 31, at 72–73 (also noting that “[w]ithout the support of the manufacturer, it can be difficult to preserve a video game of a particular system”).} Cartridge-based games can further complicate matters if the cartridge itself contains special hardware that also needs to be separately emulated or otherwise accounted for.\footnote{102See, e.g., Orland, supra note 39 (noting the difficulties posed by emulating SNES games that used the SuperFX chip).} The digital games industry is often very protective of source code and trade secrets laws often keep source code from ever reaching the public.\footnote{103See infra Part III.B.2.} Despite this, many individual publishers and developers do not take good care of their source code, and there have already been cases of major publishers losing the source code to high-profile games and severely hindering even their own ability...
to adapt a game from a single generation of consoles to the next.\textsuperscript{104} While there is evidence that many of the largest developers and publishers have processes for preserving their own source code,\textsuperscript{105} even the most successful companies cannot be expected to last forever, and such preservation may not always survive the sale or dissolution of a company.\textsuperscript{106} Further, as leadership changes, executives of larger game companies may not always respect or understand the value of preserving their back catalogue—especially when there may be no profit motive in preservation.\textsuperscript{107}

\textsuperscript{104}See, e.g., Anthony John Agnello, \textit{The Problem with Preservation}, 1UP (May 14, 2012), https://web.archive.org/web/20150606214308/http://www.1up.com/features/the-problem-with-preservation [https://perma.cc/XR2Z-H559] (revealing that game publisher Konami lost the final source code to seminal survival horror games \textit{Silent Hill} 2 and \textit{Silent Hill} 3, one of many factors leading to the buggy and widely criticized HD remakes only a single console generation after their original releases); Alex Donaldson, \textit{Why Isn’t There a PS4, Xbox and Switch Port of Final Fantasy 8? Preservation May Be the Answer}, VG247 (Sept. 14, 2018), https://www.vg247.com/2018/09/14/isnt-ps4-xbox-switch-port-final-fantasy-8-preservation-may-answer/ [https://perma.cc/638F-TPVS] (describing how the major developer Square (now Square Enix) lost the source code to \textit{Final Fantasy VII}, a popular and massively successful entry in their flagship franchise, almost immediately after development. This greatly hindered their ability to port the game to personal computers (PCs) only a year after the initial release. After they rebuilt much of the game for the PC release, that new source code was also lost, meaning they had to “reverse-engineer and frankenstein” the game once again to release it for modern consoles); Chris Kerr, \textit{Square Enix Working to Preserve and Release Entire Game Library}, Gamasutra (June 13, 2019), https://www.gamasutra.com/view/news/344710/Square_Enix_working_to_preserve_and_release_entire_game_library.php [https://perma.cc/HGQ9-G6ZH] (quoting Square Enix president and chief executive, Yosuke Matsuda: “[I]n some cases, we don’t know where the code is anymore. . . . [B]ack in the day you just made [games] and put them out there and you were done—you didn’t think of how you were going to sell them down the road. . . . Sometimes customers ask, ‘Why haven’t you released that [game] yet?’ And the truth of the matter is it’s because we don’t know where it has gone.”).

\textsuperscript{105}John Andersen, \textit{Where Games Go to Sleep: The Game Preservation Crisis, Part 3}, Gamasutra (Mar. 10, 2011), http://www.gamasutra.com/view/feature/134671/where-games_go_to_sleep_the_game_.php [https://perma.cc/H92X-GDRE] (indicating that developers such as the “big three” hardware giants, Sony, Nintendo, and Microsoft claim to have their own preservation strategies in place).

\textsuperscript{106}See, e.g., \textit{id.} (revealing that when Mattel shut down game production in 1984, it went out of its way to archive its source code in the hopes that they would find a buyer for Intellivision; only a year after the purchase, the buyer was unable to read the non-standard disks the code had been saved to, and the code had to be retrieved through a long process of acquiring old hardware and transferring the code to a new format); 7800 Games & Development, \textit{Atari Museum}, http://www.atarimuseum.com/ videogames/consoles/7800/games/ [https://perma.cc/336J-KZZ8] (explaining that much of the source code currently available for Atari 7800 games only exists because it was rescued from dumpsters behind Atari’s California office when Atari shut down in 1996).

\textsuperscript{107}See, e.g., Matt Peckham, \textit{Everything Sony Told Us About the Future of PlayStation}, TIME (June 5, 2017), http://time.com/4804768/playstation-4-ps4-pro-psvr-sales/ [https://perma.cc/N57U-EJFC] (quoting Sony global sales chief Jim Ryan regarding the importance
development spectrum, many independent developers simply do not have the
time or resources to devote to long-term preservation.108

B. Legal Hurdles for Game Preservation

One of the unfortunate realities facing digital game preservation is how
often the games industry itself stands in direct opposition to preservation efforts.
While current copyright law is more favorable to game preservation than it has
been in the past, the volatility of copyright law around archiving games still
presents major issues. Meanwhile, trade secrets laws stand in the way of
archiving source code, and the new model of “games as service” and “always
on” DRM and server-side authentication present a host of new legal issues. One
reaction to these hurdles has been a grassroots movement to preserve games
through the private collection and distribution of ROMs—but such an approach
provides unstable and uncertain results, in addition to its illegality.

1. Copyright

One of the legal hurdles to transferring games from a dying format to a more
stable one is the Digital Millennium Copyright Act of 1998 (DMCA).109 One of
the DMCA’s main functions regarding digital games is to prevent users from
circumventing the devices that copyright holders use to protect games from
unauthorized copying.110 However, even if an archive looking to transfer game
data off of dying media had the time and resources to formally reach out and
request permission from each and every rightsholder (and optimistically
assuming that the rightsholders would always be willing to grant such

of backwards compatibility: “I was at a Gran Turismo event recently where they had PS1,
PS2, PS3 and PS4 games, and the PS1 and the PS2 games, they looked ancient, like why
would anybody play this?”); see also Lewis Gordon, Beware the Corporate Video Game
Canon, AV Club (Nov. 29, 2018), https://games.avclub.com/beware-the-corporate-
video-game-canon-1830684470 [https://perma.cc/8CG4-L8Y3] (discussing how game
manufacturers benefit from perpetuating a culture of planned obsolescence and restrictive IP
laws that prevent public access to the history of digital games, and arguing that within
Nintendo and Sony’s attempts to financially capitalize on growing public interest in their
back catalogues, “[t]he vast majority of the quirks, experiments, and other cultural artifacts
that fail to meet Sony’s and Nintendo’s own teleological narratives have been scrubbed out
of their histories”).

