

**NEW YORK UNIVERSITY  
ANNUAL SURVEY  
OF AMERICAN LAW**

**VOLUME 74  
ISSUE 1**

**NEW YORK UNIVERSITY SCHOOL OF LAW**  
ARTHUR T. VANDERBILT HALL  
Washington Square  
New York City

*New York University Annual Survey of American Law*  
is in its seventy-fifth year of publication.

L.C. Cat. Card No.: 46-30523  
ISSN 0066-4413  
*All Rights Reserved*

---

*New York University Annual Survey of American Law* is published biannually at 110 West 3rd Street, New York, New York 10012. Subscription price: \$30.00 per year (plus \$4.00 for foreign mailing). Single issues are available at \$16.00 per issue (plus \$1.00 for foreign mailing). For regular subscriptions or single issues, contact the *Annual Survey* editorial office. Back issues may be ordered directly from William S. Hein & Co., Inc., by mail (2350 North Forest Rd., Getzville, NY 14068), phone (800-828-7571), fax (716-883-8100), or email (order@wshein.com). Back issues are also available in PDF format through HeinOnline (<http://heinonline.org>).

All works copyright © 2018 by the author, except when otherwise expressly indicated. For permission to reprint an article or any portion thereof, please address your written request to the *New York University Annual Survey of American Law*.

Copyright: Except as otherwise provided, the author of each article in this issue has granted permission for copies of that article to be made for classroom use, provided that: (1) copies are distributed to students at or below cost; (2) the author and journal are identified on each copy; and (3) proper notice of copyright is affixed to each copy.

Manuscripts: The *Annual Survey* invites the submission of unsolicited manuscripts. Text and citations should conform to the 20th edition of *A Uniform System of Citation*. Please enclose an envelope with return postage if you would like your manuscript returned after consideration.

---

Editorial Office: 110 West 3rd Street, New York, N.Y. 10012  
(212) 998-6540  
(212) 995-4032 Fax  
<http://www.annualsurveyofamericanlaw.org>



*For what avail the plough or sail  
Or land or life, if freedom fail?*

EMERSON



# NEW YORK UNIVERSITY ANNUAL SURVEY OF AMERICAN LAW

## 2018–2019 BOARD OF EDITORS

### *Editor-in-Chief*

KATHRYN MORRIS

### *Managing Editors*

DANIEL W. BLAZE  
MICHAEL FREEDMAN  
MARISSA L. KIBLER  
DAVID MOOSMANN  
H. CHASE WEIDNER

### *Development Editors*

BRIAN R. GOTTLIEB  
SAMUEL A. SEHAM

### *Note Editors*

NATALIE GOW  
JOSEPH M. LEVY

### *Inclusivity Editor*

ALLYSON N. KALEITA

### *Executive Article Editors*

EDWIN D. ABUNDIS  
ALEX INMAN  
ALLYSON N. KALEITA  
MICHAEL D. REBUCK  
AARON C.F. SALERNO  
SUZANNE D. TRIVETTE  
KIONE WONG  
KIRSTIE YU

### *Senior Articles Editors*

KATHERINE E. CORIC  
ISA M. HERRERA MOYA  
AMY I. WANN

### *Online Editor*

JOSHUA M. PIRUTINSKY

### *Articles Editors*

CLARA ARAGONE-DIAZ  
DEVIKA M. BALARAM  
MIKHAIL BATSURA  
ELLIOTT P. BARRON  
JACOB A. BRALY  
ALLISON S. CANDAL  
ROBERT COBBS  
CAROLINE DIAZ

KIMBERLY T. DREXLER  
YIRAN DU  
RANDOLL THEODORA FISHER  
LILLIAN GRINNELL  
JOSHUA HABERL  
SAMUEL F. HIMEL  
AVI LAHAM  
CHRISTINE J. LEE  
DYLAN LONERGAN

IGNATIUS NAM  
NEENA DEB SEN  
NICOLE H. SNYDER  
JACOB L. STANLEY  
KATHERINE T. STEIN  
IAN M. SWENSON  
BRETT M. WEINSTEIN  
DARREN D. YANG

### *Staff Editors*

JOSHUA BARKOW  
YOSEF BARUH  
LAUREN RACHEL BOBERSKY  
SAMUEL BOORSTEIN  
MELANIE D. BORKER  
TIFFANI CHANROO  
EKATERINA CHERNOVA  
AUDREY CRABTREE-HANNIGAN  
ANNE S. COVENTRY  
AMANDA DEMASI  
GABRIEL EISENBERGER  
JOANNA M. FALEY  
KIMBERLY FAYETTE  
HANNAH GOLDBERG  
DINO ILIEVSKI  
JONATHAN P. VARDEH JAFFE

SASHA KAWAKAMI  
KRISTEN LIANG  
KATHLEEN M. LEWIS  
KAREN LUO  
AARON MATTIS  
ELIZABETH A. MORGAN  
SIDNEY MOSKOWITZ  
KELLY NABAGLO  
JOEL NAIMAN  
YELENA NIAZYAN  
TRENTON S. PACER  
GABRIELLE N. PACIA  
ERIC S. PADILLA  
JORDANA PALGON  
MEGHAN PATZER  
JUSTIN PEROFF  
OLIVIA POSS

CALEB ROCHE  
JORDAN ROSMAN  
JESSICA RUBIN  
SANDER SABA  
ELIZABETH SAMIOS  
JESSICA SANDOR  
DANIELA SERNA BERRIO  
JOSEPH SHUI  
ELINA SIGAL  
YIANNI TSESMELIS  
IBOH UMODU  
ALYSSA WAARAMAA  
KASI L. WAUTLET  
SAMANTHA WILHELM  
KE YANG XIA  
EVAN T. YOO

This Volume of  
New York University Annual Survey of American Law  
is respectfully dedicated to  
**JUDGE PAULINE NEWMAN**



JUDGE PAULINE NEWMAN

## **JUDGE PAULINE NEWMAN**

Pauline Newman was appointed by President Ronald Reagan in 1984. From 1982 to 1984, Judge Newman was Special Adviser to the United States Delegation to the Diplomatic Conference on the Revision of the Paris Convention for the Protection of Industrial Property. She served on the advisory committee to the Domestic Policy Review of Industrial Innovation from 1978 to 1979 and on the State Department Advisory Committee on International Intellectual Property from 1974 to 1984. From 1969 to 1984, Judge Newman served as director, Patent, Trademark and Licensing Department, FMC Corp. From 1961 to 1962 she worked for the United Nations Educational, Scientific and Cultural Organization as a science policy specialist in the Department of Natural Sciences. She served as patent attorney and house counsel of FMC Corp. from 1954 to 1969 and as research scientist, American Cyanamid Co. from 1951 to 1954. Judge Newman received a B.A. from Vassar College in 1947, an M.A. from Columbia University in 1948, a Ph.D. from Yale University in 1952 and an LL.B. from New York University School of Law in 1958.

**DEDICATEES**  
**OF**  
**NYU ANNUAL SURVEY OF AMERICAN LAW**

1942	Harry Woodburn Chase	1979	David L. Bazelon
1943	Frank H. Sommer	1980	Edward Weinfeld
1944	Manley O. Hudson	1981	William J. Brennan, Jr.
1945	Carl McFarland	1982	Shirley M. Hufstедler
1946	George B. Galloway	1983	Thurgood Marshall
	Robert M. LaFollette, Jr.	1984	Hans A. Linde
	A.S. Mike Monroney	1985	J. Skelly Wright
1947	Roscoe Pound	1986	William Wayne Justice
1948	Arthur T. Vanderbilt	1987	Frank M. Johnson, Jr.
1949	Herbert Hoover	1988	Bernard Schwartz
1950	Bernard Baruch	1989	Barbara Jordan
1951	Robert P. Patterson*	1990	Harry A. Blackmun
1952	Phanor J. Ede	1991	Martin Lipton
1953	Edward S. Corwin	1992/93	John Paul Stevens
1954	Arthur Lehman Goodhart	1994	Judith S. Kaye
1955	John Johnston Parker	1995	Hillary Rodham Clinton
1956	Henry T. Heald	1996	Sandra Day O'Connor
1957	Herbert F. Goodrich	1997	Ruth Bader Ginsburg
1958	Harold H. Burton	1998	Janet Reno
1959	Charles E. Clark	1999	Alexander L. Boraine
1960	Whitney North Seymour		Desmond M. Tutu
1961	Austin Wakeman Scott	2000	George Mitchell
1962	Fred H. Blume	2001	Norman Dorsen
1963	Laurence P. Simpson	2002	Laurence H. Tribe
1964	Edmond Cahn*	2003	John Sexton
1965	Charles S. Desmond	2004	Richard Posner
1966	Tom C. Clark	2005	Antonin Scalia
1967	Francis J. Putnam	2006	Ronald Dworkin
1968/69	Russell D. Niles	2007	Stephen G. Breyer
1969/70	Jack L. Kroner*	2008	Anthony G. Amsterdam
1970/71	Frank Rowe Kenison	2009	Patricia M. Wald
1971/72	Robert A. Leflar	2010	Arthur R. Miller
1972/73	Justine Wise Polier	2011	Cass R. Sunstein
1973/74	Walter J. Derenberg	2012	Derrick Bell*
1974/75	Robert B. McKay	2013	Guido Calabresi
1976	Herbert Peterfreund	2014	Chief Judge Diane P. Wood
1977	Charles D. Breitell	2015	Judge Jack B. Weinstein
1978	Henry J. Friendly	2016	Sonia Sotomayor

\* In memoriam

## SUMMARY OF CONTENTS

### TRIBUTES

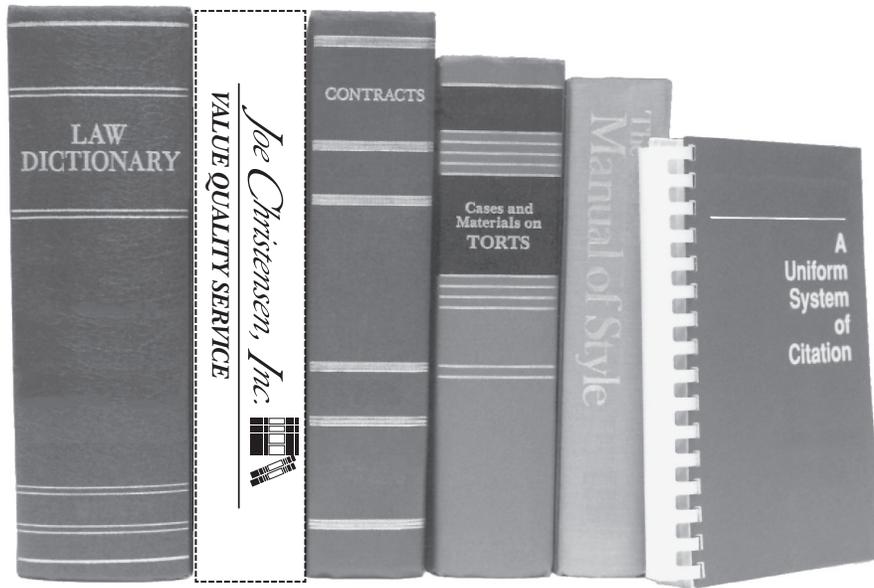
<i>Introduction by Alison Wall</i>	1
<i>Judge Raymond Chen</i>	3
<i>Professor Rochelle Dreyfuss</i>	7
<i>Rachel Elsby</i>	12
<i>Daniel Klein</i>	15
<i>James A. Worth</i>	18
<i>Acknowledgment by Judge Pauline Newman</i>	20

### ARTICLES

THE LOST STORY OF NOTICE AND PERSONAL JURISDICTION	
<i>Robin J. Effron</i>	23
AN ORAL HISTORY OF RULE 23	
<i>Interview with Arthur Miller by Samuel Issacharoff</i>	105

### NOTES

PREDICTIVE POLICING: THE ARGUMENT FOR PUBLIC TRANSPARENCY	
<i>Erik Bakke</i>	131
END-TO-END AUTHENTICATION: A FIRST AMENDMENT HOOK TO THE ENCRYPTION DEBATE	
<i>Leonid Grinberg</i>	173
CRAMDOWN INTEREST RATES IN THE COMMERCIAL CREDIT CONTEXT: ARGUMENTS FOR THE TWO-STEP APPROACH IN CHAPTER 11	
<i>Matthew Wiener</i>	223



## We Complete the Picture.

In 1932, Joe Christensen founded a company based on Value, Quality and Service. Joe Christensen, Inc. remains the most experienced Law Review printer in the country.

Our printing services bridge the gap between your editorial skills and the production of a high-quality publication. We ease the demands of your assignment by offering you the basis of our business—customer service.

*Joe Christensen, Inc.* 

1540 Adams Street  
Lincoln, Nebraska 68521-1819  
Phone: 1-800-228-5030  
FAX: 402-476-3094  
email: sales@christensen.com

**Value**

**Quality**

**Service**

Your Service Specialists

# TRIBUTE TO JUDGE PAULINE NEWMAN

ALISON WALL

Thank you, Dean Morrison. On behalf of the entire *Annual Survey of American Law*, I want to welcome you to our Dedication Ceremony in honor of Judge Pauline Newman of the United States Court of Appeals for the Federal Circuit.

Each year, the journal dedicates our upcoming volume to a person who has made a significant contribution to the law. A year ago, I sat with my peers as we brainstormed the type of person we wanted to give this honor to.

We wanted someone with a passion for the law. Judge Newman's passion for the law is indisputable. As the first judge appointed directly to the Federal Circuit in 1984, she still has yet to take senior status over 30 years later. She vehemently defends and advocates for the stability of patent laws, the protection of investments and innovation, and the patent community at large.

We wanted someone we look to as a role model for the type of lawyer, scholar, or judge we will strive to become. In Judge Newman, we found more than that: we found a trailblazer for not only women's rights but for her entire field. She has advocated for a specialized court focused in patent litigation; in fact, she served an integral role in the creation of the very court she still sits on today. She continues to serve as a constant resource to younger women judges and as a source of perspective and insight among international judicial communities.

We wanted someone accomplished and brilliant. This is evidenced by Judge Newman's degrees alone, though her accomplishments and intellect extend far beyond her diplomas. She has a Bachelor of Arts from Vassar College, a Master's from Columbia, a Ph.D. in Chemistry from Yale, and an LL.B. from NYU School of Law. Moreover, she has been a loyal alumna to us at NYU, having served as an important advisor for the Engelberg Center on Innovation Law & Policy and by continuing to return to these halls for the various awards and honors she has earned.

We wanted someone dynamic, unique, perhaps a little uncommon. In Judge Newman, we found a chemist, a scholar, an IP lawyer, a judge. She is a woman who took the road less traveled. That would hold true today and it certainly held true in the 1950's and 60's. Her fearlessness and fervent pursuit of her beliefs and passions remind us all why we came to law school and where we hope to go.

Needless to say, it is no surprise that one year ago, Judge Newman's name climbed to the top of our list, and I am incredibly excited and honored to have her here with us today. With that, I turn to Dean Morrison and our wonderful speakers to expound more personally on Judge Newman's accomplishments. Thank you.

# TRIBUTE TO JUDGE PAULINE NEWMAN

*JUDGE RAYMOND CHEN*

Thank you, Dean Morrison, and thank you all for that warm NYU welcome. I'm delighted to be here tonight for many reasons. First, as the Dean mentioned, I'm a proud alum here at the NYU law school and second, Judge Newman is the most beloved colleague on our court, and so I'm very happy to represent the Federal Circuit at tonight's *Annual Survey* dedication. Another reason I'm happy to be here is because my wife Lisa Hsiao was the editor-in-chief of *Annual Survey* in 1993, so I fondly recall watching her lead the dedication of that year's volume to Justice John Paul Stevens. And Dean Sexton was there that night as well. Lisa, of course, wishes that she could be here today, but, you know, someone had to stay at home with the kids. As you can see, tonight's event resonates for me on a lot of different levels. And to me, events like this dedication are just another example of what makes NYU a great law school and community, and so I'm thrilled to see that this *Annual Survey* tradition continues, and continues to thrive. So, to the current editors of the *Annual Survey*, let me say, you've made an excellent choice by recognizing Judge Newman and her considerable contributions to American law.

While she may not realize it, my relationship with her dates back twenty-five years, starting as a student of her work on the court, then as an advocate for the court, and now as her colleague on the bench. Starting in 1992, when I was reading Judge Newman's opinions in Professor Dreyfuss's patent law class, we read many of the court's opinions, including, for example, *Paulik v. Rizkalla*,<sup>1</sup> which is authored by Judge Newman. It's an important en banc Federal Circuit opinion on determining which of two inventors who happen to invent the same invention around the same time should be regarded as the first inventor and thereby be awarded a patent grant. Now, Judge Newman's majority opinion reached an appropriate, common-sense conclusion that a first inventor does not "abandon or suppress" that invention, as those terms are used in the Patent Act,<sup>2</sup> if he steps away from his work for

---

1. *Paulik v. Rizkalla*, 760 F.2d 1270 (Fed. Cir. 1985).

2. Patent Act, 35 U.S.C. § 102(g) (1952) (amended 2012).

a period of time but then returns to his work on the invention in time before the second inventor begins his work on the invention.<sup>3</sup>

When you read this opinion, you will see a judge's stubborn commitment in carefully, scrupulously reviewing all the relevant precedent, of which there was much, and after announcing the court's en banc holding, Judge Newman went on to make the following observations: "This result furthers the basic purpose of the patent system. The exclusive right, constitutionally derived, was for the national purpose of advancing the useful arts—the process today called technological innovation."<sup>4</sup> Judge Newman then goes on to say, "The reason for the patent system is to encourage innovation and its fruits: new jobs and new industries, new consumer goods and trade benefits. We must keep this purpose in plain view as we consider the consequence of interpretations of patent law. . ."<sup>5</sup> Judge Newman wrote the *Paulik* decision thirty-two years ago in 1985, after only one year on the court. And I think it serves as a calling card for how she has been approaching cases during her decades on the court. All along, she has consistently been able to work through the sometimes convoluted legal frameworks and technologically difficult factual records to reach wise and just decisions that always help to remind us of the core principles underlying our laws. There's a persistence, a level of scholarship, and an eloquence in her writings that are second to none. Like so many students who have studied patent law in the last thirty years, I deeply admired her work on the court while I was at NYU.

My next intersection with Judge Newman came after law school, when I joined the Patent and Trademark Office (PTO) as an associate solicitor. In that position, I had the opportunity to argue several cases at the Federal Circuit on behalf of the PTO. In these cases, the patent applicant or patent owner typically receives an adverse decision from the agency and is appealing that decision to the Federal Circuit. My office's role was to explain why the agency's decision was eminently reasonable and should be given considerable amounts of deference. Judge Newman, however, seemed to take a different view of the agency's work and she stood out as the most active and relentless questioner of PTO attorneys. In fact, after one particularly withering oral argument in front of Judge Newman in 2001 or 2002, I remember walking out of court, grumbling out loud, "My God, when is Newman going to retire?" Polly, I don't feel that way anymore and I hope we are on the court

---

3. *Paulik*, 760 F.2d at 1273.

4. *Id.* at 1276.

5. *Id.*

together for the next twenty years until I retire. What I came to appreciate eventually was that Judge Newman's persistent scrutiny was driven by a broader concern to ensure that the PTO, as a large bureaucracy, refrained from applying the patent laws in an overly rigid, mechanical way that would harm inventors and undermine the very purposes of the patent system. And Polly, I can tell you that the PTO lawyers would often go back and transmit your messages that you expressed during those oral arguments.

Judge Newman had a profound impact, not only through her many opinions, for example, *In re Sang Su Lee*,<sup>6</sup> but also through her comments during oral arguments. Finally, I'm proud to be Judge Newman's colleague and friend, and it's been quite an adjustment for me to begin thinking of someone I looked up to for as long as I did as my coworker. But I have to say that she's been exceptionally warm and gracious in welcoming me and sharing her wisdom and experiences on the court.

As the longest serving judge on our court, she is the embodiment of the Federal Circuit, and I also think of her as the conscience of the court. More than anyone else on our court, she is sensitive to ensuring that the rights of the pro se individual are upheld and protected whether that person is an inventor or a federal employee that is seeking to invoke his or her civil service protections. Of course, we often see patent cases in which hundreds of millions of dollars are on the line, but we have other cases as well, and Judge Newman is always there, commenting on the court's draft opinions where she sees an injustice that warrants deeper scrutiny.

What I've also seen up close is the courage she has in following her convictions. There is no doubt Judge Newman is an independent thinker. I don't have any statistics with me, but I would bet that she dissents more than the average judge, and I have no problem with that. I think it's a healthy thing for a court to see and consider diverging views on an issue. Knowing when to dissent, however, is something I've had to feel my way through during my first couple years. Polly and I had a case with Judge Lourie recently called *Amgen v. Sandoz*.<sup>7</sup> This is a statutory interpretation case that's now pending at the Supreme Court dealing with a very complex, new statute known as the Biologics Price Competition and Innovation Act,<sup>8</sup> and when it comes to the intersection of life sciences and

---

6. *In re Sang Su Lee*, 277 F.3d 1338 (Fed. Cir. 2002).

7. *Amgen Inc. v. Sandoz Inc.*, 794 F.3d 1347 (Fed. Cir. 2015), *cert. granted*, 137 S.Ct. 1664 (2017).

8. Biologics Price Competition and Innovation Act, 42 U.S.C § 262 (2010).

patent law, I consider Judge Newman and Judge Lourie as the two foremost experts in our country. But for some reason, I felt differently than they did on a key issue, and I had to decide whether to dissent, and I realized even though Polly didn't agree with me on the merits, she would want me to write a dissent and express another way of thinking about the statute. So I did. Since then, I sometimes say to myself, "WWPD—What would Polly do?" And then act accordingly. After speaking to several of my colleagues on the court, I can say the court is thrilled by *Annual Survey's* selection this year. She is beloved by all of us for her collegiality, her dynamic intellect, her spirit, and her deep devotion to law and justice. The court and the country are lucky to have Judge Newman. Polly, congratulations.

## TRIBUTE TO JUDGE PAULINE NEWMAN

*PROFESSOR ROCHELLE DREYFUSS*

It's fabulous that the *Annual Survey* has dedicated a volume to Judge Newman, a very loyal alum, a great friend of the school, and a terrific and terrifically interesting judge. As Trevor said, I'm the Pauline Newman Professor of Law, so I especially wanted to say a few words about why I'm so happy to hold this chair. The first reason is this: I was appointed to the chair not all that long after I entered the academy, in a day when only very senior faculty were chaired. People were extremely impressed that I was honored so early in my career. Little did they know that Judge Newman had styled the chair in such a way that at the time, I was the only one at the entire university who met the criteria. So they had no choice, for which I thank her. And I've managed to fool a lot of people, including sometimes myself.

The second reason it's so fun to have this chair is because the Judge and I are interested in many of the same things. We're often invited to the same conferences and what's more, find ourselves on the same panel. First the audience hears from Pauline Newman and then my PowerPoint slides come up and they say, "Pauline Newman Professor," and since we often disagree with one another, the contents of those slides are often diametrically opposed to what the Judge just said. We wind up in a great dialogue, a dialogue that can only happen because Judge Newman is so gracious about, and as we'll see may even relish, disagreements. So I'm truly lucky to have her chair.

But the most important reason I love having this chair is that it's brought us close and given me a fantastic opportunity to learn from her. Like many academics, I tend to go spinning off into the world of doctrine and theory. The judge, in contrast, is eminently down to earth and practical. As you heard, she was a chemist before she became a lawyer, so she knows what it takes to work in the creative sector, to make progress at the lab bench, and to bring the products developed there to market. And as you heard, she never loses sight of the goals of intellectual property protection, which is to encourage and facilitate that kind of work. I know my students think I emphasize her cases because of our connection, but I teach them because they're in the casebook. And they're in the casebook because textbook authors recognize how often her approach is the right approach. Or I should say, the right approach normatively.

And the right approach eventually, but not always the law of the moment.