108 See TABORELLI ET AL., supra note 93, at 66 (quoting journalist and game historian,
John Anderson: “Many developers and publishers care about their legacies, but the feeling I
get from indie developers was ‘We have to worry about right now, not the past.’ I think that
echoes to major developers and publishers as well since this is understandably (and
obviously) such a competitive industry.”).


110 ASHLEY SAUNDERS LIPOW & ROBERT D. BRAIN, VIDEOGAME LAW: CASES,
archives would still struggle with a number of games where identifying and locating the rightsholders decades after its release is virtually impossible.\footnote{See Bertz, supra note 91 (finding that rightsholders are often uncooperative even when presented with financial incentives, pointing to diverse issues including “[b]itter divorcees uninterested in propping up their former partner’s legacy, relatives who have little to no interest in video games, and people too rich to bother”).}

Fortunately, a DMCA exemption currently allows libraries, archives, and museums to circumvent technological copyright protection measures for games that are “no longer reasonably available in the commercial marketplace, solely for the purpose of preservation of the game in a playable form”\footnote{37 C.F.R. § 201.40(b)(12)(ii) (2019).}—but that exemption remains vulnerable and faces continuous pushback from industry organizations. Exemptions to the DMCA must be renewed every three years,\footnote{Sarah Jeong, Why DMCA Rulemaking Is an Unsustainable Garbage Train, VICE (Nov. 3, 2015), https://www.vice.com/en_us/article/9a33wv/why-dmca-rulemaking-is-an-unsustainable-garbage-train [https://perma.cc/DXH7-GLLS] (covering how the triennial exemption process not only hinders game preservation but also disrupts security research).} leading to inconsistency and uncertainty. Prior to 2018, these exemptions had to be reestablished de novo after each three-year period, but as of October 2018, the U.S. Copyright Office streamlined the proceedings to allow for the automatic renewal of exemptions “to which there is no meaningful opposition.”\footnote{Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies, 83 Fed. Reg. 54,010, 54,011 (Oct. 26, 2018).}

The exemption for archives, libraries, and museums was first introduced in 2003.\footnote{37 C.F.R. § 201.40(b)(3) (2003) (expired 2006).} It disappeared in 2010, replaced with an exemption for “good faith testing, investigating, or correcting security flaws or vulnerabilities.”\footnote{37 C.F.R. § 201.40(b)(4) (2010) (expired 2012).} In 2012, there were absolutely no exemptions for digital games\footnote{See 37 C.F.R. § 201.40 (2012) (expired 2014).} in the wake of the Electronic Frontier Foundation’s (EFF) failed effort to obtain an exception for circumvention of access controls on game consoles.\footnote{See Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies, 77 Fed. Reg. 65,260, 65,272–74 (Oct. 26, 2012).} This was largely a result of “vigorous” opposition from the Entertainment Software Association.
An exemption was not reintroduced until 2015 to allow circumvention of authentication servers that had been taken down. This new exemption also faced resistance by the ESA. Finally, the current exemption allowing approved libraries, archives, and museums to circumvent copy protection for the preservation of digital games was reintroduced in October 2018—eight years after it had disappeared. This newest exemption also faced opposition when it was presented alongside a more general exemption for computer software, with opponents claiming that the overall proposal was “overbroad.” In 2021, the current exemption may again have to be reconsidered if the ESA or other groups oppose it, which would mean that preservationists could once again have to put valuable time and resources into retaining the exemption or else be thrust back into a complicated, legal grey space.

This legal uncertainty and adversarial positioning towards the industry is not conducive to long-term preservation efforts. At the moment, the Library of Congress does not transfer games off of their original media and seems to be nervous about doing so. However, some archives have begun to back up certain games. The Stephen M. Cabrinety Collection at Stanford University is in the process of ripping ROM files and ISOs for digital preservation, focusing on software from 1975 to 1995. Even further along the spectrum is

---


121 37 C.F.R. § 201.40(b)(8)(i) (2015) (allowing circumvention “when the copyright owner or its authorized representative has ceased to provide access to an external computer server necessary to facilitate an authentication process to enable local gameplay”).

122 Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies, 80 Fed. Reg. 65,944, 65,957 (Oct. 28, 2015) (stating that the ESA argued that the proposed exemption “was too broad, would not facilitate any noninfringing uses, and could adversely impact the market for video games”).


125 Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies, 83 Fed. Reg. at 54,011 (allowing for the automatic renewal of DMCA exceptions unless there is “meaningful opposition”).

126 See Interview with Dave Gibson, supra note 94 (“[R]ight now, at least the Library of Congress, we’re not going down the emulation road at all for a number of reasons, the main one being that the fact is copyright and we share—we’re under one roof together.”).

127 See Thai, supra note 100 (presenting the “method for retrieving forensically viable data from cartridge-based media in the Cabrinety collection”).

the Internet Archive, which has long claimed that copyright law allows it to archive game software and which also openly hosts a prodigious amount of copyrighted games for public use. The legality of some of the Internet Archive’s actions remains in question despite the organization’s ties to the Library of Congress, and it is worth noting that the Internet Archive has openly broken other, similar copyright laws for long periods of time. As long as the legality of these and other preservation efforts remain in question or could be undone by future changes to the exceptions, long-term game preservation stands on shaky grounds, and many archives may simply not wish to take on the legal risks without the explicit approval of rightsholders.