The first Justice Harlan is often called the great dissenter, but within the world of patents, that title unquestionably goes to Judge Newman. And I do have the statistics. In a recent empirical study, Daryl Lim reviewed her thirty years of work on the Federal Circuit.<sup>9</sup> During that time, she wrote 202 dissents, the most of any Federal Circuit judge. She has dissented on substantive questions: patentability, infringement, inequitable conduct, remedies. She has dissented on procedural issues: reviewability, joinder of parties, evidentiary questions. And she's dissented on institutional matters, such as overreaching by the Federal Circuit into the discretionary authority of the district courts and most recently on the relationship between the Federal Circuit and the Patent Trial and Appeal Board, the PTAB. These dissents are of crucial importance. The Federal Circuit is almost unique in the federal judicial system because its jurisdiction is based on subject matter, not geography, and thus it has near plenary authority over patent law. That means there's no forum shopping at the appellate level on patent cases, which is a good thing, but it also causes a significant problem. There's no dialogue, no percolation among appellate courts, no opportunity for courts to debate open questions, distill the issues, experiment with different approaches. And there's no way to serve up a rich, detailed understanding of these issues to the Supreme Court for ultimate resolution. Instead of conflicts among the circuits, patent law relies on conflicts within the Federal Circuit. So for those of you sitting here within CA2, you should think of her as embodying the entirety of the Ninth Circuit. And she does that job brilliantly. As I said, her views tend to be the ones that ultimately prevail. Sometimes the Federal Circuit simply ignores the majority opinion and goes with her dissent in later cases. Sometimes her dissents signal the Supreme Court to intervene. That's happened nine times and in eight of those cases, the Supreme Court reversed or vacated the majority's opinion, and sided with Judge Newman.<sup>10</sup> The only one who comes even close to that is Judge Mayer, and his dissents were sustained on certiorari only three times.<sup>11</sup>

But in the end, sometimes it takes an act of Congress to get where she wants to go, but she does tend to get there. So the best current example here lies in the recent set of Supreme Court cases

---

9. Daryl Lim, *I Dissent: The Federal Circuit's 'Great Dissenter,' Her Influence on the Patent Dialogue, and Why It Matters*, 19 VAND. J. ENT. & TECH. L. 1 (2016).

10. *Id.* at 51.

11. *Id.*

stripping patent eligibility from a whole slew of technical advances.<sup>12</sup> At the Supreme Court, the effort began with a cry of anguish from Justice Breyer that sometimes patents can impede rather than promote the progress of science. Now he thought that for a specific reason. He thought that because the Federal Circuit had significantly compromised the ability of scientists to build on one another's work.<sup>13</sup> In a series of cases, it narrowed nearly to extinction the experimental use defenses in both common and statutory law. Judge Newman dissented every time she had the chance. She knew as a former scientist just how important experimentation is to pushing forward the frontiers of knowledge. Take, for example, *Integra v. Merck*, where the Federal Circuit majority narrowed the scope of the statutory experimental use defense.<sup>14</sup> Judge Newman dissented. "The purpose of the patent system," she said, "is not only to provide a financial incentive to create new knowledge and bring it to public benefit through new products; it also serves to add to the body of published scientific/technical knowledge. The requirement of disclosure of the details of patented inventions facilitates further knowledge and understanding of what was done by the patentee, and may lead to further technologic advance. The right to conduct research to achieve such knowledge, need not, and should not, await expiration of the patent."<sup>15</sup> Now she was right. The Federal Circuit's *Integra* decision was vacated and remanded nine-zip by the Supreme Court.<sup>16</sup> Unfortunately, that reversal failed to solve the larger problem of experimentation. The statutory exemption sustained by the Supreme Court in *Integra* applies only to research in the life sciences, and then only for work required to secure FDA approval of new drugs and medical devices.<sup>17</sup> To prevent patents from impeding, rather than promoting the progress of science elsewhere, what the Supreme Court has done is to categorize many important advances in genetics, biotechnology, and computer sciences as simply non-patentable.<sup>18</sup> Great, perhaps, for

---

12. See, e.g., *Mayo Collaborative Servs. v. Prometheus Labs., Inc.*, 566 U.S. 66 (2012); *Ass'n for Molecular Pathology v. Myriad Genetics*, 569 U.S. 576 (2013); *Alice Corp. v. CLS Bank Int'l*, 134 S. Ct. 2347 (2014).

13. *Lab. Corp. of Am. Holdings v. Metabolite Labs., Inc.*, 548 U.S. 124, 132-33 (2006) (Breyer, J., dissenting).

14. *Integra Lifesciences I, Ltd. v. Merck KGaA*, 331 F.3d 860, 867 (Fed. Cir. 2003), *vacated*, 545 U.S. 193 (2005).

15. *Id.* at 873 (Newman, J., dissenting).

16. *Merck CGaA v. Integra Lifesciences I, Ltd.*, 545 U.S. 193 (2005).

17. *Id.* at 208.

18. See, e.g., *Mayo Collaborative Servs. v. Prometheus Labs., Inc.*, 595 U.S. 66 (2012); *Alice Corp. v. CLS Bank Int'l*, 134 S. Ct. 2357 (2014).

fundamental scientific advances, but not so good for firms that want to invest in bringing things like new diagnostics and computer technology to market. So Judge Newman has continued to dissent. In *CLS Bank v. Alice* on the patentability of computerized business methods, she opined as follows: “This [issue on patentable subject matter] appears to have its foundation in a misunderstanding of patent policy. For the debate about patent eligibility . . . swirls around concerns for the public’s right to study the scientific and technological knowledge contained in patents. The premise of the debate is incorrect, for patented information is not barred from further study and experimentation in order to understand and build upon the knowledge disclosed in the patent.”<sup>19</sup> She went on to suggest that the problem of destroying patentability along with incentives to innovate would fade away with clarification on experimental use. And you know what’s happening? The patent industries can’t live with the current state of affairs so they’re lobbying Congress and their proposal is to couple reform of patentable subject matter to restoration of the experimental use defense.<sup>20</sup> Exactly Judge Newman’s point.

We’ve had a very patent-heavy discussion so far tonight, but the Federal Circuit also has jurisdiction over appeals from the Trademark Office, so I’ll give as one last example a trademark case: *In re Bayer*, a case that nicely demonstrates the difference between a theoretical or purely doctrinal approach and the hardcore, practical common sense views of Judge Newman.<sup>21</sup> The issue there was whether Aspirina, a trademark for Bayer Aspirin in almost the entire world, is generic and therefore un-protectable here in the United States. The majority opinion by Judge Moore is a ton of fun. To see what the word means in the U.S., she Googled it. She then wrote an extremely learned opinion on what one can deduce from Google. Judge Moore was a professor before she went on the bench. She’s an academic at heart and wrote the kind of opinion that, I have to confess, I aspire to write. In the end, she decided Aspirina was generic and could not be protected as a trademark.<sup>22</sup> Judge Newman’s dissent, I’m paraphrasing, is basically this: everyone who has travelled anywhere in Europe or South America knows

---

19. *CLS Bank v. Alice Corp.* 717 F.3d 1269, 1321 (Fed. Cir. 2013) (*en banc*) (Newman, J., dissenting).

20. *See*, IPO Daily News, INTELLECTUAL PROPERTY OWNERS ASSOCIATION (Jan. 31, 2017, 1:40 PM), [https://www.ipo.org/index.php/daily\\_news/31-january-2017/](https://www.ipo.org/index.php/daily_news/31-january-2017/).

21. *In re Bayer Aktiengesellschaft*, 488 F.3d 960 (Fed. Cir. 2007).

22. *Id.* at 970.

Aspirina is Bayer's trademark. The use by anybody but Bayer will be confusing. The goal of trademark law is to protect against consumer confusion. So why in the world would we want to let anyone but Bayer use the mark?<sup>23</sup> Now that was a dissent. Aspirina didn't get registered per Judge Moore's opinion on genericity. But, you know what? The notion that we should pay more attention to what words mean abroad is becoming the law. Bayer won its next case about the trademark for an analgesic based on knowledge of the mark in Mexico.<sup>24</sup> That case, decided by the Fourth Circuit, is now on appeal to the Supreme Court.<sup>25</sup> We'll see who it is the Supreme Court agrees with this time.

I should add that Judge Newman is an independent thinker in real life as well as in her opinion writing. My husband and I once took her from NYU, from an NYU event, to her hotel, and on the way we passed Hogs and Heifers, a sleazy bar in what was, at that time, a sleazy meatpacking district. As we passed, she rolled down the window and called out, "Hogs! Heifers! Hogs! Heifers!" I said, "Judge Newman, do you know that's a biker bar?" She said, "Of course I know it's a biker bar. I was a biker."

NYU has been amazingly lucky having Judge Newman as our grad. She has taken time to be in residence and taught classes. She has done fireside chats, often with Judge Dyk, whose opinions often provoke her dissents. She has given many of our students a leg up in the profession by hiring them as clerks. We'll have one of them speaking soon. And she's been involved in many alumni programs. It's hard to think of anyone more deserving of this award, so, congratulations.

---

23. *Id.* at 972.

24. *Belmora LLC v. Bayer Consumer Care AG*, 819 F.3d 697 (4th Cir. 2016).

25. On February 27, 2017, the Supreme Court denied the petition for *writ of certiorari*. *Belmora LLC v. Bayer Consumer Care AG*, 819 F.3d 697 (4th Cir. 2016), *cert. denied*, 37 S.Ct. 1202 (2017).

## TRIBUTE TO JUDGE PAULINE NEWMAN

*RACHEL ELSBY*

Thank you for the introduction and for having me here today. If you Google “Pauline Newman”, you’ll find the judge shares her name with a famous labor activist. And if you read some of the articles, you’ll find that many of the attributes used to describe Pauline Newman the labor activist are also synonymous with Pauline Newman the judge. An educator, an organizer, an expert, a writer, a mentor: all words that aptly describe Judge Newman. Both women also achieved remarkable firsts as females, but also more broadly as leaders in their respective fields. And both women could undoubtedly be described, as Judge Newman’s namesake once was, as capable of smoking a cigar with the best of them.

Throughout her career, Judge Newman has found success both in science and law. But the thing that impresses me most, the thing that I’m going to attempt to talk to you about today, is who Judge Newman is as a person. Those who know Judge Newman well know her to be incredibly humble and modest. She’s very quick to recognize the achievements of others, but much more reserved when it comes to her own accomplishments. Once a year, she gathers her clerks together for a reunion, and as part of the gathering we stand in a circle and talk about what’s going on in our lives and our careers. But the judge has never touted a single one of her accomplishments at these events, including the one that was specifically intended to celebrate her thirty years on the bench. Instead, she typically ends the conversation by encouraging her clerks to be leaders and by reminding us, to borrow her words, of the enormous amount of work that remains to be done.

Another thing you learn quickly as a clerk for Judge Newman is that she has a very good sense of humor and a quick wit. She also enjoys pranks. For most of my clerkship, we had two girls and two guys in chambers. And the two girls were “disorganized”, not totally unlike Judge Newman. And the two guys were “particular” about their space. And so every once in a while, Christine, my co-clerk, and I would go over to the guys’ desks when they were out of chambers and reorganize them for them just to play a prank. Judge Newman always got a great laugh out of this. And, like any good boss does, she would offer suggestions about which articles might look best upside down.

Judge Newman also possesses a tremendous sense of adventure, and I'm not just saying that because she regularly orders the cheesesteak sandwich from the deli next to the court. I don't think it's an understatement to say she's lived five lives. She's built and driven fast cars, as well as the occasional tank. She's flown airplanes and jumped out of them. She's traveled extensively and met many world leaders, including the Queen of England, who once complimented her curtsy. She's even judged a science competition with Steve Jobs. And as a clerk, this leads to the occasional unexpected interaction. I remember driving with the judge to court one day and asking her what she'd done the night before. She responded that she was reading *Wired* magazine. She said, "Rachel, did you know all those new millionaires in Silicon Valley are buying Ferraris?" I said, "Yes, Judge, I think I heard that." To which she responded, "In my opinion, the Lamborghini and Maserati are far superior cars." I think she still laments not having a Maserati.

This, of course, was not her first dissent, as you have heard tonight. The judge has, throughout her career, not shied from expressing her opinion, either in majority or dissent, and has always done so with the singular objective of promoting a strong patent system. But despite being steadfast in her convictions, and from time to time at odds with her colleagues, Judge Newman harbors no animosity and bears no grudges. Instead her opinions, like herself, are forward-looking, sincere, and always cognizant of the need to strive for excellence. The words "we must do better" are not uncommon in a Judge Newman opinion, but when used they are couched in a respect for both the judges with whom she works and the institution she represents. But the thing that might be most remarkable about Judge Newman is that, despite her many accomplishments, she's approachable and able to interact with all sorts of different types of people. To get a sense of this, you can look at her clerk family. No two are alike. Judge Newman's clerks come from every background, with varied experiences in law and life, and yet she runs a chambers in which everyone is equal. There is no favorite clerk or special treatment. Her only expectations of her clerks are that they work hard and think critically. And, although she would never say it, I suspect this is because the Judge can do most of the work herself. But the Judge has no ego. In fact, it's not uncommon for a clerk and the Judge to exchange up to twenty drafts of an opinion before it leaves chambers. The Judge editing the clerk, and the clerk editing the Judge, exchanging ideas openly and without criticism. Judge Newman has a patience and a respect for others that make her truly unique. In conclusion, I just want to say: con-

gratulations, Judge Newman. I'm grateful for the opportunity you gave me as a clerk, and I thank you for being a role model, a mentor, and a friend.

## TRIBUTE TO JUDGE PAULINE NEWMAN

*DANIEL KLEIN*

Thank you for the introduction. As the fourth speaker up here tonight, you've heard all the high notes about the remarkable person that is Judge Newman. I should say that when I initially was invited to speak, it was somewhat lost on me the magnanimity of this event until I saw who the previous years' dedicatees were and that I would be standing up here with Judge Chen, Professor Dreyfuss, Judge Worth, and Rachel. It is humbling to be up here alongside them, but an even greater honor is standing up here to pay tribute to Judge Newman, a person who in my opinion has made a bigger impact on our country than anyone I have personally ever met, and she could not be more deserving of this honor.

The theme of my comments are Judge Newman the patriot, and I will get into how I came to that in a minute, but before I do, I would like to spend just a moment on the human being, the Judge. When I started clerking for Judge Newman, I had only a cursory understanding of the internal machinations of the judiciary, and I had virtually no opinions on how I thought judges should judge or courts should rule. I could, if asked, tell you what I thought should be the qualities of a good judge: integrity, fidelity, honesty, intellect. But these were just abstract ideas, disembodied from concepts and attributes that we hope apply to all our sacred institutions, not just our judges and courts and civil servants. Once I started clerking for Judge Newman, the view I quickly developed was that she defined my notion of the ideal judge. I'm quick to note, however, that with Judge Chen being in attendance, this is not intended as a comparison to him or any other judge at the Federal Circuit. Judge Chen truly is a remarkable jurist and an undeniable asset to the court. Judge Newman is unwavering in her sense and pursuit of justice through her judicial opinions. She's impervious to grandstanding and rhetoric, and her integrity is beyond question. Her commitment to the judicial process is evident in all of her opinions and particularly so in her dissents, which give the Bar and the Supreme Court a window into the different interpretations on a particular issue. But quote Judge Newman's dissents to her at oral argument, and she will quickly tell you that her dissents are not the law.

Now, Judge Newman the patriot. I thought this theme might capture just some of the essence of Judge Newman, a lens through

which I think we can view a lot of aspects of her life. For it is a deep-rooted sense of patriotism and civic commitment that I think informs Judge Newman's journey to the bench, her judgeship, and her jurisprudence. To hear her speak or read her opinions, one is likely to see Judge Newman refer to the nation, the important role innovators and patents have in it, and the impact courts have on civil society when they interpret the law. It was not out of a sense of aggrandizement, self-aggrandizement, or simple corporate protectionism that Judge Newman in the early '80's testified before Congress in support of a bill that would create the Federal Circuit, to which she would be the court's first direct appointee just a few years later.<sup>26</sup> Although if that had been the case, a better example of manifest destiny I've never seen. Asking Congress to create a new court for you and then appoint you to it is not a realistic life goal for anyone. Rather, her testimony to Congress was on the heels of her work on the advisory committee of President Carter's domestic policy review, which was tasked with finding ways to pull the country out of its deep economic recession.<sup>27</sup> Unpredictability in the application of patent law was seen as frustrating the recovery by encumbering innovation and investment, and it was the committee's proposal to create a uniform appellate court to administer these laws.<sup>28</sup> So it was with her understanding of how patents can help a nation thrive that Judge Newman joined the bench in 1984. Judge Newman believes strongly in the founding principle that the framers memorialized in our Constitution that an industrial nation depends on the system of protecting its dreamers' intellectual property.<sup>29</sup> But it is not, I think, simply because the framers provided for a patent system in the Constitution that Judge Newman holds this view. Rather, it is because of the essential role innovation plays in our society. When we interpret the law, when we apply the law to a new set of facts, we should be mindful of where these laws stand in relation to the nation that they serve. This, I believe, is Judge Newman's judicial philosophy. Judge Newman's civic commitment extends far beyond her support for a robust patent system. Her thoughtful judicial writings in cases involving veterans' benefits and government employees are a hallmark of her judgeship. Re-

---

26. Hearings Before the Subcomm. on Courts, Civil Liberties, and the Admin. Of Justice of the Comm. on the Judiciary H.R., 97th Cong. 1 (1981) (statement of Pauline Newman, Witness, H. Comm. on the Judiciary).

27. Pauline Newman, *The Sixth Abraham L. Pomerantz Lecture: Commentary on the Paper by Professor Dreyfuss*, 61 Brook. L. Rev. 53, 55-56 (1995).

28. S. REP. NO. 97-275 at 2 (1981).

29. U.S. CONST. art. I, § 8, cl. 8.

ardless of the outcome on the merits, Judge Newman's writing is mindful and sensitive to the sacrifices veterans and government workers make for their country. I think it is with this devotion to the country that Judge Newman has served tirelessly on the bench for over 30 years. Why, when I served as her clerk, she took my calls late on the weekend to discuss a case, and why it was not uncommon to receive emails from her at three in the morning. Why, just two years ago, Judge Newman embarked on a new endeavor to create an organization that would teach science to judges, with the view that it would help them across the entire spectrum of their dockets, from forensic issues in criminal cases to technical product warranty disputes.

In short, Judge Newman has not devoted the latter part of her life to the improvement of the law and judiciary as an ends in and of itself, although, if she had, it would be a more noble endeavor than most. Rather, Judge Newman has devoted herself to the betterment of the nation through improvements to the judiciary, our nation's capacity to lift up explorers and inventors, and our system of governing rules we call laws. She's a constant inspiration to me, and I thank her from the bottom of my heart for letting me serve as her law clerk. Congratulations.

## TRIBUTE TO JUDGE PAULINE NEWMAN

*JAMES A. WORTH*

Good evening Dean Morrison, Judge Newman, Judge Chen, Professor Dreyfuss, assembled guests. This is a really nice occasion. I'm honored to be a part of it. It's great to be back here at NYU.

NYU is really lucky to have Judge Newman and to call her one of its own. I know the Judge has made great contributions to the nation. She has also made great contributions to this institution, the NYU School of Law. NYU can certainly be very proud to have her on their rolls of this institution and to call her an alumna.

Judge Newman has had a great impact, not only on patent law, but on the wide-ranging jurisdiction of the Federal Circuit on takings law, tax law, trade law and international trade, trademark law, contract law, federal labor law, and veterans' law. Plus, Judge Newman has shown us that to be a good patent attorney you have to be conversant with a wide-ranging interconnecting range of discipline.

Judge Newman is a great writer. We've heard some of her own writing tonight. Her opinions are polished and many read like the spoken word. It's in a style all of her own. I can tell you that Judge Newman has had a tremendous impact, not only on the court system, not only on patent law and the other areas of the court's jurisdiction. I can tell you that Judge Newman has had a tremendous impact on me.

Pretty much every day I ask myself, what would Judge Newman do? Although I can tell you the answer to that question is not always obvious. This isn't about legal analysis either. One occasion had a profound impact on me, and I remember it well. There was a pro se litigant, a woman who requested oral argument but became choked up, was crying, and couldn't get a word out. Judge Newman reached out her hand from the bench with a glass of water and just asked her if she would like to have some water. Now at the court, the parties have their own water, not just the judges. But here, Judge Newman leaned over the bench and offered her some of hers. That gesture by the Judge won't be in an opinion, but, as they say, it speaks volumes.

Judge Newman also has incredibly high ethical standards. I remember one day it was raining, and I had more than one umbrella to choose from. Judge Newman suggested that I use one without a logo, just in case some day the entity on that logo might become a

party. That's just one example of how Judge Newman takes her work seriously and applies it to her own life.

I can tell you that as her former clerk, research had to be thorough. What did Jefferson say? What did Blackstone say? What did the Supreme Court say over the last 200 years? What did the English courts say at common law? My wife says that Judge Newman lit the fire of curiosity in me to get to the bottom of the issue. Judge Newman was a great boss. Sure, we worked hard. Sometimes we worked late. But Judge Newman frowned on working on weekends, if you can imagine. Except one time she gave me an assignment, one weekend in April. Homework, if you will. I was ordered to visit the cherry blossoms in the D.C. area, and you know you have to comply with the Judge's orders.

Judge Newman works incredibly hard, you may have gathered. She was a pioneer in the law, and that might have been enough for some, but she continues to work hard day in and day out. She definitely has a vision for the law and she works to carry it out.

## ACKNOWLEDGMENT

*JUDGE PAULINE NEWMAN*

Well, I think it's fair to say that I'm overwhelmed. This honor from my law school, from all of you, and your words, they warm my heart more than I can say, and that beautiful crystal gavel souvenir I shall put to very good use. I like its weight.

I should also say an appreciation to my dear colleague Judge Chen, my friend and mentor Professor Dreyfuss, my law clerks who started as law clerks and who became friends. I so much appreciate and treasure all that you've been saying this evening.

I must say I also believe every nice word that you've said. Just about only half of those who appear before me have nice things to say, I would say exactly half. The other half probably would not survive in polite society.

However, also, just in keeping with what we've been talking about, the ways of our court, I'll talk a little bit about the patent law. The ways of our court in patent law has convinced me that the judiciary perhaps is no longer the least dangerous branch. James Madison may have to reconsider. In fact, I had this inkling about the dangers of the judiciary some thirty-five years ago, when I was working in industry. When I thought that the thing to do was to remove patent cases from the judges who were deciding them. To shake up the judiciary, because there were so many dramatic new sciences, there was so much going on in the world. The nation was trying to recover from the recession of the Vietnam War. Industry viewed that uncertain law was the worst of commercial disincentives. Industry could handle competition alright, but there seemed to be no reason in the world why we at the time—I included myself in the industrial “we”—should cope with uncertainty as to what the law is, how the law is applied, particularly when the issues arise only after you've done all of your work and science, the research, the development, the marketing, the commercial aspects, only to find that someone comes along and says that you aren't entitled to whatever you based your evolution on. We in industry at that time, we've mentioned this evening some of the origins of the Federal Circuit. We in industry knew very well, we were convinced that judges just didn't understand how commerce worked. Not all judges, of course, but with forum shopping, it didn't matter if it was all, there were enough. So we in industry undertook to shake up the federal judiciary. Perhaps we didn't know better, but here we

are. In fact, that reaction of the people against the rulers would very much have pleased James Madison. I think we're very much in his tradition, as industry and universities, scholars, scientists, government agencies, and the Bar, perhaps not unanimous, but with enough weight behind them to convince the legislature that it was time to change the way cases were decided in the federal judicial structure.

Of course, the goal was to stabilize the law, to stabilize the judge-made law. We were looking at the patent law, but that certainly applies to any area that you can think of, any area of the law that produces behavior that produces a basis on which for one reason or another people rely and ought to be entitled to rely. The simple idea that we had as far as patent law was concerned was to take patent decisions away from the federal judges who were making the important law—the circuit judges of the nation and put it in the hands of a single national court that, by the very fact of this being a single national court, would provide consistency, and the product was the Federal Circuit. Actually, I think it worked. It worked pretty well. It worked pretty well for a couple of decades. There was much talk at that time in the early years of the new strength of patents. There was a burst of industrial investment, research and development. There were new products, new industries at a time when the science and technology was undergoing the kind of renaissance that you see only once every half century or so. And the court did get credit for—part of the credit, at least—for pulling the nation out of the post-Vietnam recession.<sup>30</sup> But after a couple of decades, that too wasn't working quite as well as we had hoped, and it looked as if another fresh look was appropriate. Another upheaval, at least in this same era of patent law, which was turning out to be a foundation of the industrial activity of the nation.