2. Source Code and Trade Secret Law

Frequently, companies seek to protect source code via trade secret law. However, the very nature of trade secret law requires companies to ensure that people outside of the company do not have access to the source code. Under the Uniform Trade Secrets Act (UTSA), secrets are “entitled to indefinite protection

---


130 See, e.g., Internet Arcade, INTERNET ARCHIVE, https://archive.org/details/internetarcade [https://perma.cc/CRB7-XZYV] (allowing users to play many copyrighted arcade games from their browsers, including Sega’s Out Run, Atari’s The Empire Strikes Back, and Capcom’s Street Fighter II).

131 See Dan Whitehead, Opinion, We Need to Talk About Emulation, EUROGAMER (May 31, 2015), https://www.eurogamer.net/articles/2015-05-31-we-need-to-talk-about-emulation [https://perma.cc/4DUU-2BWK] (arguing that despite the value that the Internet Archive provides, the community around digital games must sooner or later talk about the obvious copyright violations involved).


133 See Nate Hoffelder, SFWA Finally Notices Internet Archive’s Decade Old Open Library, Decides It’s Piracy, DIGITAL READER (Jan. 9, 2018), https://the-digital-reader.com/2018/01/09/sfwa-finally-notices-internet-archives-decade-old-open-library-decides-piracy/ [https://perma.cc/URG3-PS5U] (covering the recent legal issues surrounding the Internet Archive’s decade old Open Library, “where members can borrow digital copies of the print books that either the IA has archived in its warehouses, or a partner library has in its catalog”).

as long as [they are] kept a true secret.” Over forty-five states have adopted the UTSA, which defines a trade secret as:

[I]nformation, including a formula, pattern, compilation, program, device, method, technique, or process, that: (i) derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable by proper means by, other persons who can obtain economic value from its disclosure or use, and (ii) is the subject of efforts that are reasonable under the circumstances to maintain its secrecy.

When considering whether particular information is a trade secret, three of the major factors a court will consider are (1) the extent to which the information is known outside the business, (2) the extent of measures taken to guard the secrecy of the information, and (3) the difficulty with which others could acquire the information.

Companies often go to great lengths to ensure that their trade secrets remain confidential. If they succeed, one of the major benefits of trade secret protection to rightsholders (and one of the hurdles to preservationists) is that trade secret protection truly can be indefinite—so long as the secret is kept from the public, protection will not expire with time.

For these reasons, it is easy to understand why companies would be nervous about sharing their source code with archives and potentially losing their trade secret protection. Any long-term preservation strategy wishing to gain the cooperation of the digital games industry will need to work with companies to guarantee that any archived source code is kept sufficiently secret, at least so long as the source code remains meaningfully valuable to those companies.

3. **DRM and Server-Side Authentication**

The shift to digital distribution and “games as service” presents a host of new legal issues in game preservation and may mean that many newer “digital-only” games that are never released on physical media may be at a
paradoxically greater risk of being lost than many older games. Further, games that require a connection to a developer or publisher’s server to function or access certain features can quickly be rendered nonfunctional or incomplete once those servers are taken down. While this is already a large technical hurdle for preservation of many modern games, it is further complicated by additional legal hurdles.

A recent decision by the Librarian of Congress allows museums and archives to recreate servers for games after the official servers have been taken down, but only if (1) the museum or archive legally obtains the original server code (2) the museum does so without the purpose of obtaining a commercial advantage, and (3) the servers are only locally accessible on the museum or archive’s premises. Because of this, emulation or “reconstruction” of server code is not currently a valid exception to the DMCA. However, it is rare for companies to make server code available and the original server code is often lost as soon as a company takes the servers down, making this provision effectively useless in most cases.


143 See supra Part II.D.


145 Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies, 83 Fed. Reg. at 54,024; 37 C.F.R. § 201.40(b)(12)(iv)(A) (2019) (defining “complete games” for purposes of this regulation as those “that can be played by users without accessing or reproducing copyrightable content stored or previously stored on an external computer server”).

146 Ian Birnbaum & Matthew Gault, Copyright Law Just Got Better for Video Game History, VICE (Oct. 25, 2018), https://www.vice.com/en_us/article/zm9asz5/copyright-law-just-got-better-for-video-game-history (quoting John Hardie, Director of the National Videogame Museum in Frisco, Texas: “It’s very unlikely that anyone saved the server code . . . . In all the archiving we’ve done, we’ve never had a
This is another situation where the collaboration between rightsholder and preservationist is critical to long-term preservation. Not only does the law need to change to allow for circumvention in cases where the server code has been lost, but there needs to be channels and agreed upon best practices for preserving server code in the first place. However, the current legal environment is unlikely to foster such collaboration. Even this small exception was resisted by the ESA who, alongside other opponents, “contended that proponents wish[ed] to engage in recreational play that could function as a market substitute.”

4. What of Private Preservation Efforts?

While many of the legal issues revolving around the private creation, consumption, and distribution of ROMs fall outside the scope of this Note, it is sufficient to say that the private distribution of ROMs over the internet is “unambiguously illegal.” However, it is worth briefly mentioning the grassroots efforts many individuals are making to preserve games through the admittedly illegal collection and distribution of ROMs.

Some have begun to argue that without legal alternatives, the illegal collection and distribution of ROMs is the only available method for comprehensive and long-term digital game preservation. In addition to its
illegality, the major issue with this approach is the vulnerability of ROM sites
to takedown notices and other legal action, as well as the tremendous legal and
financial risks that such private preservationists might take on. This method
also does nothing to ensure that source code will become more accessible,
meaning that obsolescence and certain technical problems will remain pressing
issues. There is also the reality that much of this preservation is driven by
individuals with nostalgic and personal attachments to the games they are
preserving through these methods, and it is impossible to say whether future
generations will pick up the torch of private, illegal preservation or if all those
ROMs will simply fade into obscurity.