We had become in the nation a technology-based economy. If you look at the traditional measures of the gross domestic product, it's now measured, it's somewhere like seventy percent technology-based, whereas it was only fourteen or fifteen percent a few decades ago. So besides being heavily involved with patents, it was the way this economy was developing. Aggressive; litigious; and ingenious in the way litigation was being intertwined with commerce. Wasteful litigation, even those of you whose livelihood perhaps is intertwined with it. Unproductive; unpredictable. And so the concerned public again stepped in. This took another ten years, the past decade or so, with volumes of studies and debates and hearings. There was

---

30. *The Federal Circuit, the National Appellate Court, Celebration and Introspective Symposium*, 19 FED. CIR. B.J. 327 (2010).

always something called the America Invents Act.<sup>31</sup> It didn't change the substantive law—it changed the procedure. It took a procedural leap, moving many patent disputes, removing them from the trial, from the district courts of the nation, putting them in the hands of technical experts. And a new tribunal staffed by people knowledgeable, extremely knowledgeable in technology and the law. This is the tribunal on which Judge Worth now sits as an example of the attention that is being given to this national need. The goal was simple. Perhaps it applies to all areas of law. Consistency, predictability, harmony, confidence. Confidence in the rule of law, in a nation ruled by law where we care about the objective and subjective excellence of our situation.

I must say, I very much applaud this move. In fact, I put it in the same context, the same mindset of the modest revolution that I tried to have a hand in some decades earlier. So James, I think now that we are fixing the procedural law, I think it's probably time for us to return to the substantive law. There's a lot to be done. Our Supreme Court is much involved. It has come to be understood and appreciated that the technology, the world in which we live, requires optimum incentive; optimum human interaction; optimum betterment of the human condition as we are learning may be available, and available only through these advances of science.

I must say, these are not simple concepts. And sometimes I wonder whether we'll ever finish. But whenever I start to feel discouraged about the weight of what's ahead of us and what confronts the judiciary our court is visited by a delegation of foreign judges and lawyers. Judge Chen knows, he has entertained many such delegations. And they explain to us that the United States system, and law, and judiciary, and attitudes explain why we are indeed in the position of world leadership that we occupy. And I must say, I agree with them. I think we can and should take all of that credit. Someone has to. So why not us? And I must say now to add extra spice in the current political situation, I'm happy to see that again perhaps it is finally coming to be understood that the judiciary really is not the least dangerous branch after all.

It was an extraordinary law school that put me on this path, and I treasure the honor of this dedication. I so much appreciate all that all of you have been saying, and all of you being here. Dean Morrison, President Sexton, my old friends, new friends. Thank you.

---

31. Leahy-Smith America Invents Act, Pub. L. No. 112-29, 125 Stat. 284 (2011).

# THE LOST STORY OF NOTICE AND PERSONAL JURISDICTION

ROBIN J. EFFRON\*

*Notice and personal jurisdiction have long been closely-tied procedural law concepts because of their common origins in the mechanics of service of process and their shared due process ancestor in Pennoyer v. Neff. Notice was once a reliable feature of personal jurisdiction jurisprudence, but slowly faded from prominence in personal jurisdiction analysis after the International Shoe and Mullane decisions, and then fell away almost completely in the post-Asahi era.*

*Once the Supreme Court tied personal jurisdiction to due process, notice was critical in shaping the direction of jurisdictional doctrine. Its role extended beyond that of a mere instrument of doctrinal development. The use of notice was integral to the mode of legal reasoning that the Court employed in its personal jurisdiction journey. Notice, with its tangibility and dependence on mechanical service of process, allowed the Court to navigate the strict formalism of the pre-International Shoe era and the Court's many returns to formalism in the era of minimum contacts. Moreover, when the Court wanted to engage in a more functional mode of analysis, notice allowed the Court to continually tie personal jurisdiction to due process because of the intuitive fairness appeal of the ideas of notice and opportunity to be heard. When the Court made several efforts to limit the scope of personal jurisdiction between International Shoe and the early 1990s, the Court seized upon a different but related concept of notice—notice of jurisdiction—as a due process justification for restricting personal jurisdiction.*

*This Article advocates for a “notice-inclusive approach” to personal jurisdiction. It focuses on reestablishing comfort with the inclusion of easily-satisfied due process considerations while also stressing that constitutional notice doctrine itself might be strengthened in small but strategic ways, thus adding some additional due process protections both to notice and to personal jurisdiction.*

---

\* Professor of Law, Brooklyn Law School. Thanks to Alexandra Lahav, Jonathan Remy Nash, Zach Clopton, Pam Bookman, Brooke Coleman, Linda Silberman, the Hon. Andrew Effron, David Noll, Allan Stein, Rick Swedloff, John Leubsdorf and participants at the Third Annual Constitutional Law Scholars' Conference, the Fourth Annual Civil Procedure Workshop, and the Rutgers-Newark Faculty Workshop. Nathalie Gorman, David Moosmann, and Sander Saba provided exceptional editorial and research support throughout this process. Thanks also to Dean Maryellen Fullerton for support from the Dean's Summer Research Fund.

Introduction .....	25
I. Notice and Personal Jurisdiction from <i>Pennoyer</i> Through <i>International Shoe</i> and <i>Mullane</i> .....	30
A. The Relationship of Personal Jurisdiction and Notice Prior to <i>Pennoyer</i> .....	30
B. Personal Jurisdiction and Notice in <i>Pennoyer</i> .....	33
C. Personal Jurisdiction and Notice from <i>Pennoyer</i> Through <i>International Shoe</i> and <i>Mullane</i> .....	35
1. The <i>In Rem</i> Cases .....	36
2. The Marriage Exception Cases .....	39
3. Corporations and Consent Cases .....	41
4. The End of the Journey to <i>International Shoe</i> .	43
II. Personal Jurisdiction in <i>International Shoe</i> and <i>Mullane</i> .....	48
A. <i>International Shoe</i> and the Alternative History of a Minimum Contacts Test for Notice .....	48
B. <i>Mullane</i> and the New Trajectory of a Distinct Standard for Due Process in Notice .....	51
III. The Shifting and Fading Role of Notice in Personal Jurisdiction from <i>International Shoe</i> and <i>Mullane</i> through <i>Asahi</i> .....	55
A. Early Jurisdictional Expansion in <i>Perkins</i> and <i>McGee</i> : Continued Use of Notice as a Fairness and Due Process Crutch .....	57
B. From Notice of Suit to a Tool of Due Process Expansion .....	61
C. A New Use for Notice in a New World of <i>In Rem</i> Jurisdictional Problems .....	71
D. Jurisdiction Over Plaintiffs and the Return to the Touchstone of Notice of Suit .....	74
IV. Whither Notice .....	80
A. Notice Vanishes from General Jurisdiction .....	80
B. And Then Notice Disappears from Specific Jurisdiction .....	83
V. Notice Resurrected .....	89
A. Establishing Comfort with Easily Satisfied Due Process Criteria .....	91
B. Pressing for Deeper Constitutional Scrutiny of Notice and Service of Process Practices .....	92
C. Sharpening Specific Jurisdiction with an “Additional Procedural Protections” Approach ...	95
1. Specific Jurisdiction in Mass Actions and Class Actions .....	96

2. Using Registration Statutes to Broaden the Availability of Specific Doctrine ..... 98

D. Restoring Notice to a Broader General Jurisdiction Doctrine ..... 100

Conclusion ..... 103

INTRODUCTION

This Article examines the curious doctrinal journey of two procedural lynchpins of due process: personal jurisdiction and notice. For the past several decades, they have been treated as more or less separate doctrines. But that was not always the case, and there are still shadowy remnants of each doctrine that remain in the case law and analysis of the other. This Article attempts to answer a few important but surprisingly overlooked questions, namely: when and why did notice break off from personal jurisdiction? And why, despite decades of nearly constant hand-wringing about confusion and chaos in personal jurisdiction, did so few commentators or jurists seem to take note of this development?

The link between personal jurisdiction and notice is mechanical and conceptual. Both doctrines have common roots in the mechanics of service of process. The reason for this is fairly evident in the case of notice; service of process is the means by which a party is apprised of the pendency of an action. Constitutionally sufficient notice depends on the proper execution of service of process that is “reasonably calculated under the circumstances”<sup>1</sup> to apprise a party of an action. The connection between the mechanics of service of process and personal jurisdiction is less obvious. In the American system, service of process is the means by which personal jurisdiction is acquired or “perfected.”<sup>2</sup> In other legal regimes, personal jurisdiction is not directly connected to the question of how—and whether and when—a party should be served with process<sup>3</sup> and thus notified of the pendency of an action. But

---

1. *Mullane v. Cent. Hanover Bank & Tr. Co.*, 339 U.S. 306, 314 (1950).

2. *See Omni Capital Int’l, Ltd. v. Rudolf Wolff & Co.*, 484 U.S. 97, 104 (1987) (quoting *Miss. Publ’g Corp. v. Murphree*, 326 U.S. 438, 444-45 (1946)) (“[S]ervice of summons is the procedure by which a court having venue and jurisdiction of the subject matter of the suit asserts jurisdiction over the party being served.”).

3. *See, e.g.,* Patrick J. Borchers, *Comparing Personal Jurisdiction in the United States and the European Community: Lessons for American Reform*, 40 AM. J. COMP. L. 121, 135 (1992) (criticizing the United States’ use of service of process as a basis for personal jurisdiction and noting that the Brussels Convention outlaws “such exorbitant devices”).

in the United States, the procedures for service of process undergird both notice and personal jurisdiction.

A “problem” with service of process could actually be any one of three distinct problems: a problem with the mechanics of service, a problem with notice, or a problem with personal jurisdiction. Both federal and state courts have distinct motions and procedures for redressing problems with each.<sup>4</sup> In practice, courts do not usually dwell on missteps in the formal delineation of a motion regarding a problem related to service of process.<sup>5</sup> This means that arguments and concepts about all service of process problems—mechanics, notice, and personal jurisdiction—bleed into each other. Looking at the trajectory of personal jurisdiction doctrine and notice doctrine over the past 150 years, the concepts and arguments associated with each can get tangled together. Sometimes the doctrines seem to merge or look as if one will subsume the other. At other times, the doctrines and arguments drift apart.

Requirements for proper notice and lawful personal jurisdiction predate the passage of the Fourteenth Amendment. In 1877, the Supreme Court used the due process clause to elevate both doctrines to constitutional status.<sup>6</sup> Since then, the Court has struggled to make sense of notice and personal jurisdiction, both in providing an internally coherent account of each doctrine, and also in explaining the due process basis for each as a constitutional right.<sup>7</sup> Personal jurisdiction has unquestionably been the more difficult and problematic due process doctrine, and as such, it is around personal jurisdiction that this Article is framed.

Personal jurisdiction encompasses doctrines and concepts that are not natural or obvious fits with due process.<sup>8</sup> To the extent that

---

4. See Fed. R. Civ. P. 12; N.Y. C.P.L.R. § 3211 (Consol. 2012); CAL. CIV. PROC. CODE § 418.10 (Deering 2002); Del. Super. Ct. Civ. R. 4(j).

5. See 62B AM. JUR. 2D *Process* § 99 (2018). However, unless the effect of a motion practice mistake results in the waiver of the ability to raise a defense like personal jurisdiction, the choice of device rarely has much practical significance.

6. See *Pennoyer v. Neff*, 95 U.S. 714, 733 (1877).

7. See RHONDA WASSERMAN, *PROCEDURAL DUE PROCESS: A REFERENCE GUIDE TO THE UNITED STATES CONSTITUTION* 63–161 (2004) (history and status of due process doctrine for notice and opportunity to be heard); *id.* at 207–61 (history and status of due process doctrine for personal jurisdiction).

8. See generally Jay Conison, *What Does Due Process Have to Do with Jurisdiction?*, 46 RUTGERS L. REV. 1071 (1994). See also Patrick J. Borchers, *The Death of the Constitutional Law of Personal Jurisdiction: From Pennoyer to Burnham and Back Again*, 24 U.C. HASTINGS L. REV. 19, 20 (1990) (“[T]he Court should . . . abandon the notion that state court personal jurisdiction is a matter of constitutional law.”); Stephen E. Sachs, *Pennoyer Was Right*, 95 TEX. L. REV. 1249, 1252 (2017) (arguing against “the main holding of *Pennoyer*: that the Fourteenth Amendment’s Due Process

personal jurisdiction has at least some of its pre-*Pennoyer* origins in the international and general law doctrines of territoriality, sovereignty, comity, and federalism, it has been difficult to square with an individual liberty-based understanding of due process—even accounting for the fairness rationales that emerged from the minimum contacts approach that the Court eventually established in *International Shoe*.<sup>9</sup> Notice doctrine, on the other hand, has always fit more comfortably with individual liberty and due process because of the emphasis on ensuring that a party is aware of a pending action so that she may participate and defend or vindicate her rights before a court issues a binding judgment.<sup>10</sup>

Notice and personal jurisdiction share common origins in the mechanics of service of process and the due process ancestor in *Pennoyer*. Both are due process rights that litigants can waive.<sup>11</sup> But it is not enough to casually observe the parallel development of the doctrines: notice and personal jurisdiction have a tangled history that is more than just a historical quirk. Notice was once a reliable

---

Clause . . . imposes rules for personal jurisdiction.”); Allan R. Stein, *Styles of Argument and Interstate Federalism in the Law of Personal Jurisdiction*, 65 TEX. L. REV. 689, 694 (1987) (“*Pennoyer*’s linkage of due process and jurisdictional theories outside the due process clause provoked reams of scholarly criticism, focusing primarily on the absence of any federalism component of the fourteenth amendment.”). *But see* Kenneth J. Vandeveld, *Ideology, Due Process and Civil Procedure*, 67 ST. JOHN’S L. REV. 265, 274–77 (1993) (due process formulation of personal jurisdiction in *Pennoyer* was consistent with the Supreme Court’s broader conservative due process ideology of the era).