Between the legal issues and the individual nature of private preservation,
it is impossible to know how such methods will hold up over the generations. A
more collaborative and institutional solution both respects rightsholders and
encourages developers to take an active part in the preservation process.
However, game developers and publishers should take note of this grassroots
movement both as an indicator of general interest in game preservation and as
an untapped market who may be willing to pay for games they would otherwise
pirate due to inaccessibility.

---


153 See supra Parts III.A.2 & III.B.2 for discussion on source code.

154 See Patrick O’Rourke, Retro Video Game Collecting: An Industry Built on Nostalgia but Plagued by High Prices, FIN. POST, https://business.financialpost.com/technology/gaming/retro-video-game-collecting-an-industry-built-on-nostalgia-but-plagued-by-high-prices [https://perma.cc/BJ9M-5EQB] (last updated Dec. 12, 2014) (noting that the growing market for retro games is driven largely by nostalgia); Pot, supra note 150 (noting that private collectors are often “[p]eople who obsessively scan eBay for obscure games, then buy and preserve them”).

155 See Maaz, supra note 152 (arguing that Nintendo’s aggressive legal action against ROM and emulator sites has had a deterrent effect which “is swiftly eroding a large chunk of retro gaming”).

156 For one exploration of this later idea, see generally Jethro Dean Lord IV, Comment, Would You Like to Play Again? Saving Classic Video Games from Virtual Extinction through Statutory Licensing, 35 SW. U. L. REV. 405, 405 (2006) (arguing that the digital games industry should “adopt the principle of compulsory licensing from the music industry and encourage Congress to pass legislation creating a statutory licensing system for video game ROMs”).
C. Current State of Game Preservation

Digital game preservation efforts are currently underway in some public and nonprofit archives, but the preservation movement has a long way to go. As of 2016, the Library of Congress had only approximately 3500 games in its archive. The collection is comprised entirely of physical media, and the Library’s preservation activities are limited to storing this media in temperature- and humidity-controlled environments. Despite these controls, the games remain vulnerable to media degradation, but the legal uncertainty surrounding extracting the data from the games for long-term preservation continues to prevent the Library of Congress from taking further steps.

In the United States, there are several nonprofit archives and museums that are also beginning to undertake the mission of long-term game preservation. Some of the major efforts include the Internet Archive, the Stephen M. Cabrinety Collection, and the Video Game History Foundation. In addition to helping to clear legal hurdles in the way of preservation, a national board would present these individual programs with a centralized hub for coordination and the development of universal best practices.

IV. Establishing a Federally Funded Preservation Board

When considering the preservation of digital games, it is important to remember that this is not the first time society has needed to consider whether to preserve the historical foundations of a new artistic medium and how to go about doing so. The history of film preservation presents a cautionary tale of just how quickly irreplaceable pieces of our culture and history can be lost when industry figureheads and legislators fail to take proper action. However, the eventual establishment of the National Film Preservation Board (NFPB) and the collaborative approach to film preservation in place today provides a framework.
for how to ensure the long-term preservation of digital games. By adopting the modern approach towards film preservation facilitated by the NFPB, digital games can avoid having to reinvent the wheel of media preservation and prevent the sort of losses presented by the early days of film.

A. Film Preservation: A Case Study

“Making pictures is not like writing literature or composing music or painting masterpieces. The screen story is essentially a thing of today and once it has had its run, that day is finished. So far there has never been a classic film in the sense that there is a classic novel or poem or canvas or sonata. Last year’s picture, however strong its appeal at the time, is a book that has gone out of circulation.”

–Drama critic Edwin Schallert in 1934

The failures of early film preservation present a cautionary tale about how easily the foundation of a new medium can be lost when creators, policymakers, and the public fail to take action. Like games, film is relatively young in comparison to most other major artistic mediums such as visual art or literature. Both games and film share features of both works of human expression and consumer products, and “are part of a broader media landscape that exists in the contested space between artistic freedom and economic incentives.”


167 Conway & deWinter, supra note 26, at 2.
standards and the design language of modern games,\(^\text{168}\) the era of the American silent film “established the language of modern cinema.”\(^\text{169}\)

Each early film is a piece of our shared cultural history that we can never recapture, which makes the truth that much more difficult: Nearly all those films have vanished. Estimates vary, but around 90% of all American silent films and 50% of American sound films made before 1950 are irretrievably lost.\(^\text{170}\) These early films were destroyed through a mixture of negligence, natural decomposition, and corporate policies that dictated the willful destruction of older films.\(^\text{171}\) All this is in spite of the fact that, even in the earliest days of film, voices from within the industry were already making the case for film preservation.\(^\text{172}\)

However, not all is doom and gloom. Today, film preservation is facilitated by collaboration between countless public and nonprofit archives and research centers both in the United States and around the world, including the Library of Congress, the Museum of Modern Art, the Academy Film Archive, the George Eastman House, and the UCLA Film & Television Archive.\(^\text{173}\) Through collaboration between rightsholders, legislators, and public and nonprofit archives, we have made the decision to preserve this vital art form for future

---

\(^{168}\) See Taborelli et al., supra note 93, at 80 (interviewing game designer William Novak, who discussed the difficulty of figuring out how to deal with cameras in the development of early 3D games, and how Super Mario 64’s “simple, elegant, and intuitive” solution of dragging the camera behind the player character inspired other designers); id. at 37 (quoting game designer Art Min on building the controls for System Shock: “Configurable key controls were not common and we had this interesting gameplay with learning around corners. We were the first with that tech in games, so we didn’t have some thing [sic] we could model after.”).

\(^{169}\) Pierce, supra note 164, at 5.

\(^{170}\) Dave Kehr, Film Riches, Cleaned Up for Posterity, N.Y. Times (Oct. 14, 2010), https://www.nytimes.com/2010/10/15/movies/15restore.html [https://perma.cc/EWT2-RM4F]; see also Pierce, supra note 164, at 21 (“Only 14% of American silent feature films (1,575 of 10,919 titles) survive as originally released in complete 35mm copies. Another 11% (1,174) also survive in complete form, but in less-than-ideal editions—foreign release version or small gauge-formats such as 16mm.”). But compare Anthony Slide, Nitrate Won’t Wait: Film Preservation in the United States 5–6 (1992) (claiming that “[n]o one really knows just how many films have survived and how much has been lost,” and stating that the number of surviving films may be “a lot higher”).