9. *Int’l Shoe Co. v. Washington*, 326 U.S. 310 (1945). *But see* Charles W. “Rocky” Rhodes, *Liberty, Substantive Due Process, and Personal Jurisdiction*, 82 TUL. L. REV. 567, 567 (2007) (“This Article defends—against numerous critics—the view that constitutional limits on personal jurisdiction arise from basic substantive due process principles.”).

10. *See* WASSERMAN, *supra* note 7, at 207 (describing notice as one of the “principal procedural protections afforded by due process” and personal jurisdiction as an “important corollary.”). *See also* 4A CHARLES ALAN WRIGHT ET. AL., FEDERAL PRACTICE AND PROCEDURE, § 1074 (4th ed. 2018).

11. Personal jurisdiction can be waived by consenting to the forum, either through a forum selection clause or by failing to raise a timely jurisdictional objection. *See* 4A CHARLES ALAN WRIGHT ET. AL., FEDERAL PRACTICE AND PROCEDURE § 1067.3 (4th ed. 2018) (“[P]ersonal jurisdiction can be based on the defendant’s consent to have the case adjudicated in the forum, or the defendant’s waiver of the personal jurisdiction defense.”). Notice, in many circumstances, can also be waived, such as through a “cognovit” note in which a party agrees in advance to forego ordinary notice and service of process in a debt action. *See* D.H. Overmyer Co. v. Frick Co., 405 U.S. 174, 176 (1972).

feature of personal jurisdiction jurisprudence,<sup>12</sup> but slowly faded from prominence after the *International Shoe* and *Mullane* decisions, and then fell away almost completely in the post-*Asahi* era.

At a superficial level, the due process story of personal jurisdiction and notice is a tale of historical accident and doctrinal path dependence. Perhaps the decision to drop notice from personal jurisdiction was not a conscious choice, but rather, the side effect of an emphasis on other jurisdictional values and the result of benign neglect. A closer look, however, reveals a more nuanced story. Once the Supreme Court tied personal jurisdiction to due process, notice was critical in shaping the direction of jurisdictional doctrine. Its role extended beyond that of a mere instrument of doctrinal development. The use of notice was integral to the mode of legal reasoning that the Court employed in its personal jurisdiction journey. This was on account of two key attributes of notice. First, notice, with its tangibility and dependence on mechanical service of process, allowed the Court to navigate the strict formalism of the pre-*International Shoe* era and the Court's many returns to formalism in the era of minimum contacts. Second, when the Court wanted to engage in a more functional mode of analysis, notice allowed the Court to continually tie personal jurisdiction to due process because of the intuitive fairness appeal of the ideas of notice and opportunity to be heard. Thus, the Court could lean on notice to provide a veneer of fairness and process, even while supposedly privileging arguments about sovereignty and territoriality. Finally, when the Court made several efforts to limit the scope of personal jurisdiction between *International Shoe* and the cases of the early 1990s, the Court seized upon a different but related concept of notice, notice of jurisdiction, as a due process justification for restricting personal jurisdiction.

This Article proceeds in five parts. Part I recounts the relationship of personal jurisdiction and notice from its roots in the pre-*Pennoyer* and due process era through the Court's slow evolution of personal jurisdiction doctrine that laid the groundwork for the modern minimum contacts test. Part II reconsiders the conventional wisdom of *International Shoe* and *Mullane*, arguing that these cases each analyze personal jurisdiction and notice in a way that had lasting consequences for personal jurisdiction doctrine and analysis. Part III traces the continued use of notice in personal jurisdiction analysis in the first five decades after *International Shoe*, dem-

---

12. See WRIGHT ET. AL., *supra* note 10, at n.2 ("As the discussion of *Pennoyer v. Neff* . . . reveals, the Supreme Court has long regarded 'notice' and 'power' as inseparable aspects of the due process restrictions on state court jurisdiction.").

onstrating that notice was used as a tool to expand personal jurisdiction by providing reassurances of fairness, and as a tool to restrict jurisdiction by appealing to evolving notions of due process and the role of notice therein. This was an era in which reliance on notice to quietly bolster doctrinal changes allowed the Court to plaster over an increasing incoherence in personal jurisdiction doctrine and reasoning.

Part IV examines the Court's latest round of personal jurisdiction cases in which notice has all but disappeared from the Court's menu of doctrines and values that support jurisdictional decisions. This absence lays bare the consequences of the evolution in the decades-long relationship between personal jurisdiction and notice. Notice had long been a fundamental yet little recognized partner in constitutional personal jurisdiction analysis. It helped paper over some of the difficult doctrinal inconsistencies in personal jurisdiction analysis, particularly concerning the nature of personal jurisdiction as a due process right. It propped up doctrinal innovation, sometimes to expand jurisdiction and sometimes to restrict it. When notice disappeared, the already apparent incoherence and inconsistencies in personal jurisdiction doctrine only became more obvious.

Finally, in Part V, I argue that restoring notice to personal jurisdiction might be a small yet helpful part of a strategy to impose normative and doctrinal order on personal jurisdiction chaos. This "notice-inclusive approach" has four distinct components. The first component focuses on reestablishing comfort with the inclusion of easily-satisfied due process considerations in personal jurisdiction analysis and treating these considerations as meaningful or even dispositive under appropriate circumstances. The second component suggests, in turn, that constitutional notice doctrine itself might be strengthened in small but strategic ways, thus adding some additional due process protections both to notice and to personal jurisdiction. The third component is to reincorporate notice as a factor in specific jurisdiction analysis, thus broadening the doctrine and sharpening its boundaries by refocusing analysis on the relationship between personal jurisdiction and other procedural protections with a due process component. The fourth component is to return notice to personal jurisdiction which might pave the way for a less restrictive, yet still appropriately constrained, approach to general jurisdiction.

I.  
NOTICE AND PERSONAL JURISDICTION FROM  
*PENNOYER* THROUGH *INTERNATIONAL*  
*SHOE AND MULLANE*

A. *The Relationship of Personal Jurisdiction and Notice  
Prior to Pennoyer*

The story of personal jurisdiction and notice begins long before *Pennoyer v. Neff*<sup>13</sup> constitutionalized both doctrines. In the pre-*Pennoyer* legal landscape, courts viewed personal jurisdiction primarily—although not exclusively—as a limit on the authority of a given tribunal, an authority that was first and foremost grounded in notions of territoriality. Notice, on the other hand, was viewed primarily as an issue of fairness and justice to a party, usually a defendant, whose rights were to be adjudicated before a given tribunal. Courts used an amalgam of “natural justice”<sup>14</sup> principles, the so-called “general law,” and the international law principle of comity<sup>15</sup> to develop limits on the exercise of personal jurisdiction. Courts also employed the Full Faith and Credit Clause<sup>16</sup> as a constitutional basis for refusing to enforce judgments of other state courts that purportedly lacked personal jurisdiction.<sup>17</sup> As for notice, much of the doctrinal pronouncements came in *in rem* actions, but courts “rarely had occasion to discuss the form that notice had to take in *in personam* actions . . . because [their] personal jurisdiction jurisprudence . . . ensured, as a practical matter, that defendants in such actions received notice through personal service of process.”<sup>18</sup>

There was always some shared space between personal jurisdiction and notice, in particular, the appeals to natural justice and fair-

13. 95 U.S. 714 (1877).

14. See Conison, *supra* note 8, at 1097–1103 (natural justice basis for notice as well as personal jurisdiction limitations both before and after *Pennoyer*).

15. See Conison, *supra* note 8, at 1104–11; Sachs, *supra* note 8, at 1270 (“Early American courts applied what they saw as rules of general and international law to determine whether foreign judgments deserved any respect.”); WASSERMAN, *supra* note 7, at 208–09 (pre-*Pennoyer* limitations on personal jurisdiction were “derived from international law.”).

16. U.S. CONST. art. IV, § 1 (“Full faith and Credit shall be given in each State to the public Acts, Records and judicial Proceedings of every other State.”). See Steven R. Greenberger, *Justice Scalia’s Due Process Traditionalism Applied to Territorial Jurisdiction: The Illusion of Adjudication Without Judgment*, 33 B.C. L. REV. 981, 1015–16 (1992) (“The pre-*Pennoyer* case law that incorporated those principles consequently arose entirely as a problem of the *interstate* recognition of judgments under the Full Faith and Credit Clause and statute.”).

17. WASSERMAN, *supra* note 7, at 208–09.

18. *Id.* at 130.

ness. For example, in the leading pre-*Pennoyer* case of *Lafayette Ins. Co. v. French*, the Supreme Court recognized “that principle of natural justice which requires a person to have notice of a suit before he can be conclusively bound by its result; [and] those rules of public law which protect persons and property within one State from the exercise of jurisdiction over them by another.”<sup>19</sup> Courts periodically cited *Vallee v. Dumergue*, an 1849 English Exchequer case upholding personal jurisdiction where it was supported by “natural justice.”<sup>20</sup> State courts similarly included appeals to principles of fairness and natural justice, and some of these decisions found their way into Justice Field’s *Pennoyer* decision.<sup>21</sup>

The strongest link between personal jurisdiction and notice was not conceptual, but mechanical. The procedures of service of process were, and are, simultaneously the method for notifying a party of the pendency of an action and the procedure by which personal jurisdiction is “perfected.”<sup>22</sup> This link between notice and service of process was always apparent: courts and lawmakers needed some way of dictating and then measuring how service of process should be accomplished and whether such methods were sufficient.

The link between personal jurisdiction and service of process is more a quirk of historical path dependency than one of conceptual

---

19. *Lafayette Ins. Co. v. French*, 59 U.S. 404, 406 (1855). This case also played a prominent role in the introduction of the theory that jurisdiction could be acquired by “implied consent.” See Patrick J. Borchers, *One Step Forward and Two Back: Missed Opportunities in Refining the United States Minimum Contacts Test and the European Union Brussels I Regulation*, 31 ARIZ. J. INT’L & COMP. L. 1, 7 (2014) (in *Lafayette*, “the Supreme Court invented the fiction that a corporation doing business in the forum had implicitly consented to jurisdiction there.”); Charles W. “Rocky” Rhodes, *The Predictability Principle in Personal Jurisdiction Doctrine: A Case Study in the “Generally” Too Broad but “Specifically” Too Narrow Approach to Minimum Contacts*, 57 BAYLOR L. REV. 135, 142 (2005) (citing *Lafayette* as a pre-*Pennoyer* historical precedent for consent in personal jurisdiction).

20. *Vallee v. Dumergue* (1849) 154 Eng. Rep. 1221; 4 Exch. 290. Even after *Pennoyer*, some courts continued to cite the early common law origins of notice and personal jurisdiction doctrines. See *Gilmore v. Sap*, 100 Ill. 297, 302 (1881) (justifying the validity of a substituted service statute by harkening back to the common law in the time of Richard II).

21. See *Pennoyer v. Neff*, 95 U.S. 714, 731–32 (1877). See also *Gillespie v. Commercial Mut. Marine Ins. Co.*, 78 Mass. 201, 201 (1958) (upholding in-state service of process on a non-resident defendant).