\(^{171}\) Pierce, supra note 164, at 22 (noting that MGM was the only major studio of the silent-film era that saw long-term value in preserving its entire library, whereas studios like Universal-International engaged in the “willful disposal” of its entire library).

\(^{172}\) See Slide, supra note 170, at 9–10 (presenting a 1906 editorial in industry magazines suggesting that, “[p]erhaps the day will come when motion pictures will be treasured by governments in their museums as vital documents in their historical archives”).

generations. These modern efforts to preserve film present a hopeful foundation for how to approach game preservation.

1. Early Film Preservation Efforts (Late 1800s–1988)

As with games today, the Library of Congress’s early involvement with film was haphazard, characterized by fits and starts. The earliest films can be traced back to the 1870s–1880s. In turn, one of the earliest known attempts at any sort of film preservation came in 1893, when one of Thomas Edison’s assistants registered several “Kineographic Records” for copyright protection at the Library of Congress. Those records were soon after lost.

Until the Townsend Act was passed in 1912, motion pictures were not subject to copyright protection. Instead, copyright owners had to register their films as paper-printed collections of still photographs, which were then held by the Library of Congress. Despite the eventual change in the law, many copyright owners continued this practice of registering paper-printed stills as late as 1916. These paper prints were eventually sealed away in a vault, where they were left to molder until their rediscovery in the late 1930s.

After the Townsend Act, film companies began to submit actual nitrate film for copyright. However, due in large part to the flammability of the film, the Library of Congress would then return the reels themselves to the rightsholders and only retained “descriptive material” about the film.

---

174See supra Part III.C.
175WHEELER WINSTON DIXON & GWENDOLYN AUDREY FOSTER, A SHORT HISTORY OF FILM 4–5 (2008) (citing the earliest example of a “movie” being filmed as occurring in 1878, consisting of several frames of a horse galloping, shown in quick succession. Meanwhile, the “first truly portable moving picture camera” was invented in 1882).
176SLIDE, supra note 170, at 36.
177Id.
180Id.
181SLIDE, supra note 170, at 38.
182Id. at 37–38.
183Allen, supra note 179.
184SLIDE, supra note 170, at 1, 3 (stating that nitrate film “had two major drawbacks (at least from today’s viewpoint): it was highly flammable and it would eventually decompose. . . . Because nitrate film is chemically unstable, it is in a perpetual state of decomposition. Such decomposition may take less than two decades or more than two generations. . . . If kept at a low temperature and humidity, it can be stored safely and indefinitely.”).
185Allen, supra note 179.
The earliest governmental efforts at preserving the film itself began in the mid-1920s. This preservation was limited to news reels, government films, and other records of historical events, rather than what most people today would think of as “movies.” In 1934, the National Archives and Records Service (NARS) was brought into being to establish policies and procedures for managing these records.

The preservation of films as we typically think of them did not get its start as a movement until the Museum of Modern Art began its film library in 1935. The museum had considered engaging in film preservation efforts as early as 1929, but much as many today are still hesitant to grant the artistic merits of digital games, the museum’s trustees at the time were nervous about taking the “then-revolutionary step of endorsing motion pictures as an art form.”

It was not until 1942—nearly fifty years after accepting Thomas Edison’s Kineographic Records—that the Library of Congress first “recogniz[ed] the importance of motion pictures to the historical record,” and began to request the return of actual nitrate reels for “selected works, including films made before 1942.” After consideration of 4398 feature films, newsreels, documentaries, and other “short subjects,” the Library acquired and retained 971 total films in 1945. During that process, the Motion Picture Collection was established in 1944, and was later renamed the Motion Picture Division in 1946.

Unfortunately, this progress was short lived. In 1947, the Library of Congress ran afoul of the film industry when the Library suggested that it become involved in film distribution. Much like how the modern games industry often works against digital game preservation efforts, claiming that preservation may “adversely impact the market for video games” or “function as a market substitute,” the film industry in 1947 felt threatened by the public having more open access to these older films. In response, the film industry pressured the House Appropriations Committee to eliminate the Motion Picture

---

186 SLIDE, supra note 170, at 26.
187 See id.
188 Id.
189 See id. at 18 (“The serious work of film preservation in the United States can be dated from the founding of the Film Library of the Museum of Modern Art in 1935.”).
190 Id.
191 Allen, supra note 179.
192 SLIDE, supra note 170, at 39.
193 Id.
194 Id. at 41.
197 See SLIDE, supra note 170, at 41.
Defunded, the Motion Picture Division was closed, its staff dismissed. After the Division’s closure, the films that the Library of Congress had already collected were only inspected “on a casual basis,” and any that showed signs of deterioration were simply thrown away. Over the next decade, half of the films in the collection were lost.

The Library of Congress was not able to return in earnest to film preservation until 1965. In 1968, the Library of Congress began a collaboration with the private American Film Institute (AFI) in order to obtain a “comprehensive collection of significant American films,” renaming the Library’s archive, the “American Film Institute Collection at the Library of Congress.” This collaboration attempted to garner public interest in film preservation beginning in the early 1970s, but these early attempts were largely unsuccessful. The general public simply did not seem interested—but that was about to change.

2. Establishment of the National Film Preservation Board

After facing a nearly a decade of public disinterest, the American Film Institute made a major publicity push for film preservation, declaring 1983–1993 to be “The Decade of Preservation.” Rather than working against preservation efforts, the film industry threw its weight behind this new push, holding “star-studded” parties, raising hundreds of thousands of dollars in donations, and even drawing the attention and support of former film star President Ronald Reagan.

The legislative result of this surge of attention was the 1988 Film Preservation Act. The Act explicitly recognized that films are “an indigenous American art form . . . emulated throughout the world,” that they “represent an enduring part of our Nation’s historical and cultural heritage,” and that they are a “significant American art form deserving of protection.” The Act directed the Librarian of Congress to establish the National Film Registry “for the purpose of registering films that are culturally, historically, or aesthetically

198 Id.
199 Id.
200 Id.
201 Id.
202 Id. at 42–43.
203 SLIDE, supra note 170, at 76–77.
204 Id. at 78.
205 See id.
206 Id. at 86.
207 Id. at 87.
209 Id.
It also ordered the Librarian of Congress to establish the National Film Preservation Board (NFPB), to be comprised of members selected from select industry organizations, academic institutes, and trade guilds, who were then tasked with selecting the films to be included in the National Film Registry.

The 1988 Act was set to expire after three years, but was renewed in 1992. As part of this renewal, the Librarian of Congress and the NFPB were required to “conduct a study on the current state of film preservation and restoration activities, including the activities of the Library of Congress and the other major film archives in the United States.” Following that study, the Librarian of Congress was then required to “establish a comprehensive national film preservation program . . . in conjunction with other film archivists and copyright owners.”

The results of this study were catalogued in a series of reports released in 1993 and 1994. The study ultimately concluded that films of all types were “deteriorating faster than archives [could] preserve them.” Further, the system of funding at the time was an ineffective “patchwork of federal money, institutional outlays, foundation grants, and private donations,” with archives working with increasingly shrinking budgets. In order to ensure the long-term preservation of film, the Board stressed the importance of private-public partnerships between studios, industry leaders, and public and nonprofit archives, firmly stating that “[s]olving America’s film preservation problems [was] beyond the resources of any single institution.”

To help facilitate this collaboration, the 1994 report included the first of many supporting documents outlining best practices for film preservation.

210 Id.
211 Id. at 1785–87 (identifying such organizations as the Academy of Motion Picture Arts and Sciences; the Department of Theatre, Film, and Television, College of Fine Arts at the University of California, Los Angeles; and the Directors Guild of America).
212 Id. at 1788.
214 Id.
215 Id.
218 Id. at 1.
219 Id. at 21.
220 Id. at 21–25.
221 Id. at xi.
222 See KEEPING COOL AND DRY: A NEW EMPHASIS IN FILM PRESERVATION, in NAT’L FILM PRES. BD., supra note 217, at 33 (outlining a “coherent and cost-effective” approach to film preservation).
establishing voluntary guidelines for studios and archives, and presenting a legal guide to help facilitate collaboration between studios and archives. In 1996, to further help facilitate the goals of the National Film Preservation Board, Congress created the National Film Preservation Foundation, a private “nonprofit charitable affiliate” of the NFPB charged to “encourage, accept, and administer private gifts to promote and ensure the preservation and public accessibility of the nation’s film heritage.”

The NFPB has since been renewed multiple times and is currently authorized through fiscal year 2026. In addition to selecting films for inclusion in the National film registry, the board continues to promote collaboration between rightsholders and preservationists, and provides information and resources to help preservationists adapt to changes in technology and intellectual property laws. While not everything involving the NFPB can (or should) be adopted in creating a long-term preservation strategy for digital games, the example set by the Board and its successes presents a powerful foundation for digital game preservation to build upon.

### B. Establishing a Federally Funded Preservation Board

When the NFPB released its national plan for film preservation in 1994, it acknowledged that “[s]olving America’s film preservation problems [was] beyond the resources of any single institution.” Any meaningful preservation effort required coordination between the film industry, multiple archives, and federally funded initiatives. It is no different for digital games. Industry buy-

---


226 Why the NFPF Was Created, NAT’L FILM PRESERVATION FOUND., https://www.filmpreservation.org/about/why-the-nfpf-was-created [https://perma.cc/Q8EH-ND9Y].


229 See, e.g., National Film Preservation Board: Resources, LIBR. CONGRESS, https://www.loc.gov/programs/national-film-preservation-board/resources/ [https://perma.cc/QE45-X27A] (presenting a host of resources related to film and film preservation, such as recent legislation impacting film preservation, copyright issues related to film, preservation resources and projects, and past Board initiatives).

230 NAT’L FILM PRES. BD., supra note 217, at xi.

231 See supra Part IV.A.
in is crucial both to obtain materials for preservation and to present a united front in convincing legislators and the public that digital game preservation is a worthwhile endeavor. At the moment, industry organizations often stand opposed to the most important steps towards long-term game preservation but that does not need to be the case.

A national board can establish a collaborative framework in which industry organizations such as the ESA can work with rather than against archivists to achieve a truly long-term preservation strategy. It can provide a common ground for people who genuinely care about the future of digital games to work together toward common goals that can serve to benefit both the public and the industry. Such a board should consist of representatives from archives and museums (such as the Library of Congress and the Internet Archive), academia (including scholars involved in both design and critical studies), major and minor publishers and developers, industry associations, legal experts, and other closely related groups. Among its priorities, the board must (1) develop a coherent plan to ensure the long-term preservation of source code, when that source code is still available, (2) establish methods to ensure that the games archived are the most “complete” versions available, (3) agree on a collective stance regarding the ability of libraries, archives, and museums to circumvent copyright protection measures in order to archive games for which the source code is no longer available, and (4) establish uniform model agreements and general best practices for digital game preservation, in order to ensure that preservation efforts are effective and uniform across archives.

1. Establishing and Maintaining a Source Code Archive

One of the most important priorities of the board should be the creation of a source code archive, likely through collaboration with the Library of Congress. In addition to the certainty promised by the Library of Congress’s general longevity, the use of a government agency may assuage the fears of many companies who might be worried about putting their source code in the hands of private organizations.

232 See, e.g., supra Part III.B.1 (covering the industry’s resistance to copyright exceptions for archives, libraries, and museums).