22. See *Omni Capital Int’l, Ltd. v. Rudolf Wolff & Co.*, 484 U.S. 97, 104 (1987) (quoting *Miss. Publ’g Corp. v. Murphree*, 326 U.S. 438, 444–45 (1946)) (“[S]ervice of summons is the procedure by which a court having venue and jurisdiction of the subject matter of the suit asserts jurisdiction over the party being served.”).

necessity. At common law, as received from England, the sheriff would physically arrest the defendant pursuant to the writ of *capias ad respondendum*,<sup>23</sup> meaning that “service of civil process did not differ materially from what we know today as criminal arrest. The sheriff physically restrained the person served, and then jailed him or her while he or she awaited disposition of the action.”<sup>24</sup> The act of physically restraining and confining the defendant was intimately connected to the idea that the state was exerting physical control over a person within its territory.<sup>25</sup> Thus the fact that the state *did* control and confine a defendant within its jurisdiction became the one of the foundations for the idea that the state *could* exert adjudicational authority over persons and property within its boundaries.<sup>26</sup> The *capias* was eventually replaced in the mid-eighteenth century by service of process so that by the time of the founding in 1787, “lawyers in England and America had been required to use the summons as the tool for starting suit for more than sixty years; arrest was no longer a tool for commencing suit in most civil actions.”<sup>27</sup> Thus, personal jurisdiction and notice were already on paths sometimes parallel, sometimes intertwined, long before the full constitutionalization in *Pennoyer*.

---

23. *Capias*, BLACK’S LAW DICTIONARY (10th ed. 2014) (defining *capias ad respondendum* as “[a] writ commanding the sheriff to take the defendant into custody to ensure that the defendant will appear in court.”).

24. John Martinez, *Discarding Immunity from Service of Process Doctrine*, 40 OHIO N.U. L. REV. 87, 90 (2013). See also Donald E. Wilkes, Jr., *Habeas Corpus Proceedings in the High Court of Parliament in the Reign of James I, 1603–1625*, 54 AM. J. LEGAL HIST. 200, 218 (2014) (“In seventeenth century England, unlike today, civil arrest process was an integral part of civil procedure.”). For a detailed history of the development and decline of the *capias*, see Nathan Levy, Jr., *Mesne Process in Personal Actions at Common Law and the Power Doctrine*, 78 YALE L. J. 52 (1978).

25. See RICHARD D. FREER, CIVIL PROCEDURE 45 (4th ed. 2017) (“[The *capias*] was a stark reminder that the jurisdiction was being exercised *in personam*, because it actually resulted in taking the defendant into the custody of the government.”).

26. See Levy, *supra* note 24, at 94 (“The common law courts neither exercised nor believed they could exercise jurisdiction in personal actions without either physical custody of the defendant or an appearance by him.”). But see Albert A. Ehrenzweig, *The Transient Rule of Personal Jurisdiction: The “Power” Myth and Forum Conveniens*, 65 YALE L.J. 289, 297 (1956) (“Even when [English courts] began to base [their] personal jurisdiction upon the physical arrest of the defendant, actual physical power over the defendant was not invariably required.”).

27. Bradford E. Biegon, Note, *Presidential Immunity in Civil Actions: An Analysis Based Upon Text, History, and Blackstone’s*, 82 VA. L. REV. 677, 682 (1996).

B. *Personal Jurisdiction and Notice in Pennoyer*

Personal jurisdiction and notice were at the heart of *Pennoyer*,<sup>28</sup> the seminal civil procedure due process case known to (and perhaps dreaded by) all law students over the past century.<sup>29</sup> Although the *Pennoyer* story has been told and retold in many a scholarly commentary, it's worth rehearsing the facts again here to emphasize the juxtaposition of personal jurisdiction and notice. *Pennoyer* concerned a plaintiff, Mitchell, who wanted to sue Neff for unpaid legal fees in Oregon state court. Neff, no longer a resident of Oregon, was residing in California, although he still owned land in Oregon (the land which was, in fact, the subject of the legal advice that Mitchell had tendered). Mitchell served Neff under an Oregon statute that permitted service on an out-of-state defendant via publication for six successive weeks in a newspaper.<sup>30</sup> The Supreme Court, in an opinion by Justice Field, held that service by publication on an out-of-state defendant was insufficient to establish *in personam* jurisdiction over the defendant because the defendant was not notified of the lawsuit and thus unable to defend himself before the entry of a default judgment.<sup>31</sup>

Although the case stands primarily for the proposition that personal jurisdiction is a Fourteenth Amendment due process right, it is also the genesis of locating the right of notice and opportunity to be heard in the due process clause of the Fourteenth

---

28. 95 U.S. 714, 733 (1877) (“Since the adoption of the Fourteenth Amendment to the Federal Constitution, the validity of such judgments may be directly questioned, and their enforcement in the State resisted, on the ground that proceedings in a court of justice to determine the personal rights and obligations of parties over whom that court has no jurisdiction do not constitute due process of law.”). See also *Mullane v. Cent. Hanover Bank & Tr. Co.*, 339 U.S. 306, 314 (1950); Austen L. Parrish, *Sovereignty, Not Due Process: Personal Jurisdiction Over Nonresident Alien Defendants*, 41 WAKE FOREST L. REV. 1, 4 (2006) (“In at least one respect, the doctrinal formulation is thus unmistakable: due process is the starting and ending point to any personal jurisdiction analysis.”).

29. See Linda J. Silberman, *Shaffer v. Heitner: The End of an Era*, 53 N.Y.U. L. REV. 33, 33–34 (1978) (recalling an encounter with a “derelict” in Washington Square Park who proved himself an erstwhile law student by shouting, “Pennoyer!” and recounting the facts and holding of the case).

30. *Pennoyer v. Neff*, 95 U.S. 714, 720, 736 (1877).

31. *Id.* at 727, 733–34. See also Charles W. “Rocky” Rhodes, *Nineteenth Century Personal Jurisdiction Doctrine in a Twenty-First Century World*, 64 FLA. L. REV. 387, 434 n.283 (2012) (“The problem with *Pennoyer* from a modern perspective was not that Neff should not have been subject to jurisdiction in Oregon . . . but that the notice given was not reasonably calculated to inform him of the suit.”).

Amendment.<sup>32</sup> After all, what had “gone wrong” in the original *Mitchell v. Neff* action was a problem with service of process. The *Pennoyer* decision affected both doctrines that service of process underlies: personal jurisdiction and notice.

Courts passing on pre-*Pennoyer* cases had not taken care to erect a strong or formal distinction between personal jurisdiction and notice. While the doctrines were not identical or interchangeable, there was a certain fluidity in how courts handled problematic service of process issues that sometimes implicated jurisdictional concerns, sometimes notice concerns, and sometimes both. *Pennoyer* itself has this character. Justice Field unquestionably placed power, territoriality, and sovereignty at the center of personal jurisdiction.<sup>33</sup> Fairness and natural justice, however, were not absent from the opinion. A natural fit, or even a proxy for the question of fairness, was to evaluate the actual or constructive notice that a given defendant had of a pending action. For Justice Field, notice implicated “that principle of *natural justice* which requires a person to have *notice* of a suit before he can be conclusively bound by its result.”<sup>34</sup> Thus, the *Pennoyer* opinion cemented the Fourteenth

---

32. See Frank. R. Lacy, *Service of Summons and the Resurgence of the Power Myth*, 71 OR. L. REV. 319, 344 (1992) (calling personal jurisdiction “due process I” and notice “due process II”); Sachs, *supra* note 8, at 1300 (“[I]n 1908, the Supreme Court itself identified two requirements of procedural due process: that the court ‘shall have jurisdiction’ (for which it cited *Pennoyer*, among other cases), and that the parties be given ‘notice and opportunity for hearing.’”).

33. Justice Field opens his opinion by citing *D’Arcy v. Ketchum*, 11 How. 165, the canonical pre-*Pennoyer* case establishing the territorial limits on sovereign adjudication that the Court continued to cite well into the Twentieth Century in its personal jurisdiction decisions. See, e.g., *Shaffer v. Heitner*, 433 U.S. 186, 198 n.15 (1977); *Hanson v. Denckla*, 357 U.S. 235, 255 (1958); *Baker v. Baker, Eccles, & Co.*, 242 U.S. 394, 402 (1917); *Bigelow v. Old Dominion Copper Mining & Smelting Co.*, 225 U.S. 111, 134 (1912); *Hilton v. Guyot*, 159 U.S. 113, 183 (1895). See also Rhodes, *supra* note 31, at 392 (“[The] power-based premise functioned reasonably well, at least for natural individual defendants, early in U.S. history. . . . Because travel was difficult, the parties or their property were often present in the forum where the dispute arose. Thus, courts rarely needed to consider the connection, if any, between an individual defendant and the litigation.”).

34. *Pennoyer*, 95 U.S. at 730 (quoting *Lafayette Ins. Co. v. French*, 59 U.S. 404, 406 (1856)) (emphasis added). Indeed, the Court referred on several subsequent occasions to principles of “natural” justice when discussing notice. See, e.g., *Turpin v. Lemon*, 187 U.S. 51, 57 (1902) (“[I]t would appear that the 14th Amendment would be satisfied by showing that the usual course prescribed by the state laws required notice to the taxpayer and was in conformity with natural justice.”); *Spencer v. Merchant*, 125 U.S. 345, 358 (1888) (Matthews, J., dissenting) (“[Notice] is a rule founded upon the first principles of natural justice.”); *St. Clair v. Cox*, 106 U.S. 350, 356 (1882) (citing the “principle of natural justice which requires notice of a suit to a party before he can be bound by it.”).

Amendment as the location for the limits on personal jurisdiction and the minimum requirements of notice. Justice Field treated the Fourteenth Amendment foundation as completely mundane and obvious, when in fact it was both new and not a completely intuitive fit with the due process clause.<sup>35</sup>

Because the Court did not announce personal jurisdiction and notice as formal concepts or categories, the decision reads as one that is mainly about personal jurisdiction, but also maybe about notice, weaving the justifications for both throughout the opinion. Justice Field supported the notice requirement with the preexisting notions of “fairness” and “natural justice.” Both the requirement of notice and the animating concepts behind it are used to support the Court’s conclusions about personal jurisdiction. These notice principles would continue to accompany the development of personal jurisdiction in the decades between *Pennoyer* and *International Shoe* and into the modern era.

*C. Personal Jurisdiction and Notice from Pennoyer Through International Shoe and Mullane*

The *Pennoyer* personal jurisdiction regime lasted until 1945. Although much of the jurisdictional jurisprudence from this period has receded into distant memory, the period from 1877 to 1945 was actually a time of rich doctrinal exploration and growth. The Supreme Court and lower courts struggled within *Pennoyer*’s rigid territorial framework to develop a personal jurisdiction doctrine that kept pace with the fast-changing legal and economic landscape of the United States as it entered the twentieth century.<sup>36</sup>

---

35. See Richard H. Fallon, Jr., *Some Confusion About Due Process, Judicial Review, and Constitutional Remedies*, 93 COLUM. L. REV. 309, 317 (1993) (“[T]he Supreme Court has identified some substantive due process rights that it has not even tried to fit into a two-tiered model. For example, the ‘minimum contacts standard[ ]’ from personal jurisdiction.”).

36. Other scholars have produced far more detailed histories of early personal jurisdiction doctrine. This Article highlights the intersection of personal jurisdiction and notice. For more thorough histories with a broader perspective, see Conison, *supra* note 8; Sachs, *supra* note 8. See generally Ralph U. Whitten, *The Constitutional Limitations on State-Court Jurisdiction: A Historical Interpretive Reexamination of the Full Faith and Credit and Due Process Clauses (Part One)*, 14 CREIGHTON L. REV. 499 (1981) (discussing the history of the relationship between the Full Faith and Credit Clause from early English times up until *Pennoyer*); Ralph U. Whitten, *The Constitutional Limitations on State-Court Jurisdiction: A Historical Interpretive Reexamination of the Full Faith and Credit and Due Process Clauses (Part Two)*, 14 CREIGHTON L. REV. 735, 840 (1981) (examining the history of full faith and credit with due process from early English legal history through the American Civil War and the ratification of the Fourteenth Amendment).