233 This is especially true of the “big three” hardware giants—Microsoft, Nintendo, and Sony.

234 Such as the ESA and the International Game Developers Association (IGDA).


236 See supra Part III.B.2 for a discussion how source code is protected under trade secret law and how the law requires the entity seeking trade secret protection to keep the source code “actually secret.”
The preservation of source code allows for the most complete preservation of digital games, but it also stands to provide economic benefits to the rightsholders. Whether this archive is entrusted to a single institution or delegated amongst several “approved” organizations (as is currently done with film), such an archive presents an opportunity for both archivists and rightsholders to lighten their own loads. Properly executed, it presents rightsholders with a convenient and secure way to maintain their intellectual property for future rereleases and can serve as an affordable, archival “back-up” of their projects. A source code archive has the added benefit of serving as a guaranteed source of evidence in the event of legal disputes involving source code and intellectual property infringement. In an era of constant retro rereleases and HD remakes, rightsholders are primed to see the commercial benefit in offloading much of the work of preservation onto nonprofit and public archives who would be more than happy to maintain their source code.

A source code archive also provides a necessary safety net for digital-only games and can help ensure that the discontinuation of a specific digital platform does not mean that the game is lost forever. With digital-only games, archives do not even have the option of preserving a physical copy. For a digital-only game tied to a specific platform, the only option is to download it to the system in question, leaving it inexorably tied to that piece of hardware and locked.

237 See supra Part III.A.2 for an overview of the importance of source code for long-term digital game preservation.
238 See Public Research Centers & Archives, supra note 173 (listing “a comprehensive listing of Public Moving Image Archives and Research Centers from around the world”).
239 See Devin Monnens, Why Are Video Games Worth Preserving?, in BEFORE IT’S TOO LATE, supra note 18, at 148 (“Digital games are owned, as intellectual property, by the companies that design, produce, and distribute them. . . . [That] ownership [is] jeopardized if the assets in question are too easily lost.”).
240 See Zach Vowell, What Constitutes History?, in BEFORE IT’S TOO LATE, supra note 18, at 152 (arguing that archives not only help to preserve materials but have the added benefit of helping rightsholders “work through issues of intellectual property and other legal matters that may arise”).
241 See, e.g., Gilbert, supra note 20 (covering the recent fad of miniature versions of classic game consoles pre-loaded with older games); Steven Strom, The Diminishing Value of Gaming’s “HD Remakes,” ARS TECHNICA (Apr. 13, 2015), https://arstechnica.com/gaming/2015/04/the-diminishing-value-of-gamings-hd-remakes/ [https://perma.cc/X7WF-Z3U9] (arguing that while the trend of “HD remakes” of games serves an important function for game preservation, many of the remakes being released in 2015 were for games that had “been out for as little as 10 months” and functioned as quick cash grabs as home consoles moved from one hardware generation to the next).
242 See supra note 142 and accompanying text for examples of contemporary digital-only games which are no longer available to the general public due to being removed from distribution platforms or due to entire distribution platforms being discontinued.
behind layers of DRM. Access to the source code and the ability to protect the code from degradation helps ensure that a digital-only game won’t disappear forever once it is delisted or its digital storefront is closed down.

Finally, any source code archive that is established must have an established procedure for updating the source code on file, rather than simply archiving the source code at launch. As mentioned above, physical copies of modern games are often incomplete upon release. Further, even those games that are complete and polished products upon release can often add layers of updates, expansions, and general bug fixes over time. The cooperation and collaboration of developers and publishers will be vital to maintaining such a system.

Due to the nature of trade secret laws, public access to source code archives is likely to be a contentious issue. To bring major industry associations to the table and ensure that rightsholders maintain the value protected by their trade secrets, it is likely that access to the source code archive would need to be limited to rightsholders and approved archivists, at least until the expiration of the copyrights protecting the games—with possible exceptions for approved researchers and historians. Even then, some companies might balk at the idea of their source code ever becoming publicly available. Because trade secret protection would otherwise outlaw copyrights and other IP protections so long as the source code was kept secret, many companies may be extremely hesitant about giving up control. Ultimately, the Board would be responsible for brokering a deal with major rightsholders, drafting uniform agreements for the acceptance and protection of the source code and establishing agreed upon “release mechanisms” for when the source code may become publicly available.

Preventing public access to source code archives might frustrate many libraries, archives, and museums that would like to make such information available to the public sooner rather than later, but it at least ensures that (1) the source code is available so that the rightsholders may still use it for commercial

243 See Bode & Maiberg, supra note 151 (noting that, after a digital storefront is closed down, “if the games users bought . . . are not already downloaded, or if whatever storage device users put them on is destroyed, they’ll lose [the games] for good”).
244 See supra Part III.A.
245 See supra Part II.D.
246 Returning to No Man’s Sky as discussed in supra note 74, the developer has continued to release major updates and even complete overhauls of different systems, to the point where jumping from one major update to the next is “like landing on completely foreign terrain.” Gita Jackson, No Man’s Sky’s Latest Update Makes Everything New, KOTAKU (July 24, 2018), https://kotaku.com/no-man-s-sky-s-latest-update-makes-everything-new-1827844781 [https://perma.cc/M6T6-LV8W]; see also Gita Jackson, No Man’s Sky Is Getting a ‘New Multiplayer Experience,’ KOTAKU (Mar. 15, 2019), https://kotaku.com/no-mans-sky-is-getting-a-new-multiplayer-experience-1833308940 [https://perma.cc/7SMM-QJRC].
247 See supra Part III.B.2 for a discussion of how source code is protected under trade secret law and how the law requires the entity seeking trade secret protection to keep the source code “actually secret.”
rereleases, securing some limited form of public access to the games themselves, and (2) the source code is not permanently lost to history. While finding a solution to the question of more open public access to the history of digital games is important to the development of the medium and should be a point of discussion among the Board, it is a question that ultimately falls outside the scope of this Note.

2. Establishing a Collective Stance Regaining DMCA Exceptions

The current system of establishing and maintaining exceptions to the DMCA to allow libraries, archives, and museums to circumvent copyright protection measures, while recently streamlined, still introduces an unnecessary and disruptive amount of uncertainty into the archival process. Further, the process is a forum that pits archivists and the industry against each other, rather than encouraging them to work together towards a shared goal of long-term preservation. Source code is simply not available for some games. For those games, the code stored on the physical copies of the games is the only version available. And, unfortunately, those physical copies will not last forever. Archives must have the ability to transfer the data from dying media to more stable media, and they need to be certain that their procedures are legal and will not be rendered illegal in three years’ time. Rather than potentially fighting each other every three years, a national board allows archivists and industry representatives to come together, establish shared goals, and come up with a cohesive stance regarding DMCA exceptions that allows for long-term preservation.

248 See Emanuel Maiberg, Nintendo’s Offensive, Tragic, and Totally Legal Erasure of ROM Sites, VICE: MOTHERBOARD (Aug. 10, 2018), https://motherboard.vice.com/en_us/article/bbjped/nintendos-offensive-tragic-and-totally-legal-erasure-of-rom-sites [https://perma.cc/8YLD-NAU6] (quoting game design professor Bennett Foddy: “As a professor, I very frequently see students spinning their tires trying to solve problems that were already solved in 1985. . . . And just as you would if you were teaching painting or music (or math), what you do as a teacher is you send them to the library to study the old classics, to see what they did right and wrong. That’s the only way we can make progress in the sciences, the humanities, or in the creative arts. . . . If I was teaching poetry, I could send a student to read nearly any poem written since the invention of the printing press, but in games my legal options limit me to, I would guess, less than 1 percent of the important games from history . . . .”).

249 See supra Part III.B.1 for an overview of the DMCA exception process and the history of exceptions for digital games.

250 Id.

251 See supra note 104 and accompanying text for examples of high-profile digital games where it is known that the source code has been lost.

252 Id.

253 See supra Part III.A.1 for an overview of media degradation.

254 See supra Part III.B.1 (explaining that exceptions to DMCA must be renewed every three years).
The board must also address the issues presented by the shift towards digital distribution and the “games as service” model. The current exceptions allowing the circumvention of servers that have been taken offline—but only if the archive in question has access to the original server code—ignores the fact that archives rarely have access to that server code, and that companies themselves often discard the code after the servers go down. This, of course, must be addressed and best practices must be established to preserve server code, but it also brings up an issue for the future. As companies develop and implement new forms of DRM, a national board of preservation could offer a way for developers and publishers to work with archivists to ensure that such measures do not hinder preservation efforts, while helping to assuage any fears on the part of the companies that the people attempting to circumvent the DRM are merely attempting to violate the companies’ copyrights.

3. Establishing Model Agreements and Best Practices

Beyond merely resolving contentious legal issues standing between archivists and the industry, the establishment of a national board allows its members to get a lay of the land regarding the current state of game preservation, organize priorities as needed, and establish best practices. It also provides a platform to publicly promote the importance of game preservation and deepen the conversations around digital games in general. Through collaboration with industry leaders, a national preservation board should undertake a thorough survey akin to the one undertaken by the NFPB in 1993, allowing everyone involved to get a clearer picture of the broad state of game preservation, patterns of industry practices, and the areas of game preservation that need the most immediate attention. Once that information is in hand, the board can develop established best practices for both rightsholders and archives to ensure the long-term preservation of games, determine the ultimate scale of preservation that is possible given their resources, and facilitate cooperation between archives and rightsholders.

As part of these best practices, the board should draft model agreements that facilitate collaboration between archives and rightsholders, establish a set of expectations as to what various transactions should look like, and lower the legal costs of both sides. These model agreements should cover transactions such as the secure transfer of source code and possible update schedules for that source code as the game is changed and added to. The model agreements should also help facilitate other related transactions, such as if an archive wants to make a server available for a certain game or wants to make a game publicly available

255 See supra Parts ILD, III.B.3.
256 Birnbaum & Gault, supra note 146.
257 See generally MELVIN & SIMMON, supra note 216 (describing the state of preservation for the film industry); NAT’L FILM PRES. BD., supra note 217 (prescribing a national plan for film preservation).
258 MELVIN & SIMMON, supra note 216.
for play. Combined industry and archival knowledge also provide opportunities to co-develop special “archival emulators” that can more effectively run games whose source code has been lost.259

V. CONCLUSION

Looking at the world now, it is impossible to deny the effect that digital games have had on our culture. Games are now part of the “everyday world.” But it can sometimes be difficult to understand the value of preserving parts of that everyday world until they are lost. Unless we preserve those games that are part of our everyday, it will be impossible for future generations to understand the depth of their impact on our society, history, and culture.

The games of the past few decades chart the birth of a new and exciting medium of human creativity that combines purposeful design and interactivity, story and play. It is a form that will continue to grow and build on itself, providing entertainment to future generations and exploring new and exciting avenues of human experience. Beyond their historical value, digital games have value as individual works of art and expression that deserve to be preserved as much as any film, novel, or painting. They are an achievement of our culture and we should be proud to preserve them for the future.

A national preservation board presents industry organizations, government agencies, and archives with the opportunity to relieve many of the legal issues standing in the way of truly long-term preservation and ensures that future generations have access to this history. It also presents publishers and developers with financial incentives, allowing them to offload the costs of preserving the games for future use and rerelease, as well as providing rightsholders with a platform to promote their back catalogues. But most of all, it acknowledges the importance of digital games in our lives and culture.

Bringing a national game preservation board to life will require the support of the industry, of archivists, and of the general public. Legislators and policymakers will need to be convinced that digital games are worth preserving, and without the support of all three it will be a difficult case to make. Luckily, if the first two groups are able to work together, there is no question that they will be able to convince the third. People already love games—all archivists and industry leaders need to do is convince them that such a love is worth preserving.

---

259 See supra Part III.A.2 for a brief description of emulators and the difficulty involved in creating them.