This Section contains a brief summary of the doctrinal development in personal jurisdiction between *Pennoyer* and *International Shoe*. During this time, notice was a visible and regular feature of jurisdictional analysis, although courts were not consistent in their determinations as to when notice mattered and what difference it made in the ultimate outcome of a case.<sup>37</sup>

### 1. The *In Rem* Cases

Many of the cases in the first few decades after *Pennoyer* were dedicated to clarifying the scope of, and justification for, *in rem* jurisdiction.<sup>38</sup> These cases were fertile ground for the development of personal jurisdiction doctrine because *in rem* cases provided one of the few acceptable *Pennoyer* frameworks by which states could effectively assert control over non-resident defendants. These cases were also the location of significant doctrinal development of constitutional notice doctrine since, under *Pennoyer*, substituted service was permissible with respect to *in rem* actions.<sup>39</sup>

---

37. It was also a period in which courts took a few decades to consistently fix due process as the crucial (if still hazily defined) limitation on the exercise of personal jurisdiction. See Conison, *supra* note 8, at 1141 (“For nearly forty years, courts, including the Supreme Court . . . largely failed to treat *Pennoyer* as a constitutional decision. . . . It was not until the Supreme Court’s decision in *Riverside & Dan River Cotton Mills v. Menefee* . . . that [*Pennoyer*’s] status as a constitutional decision was retroactively confirmed.”). During these years, however, some lower federal courts and state courts did, from time to time, refer to the due process basis for personal jurisdiction. See, e.g., Operative Plasterers’ & Cement Finishers’ Intl. Ass’n v. Case, 93 F.2d 56, 63 (D.C. Cir. 1937) (“[I]t is perfectly consistent with due process to provide that jurisdiction over an association doing business shall result from service upon one or more of its members.”); *Shambe v. Del. & H. R. Co.*, 135 A. 755, 757 (Pa. 1927) (“A state has no power to render a personal judgment against a foreign corporation not doing business within the state. A judgment so rendered was held a violation of the due process clause, and void.”) (internal citations omitted).

38. For example, the Court decided [several] cases in which a case that had been styled *in rem* was actually *in personam* because the res at issue had changed hands or was no longer within the territory of the State in a way which would justify the operation of *in rem* jurisdiction. See, e.g., *Nat’l Exch. Bank v. Wiley*, 195 U.S. 257 (1904) (bank notes not an appropriate res when they had been sold prior to the commencement of the suit); *Sec. Sav. Bank v. California*, 263 U.S. 282 (1923); *Wilson v. Seligman*, 144 U.S. 41 (1892) (judgment not binding against a stockholder because a proceeding against stockholders was *in personam*, not *in rem*, and stockholder had not received personal service of process as required by *Pennoyer*).

39. See WASSERMAN, *supra* note 7, at 130 (“During the nineteenth and early twentieth centuries, the Supreme Court rarely had occasion to discuss the form that notice had to take in *in personam* actions . . . because its personal jurisdiction jurisprudence . . . ensured, as a practical matter, that defendants in such actions received notice through personal service of process.”). During this period, courts

In these cases, courts would dutifully point out that actual notice and personal service upon a defendant were not constitutionally required so long as the plaintiff followed the relevant niceties of attachment at the outset of the suit.<sup>40</sup> But the concern about notice was never far from judges' minds.<sup>41</sup> Many of these opinions contain passing references to the fact that a given defendant actually did have notice of the lawsuit,<sup>42</sup> or helpful reminders that ownership of property within the territory would usually give rise to some form of notice.<sup>43</sup> Courts often noted with respect to *in rem* cases that "seizure of the property . . . is a *species of notice* to the non-resident or

---

began filling in the due process elements of notice, namely that there were minimum constitutional standards for service of process, and that the content of the notice itself should sufficiently convey the relevant information regarding the pendency of the lawsuit.

40. See, e.g., *Sec. Sav. Bank*, 263 U.S. at 287 ("the essentials of jurisdiction over the deposits are that there be seizure of the res at the commencement of the suit, and reasonable notice and opportunity to be heard."); *Grannis v. Ordean*, 234 U.S. 385, 395–97 (1914) (in an *in rem* suit, notice by publication sufficient under *Pennoyer* notwithstanding a minor misspelling of the property owner defendant's name); *Bower v. Stein*, 177 F. 673, 677 (9th Cir. 1910) (service by publication for *in rem* suit not set aside despite an error in affidavit as to resident of defendant); *Kirk v. United States*, 137 F. 753, 755 (2d Cir. 1905) ("no man can be deprived of his property without due notice and opportunity to be heard."); *Sutherland-Innes Co. v. Am. Wired Hoop Co.*, 113 F. 183, 187 (8th Cir. 1901) ("when resort is had to substituted service, there is always more or less danger that a judgment may be rendered without actual notice to the defendant"); *Bailey v. Sundberg*, 49 F. 583, 585 (2d Cir. 1892) (admiralty case addressing the notice requirements for an *in rem* [libel] of a ship noted that "notice is as indispensable as the arrest [of the ship] to confer jurisdiction."); *Palmer v. McCormick*, 28 F. 541, 544 (N.D. Iowa 1886) (clarifying standard for publisher's affidavit in constitutionally appropriate service by publication); *Porter v. Duke*, 270 P. 625, 629–30 (Ariz. 1928) (reciting the accepted justifications for notice by publication in *in rem* cases).

41. See, e.g., *Sutherland-Innes Co.*, 113 F. at 187 (value of notice sufficiently strong such that substituted service should be limited to situations "only as might be necessary to enable the courts of the state to effectually enforce . . . property [rights] within their jurisdiction.").

42. See, e.g., *Herbert v. Bicknell*, 233 U.S. 70, 74 (1914) ("[I]t appears that the defendant had knowledge of the action . . .").

43. See, e.g., *Herbert*, 233 U.S. at 73–74 (framing the question as "[r]eally the only matter before us that calls for a word is the decision that a judgment appropriating property within the jurisdiction . . . is not made bad by the short and somewhat illusory notice to the owner" and concluding that, under *Pennoyer*, the assumption that property is always in possession of the owner is sufficient); *Oswald v. Kampmann*, 28 F. 36, 38 (W.D. Tex. 1886) ("[I]f [plaintiff challenging a judgment] saw fit to abandon the country, and pay no attention to the property, she ought not be heard to complain if the law makes an exception to the general rule in her case."); *Geary v. Geary*, 6 N.E.2d 67, 72 (N.Y. 1936) (emphasizing actual notice in addition to the "possession of property" principle to justify jurisdiction).

his agent.”<sup>44</sup> As Justice Bartholomew of the Supreme Court of North Dakota explained about notice and substituted service, service is “*not mere idle form*. It serves a substantial purpose. It is the theory of the law that notice of the pendency of the action is thus brought to the defendant. . . . It is the substituted service that gives notice of the pendency of the action, and that notice is a direct challenge to the defendant to appear and protect his property, if any he have in the jurisdiction.”<sup>45</sup>

Notice, however, would never be taken so far as to overtake the *Pennoyer* barrier between *in personam* and *in rem* bases for jurisdiction. While notice could justify the exercise of jurisdiction over absent defendants in *in rem* cases, it could never, on its own, provide a basis for personal jurisdiction itself. For example, the Supreme Court of Michigan rejected a plaintiff’s argument that an Illinois court had exercised valid *in personam* jurisdiction over the defendant because defendant had actual notice of the lawsuit.<sup>46</sup> Likewise, the concern that a vulnerable defendant might not have received meaningful notice of a lawsuit did not prevent the Ninth Circuit from upholding the exercise of *in rem* jurisdiction procured by attachment and substituted service by publication.<sup>47</sup> Other courts made similar findings, namely, that *Pennoyer*’s allowance of substituted service for *in rem* cases with absent non-residents constituted a constitutionally sanctioned carve-out to the requirement, or even concern, of actual notice.<sup>48</sup>

Although most of the major *in rem* cases involved tangible property located physically within the borders of a forum state, the problem of intangible property offered courts the opportunity to explore the possibilities for boundary pushing in personal jurisdiction.<sup>49</sup> *Harris v. Balk*<sup>50</sup> was one such case. The facts of *Harris* are

---

44. *Dorr v. Gibboney*, 7 F.Cas. 923, 925 (C.C. W.D.Va. 1878) (No. 4006).

45. *Hartzell v. Vigen*, 69 N.W. 203, 207–08 (N.D. 1896) (emphasis added).

46. *See, e.g., Stewart v. Eaton*, 283 N.W 651, 657–58 (Mich. 1939).

47. *Cohen v. Portland Lodge No. 142, B.P.O.E.*, 152 F. 357, 358–62 (9th Cir. 1907).

48. *See, e.g., Bower v. Stein*, 177 F. 673, 676–77 (9th Cir. 1910). However, in one curious lower court case, a federal District Court in New York proclaimed a sort of exception to the seemingly universal rule from *Pennoyer* for an admiralty seaman’s wages case in equity, finding no violation of due process when the defendant “had full actual notice of the suit on the day when it was instituted, though not legally served with process.” *The City of New Bedford*, 20 F. 57, 60 (S.D.N.Y. 1884).

49. *See, e.g., McLaughlin v. Bahre*, 35 Del. 446, 455–56 (Del. Super. Ct. 1933) (seizure of stock without other notice is sufficient under both *Pennoyer* and the common law customary principles that seizure of property constitutes constructive notice).

fairly straightforward: Harris owed a debt to Balk and Balk owed a debt to Epstein. Harris and Balk were both North Carolina residents. While Balk was on a trip to Baltimore, Epstein sued Balk *in rem*, attaching the debt owed from Harris. Neither Harris nor Balk appeared in the Maryland court, but upon his return to North Carolina, Balk arranged for payment to Epstein pursuant to an order of the Maryland court. Harris then sued Balk for the debt in North Carolina. The Maryland judgment would be valid if Maryland was the situs of Balk's debt to Harris.<sup>51</sup> The Supreme Court held the situs of the debt traveled with the debtor. This allowed the Maryland court to "reach" Harris, via the debt that he was owed, in North Carolina. The Court did not cite *Pennoyer*, nor did it linger much on the finer points of personal jurisdiction. Instead, the opinion centers on the debtor-creditor relationship, the situs of the debt, and the obligation of a garnishee to give notice to the creditor of the attachment.<sup>52</sup> The Court did, however, note with approval that Balk did in fact have notice of the attachment, both in fact and because the Maryland attachment procedure required such notice,<sup>53</sup> and the Court ended the opinion with dicta speculating that a failure by the garnishee to notify the creditor would deprive him of using the judgment in the first action as a bar to liability in a second action.<sup>54</sup> Once again, the fact of notice fortified the exercise of jurisdiction. The fairness of jurisdictional innovations was bolstered by the assurance that no one was (or should have been) surprised by jurisdiction nor deprived of the opportunity to be heard.<sup>55</sup>

## 2. The Marriage Exception Cases

Cases concerning the status of a marriage constituted one of the exceptions to *Pennoyer*'s requirement of territorial service, and thus provided another opportunity for doctrinal development. Courts would stress the importance of a state being able to adjudicate the status of a marriage within the state,<sup>56</sup> but then temper that blunt exercise of power with the assurance that, for example, "the

---

50. 198 U.S. 215 (1905).

51. *Id.* at 221–22.

52. *Id.* at 227.

53. *Id.* at 227–28.

54. *Id.* at 228.

55. *Cf. id.* (noting that the creditor would have had "the opportunity to defend himself" in the Maryland lawsuit).

56. *See, e.g.,* *Haddock v. Haddock*, 201 U.S. 562, 572 (1906) ("[N]o question can arise on this record concerning the right of the State of Connecticut within its borders to give effect to the decree of divorce rendered in favor of the husband by

wife [had] actual notice of the suit.”<sup>57</sup> And when a party did not have notice of the suit, a lower court might opt for the most narrow reading of *Pennoyer*’s marriage status exception that emphasized the strength of the power/sovereignty theory as a justification for allowing limited substitute service to out-of-state defendants.<sup>58</sup> Curiously, the Supreme Court used the matrimonial cases as an opportunity to draw a post-*Pennoyer* wedge between the Full Faith and Credit Clause and the Fourteenth Amendment constitutionality of personal jurisdiction itself. In *Haddock v. Haddock*, the Court held that constructive service pursuant to a Connecticut statute on a non-resident spouse was sufficient for a divorce decree to be enforceable *within* Connecticut. Another state would be *permitted* to enforce the decree on public policy grounds if it so chose, but other states were not *required* to enforce such a decree as a matter of Full Faith and Credit because this would create a sort of “race to the bottom” in which states with lax residency laws for divorce would attract ill-motivated spouses seeking to abandon their marriages and obtain a divorce decree in a favorable jurisdiction.<sup>59</sup> This was likely part of a larger project in which the Court was loath for the federal courts to get too involved in questions of the state law of domestic relations,<sup>60</sup> and thus can be seen as a (perhaps unprincipled) exception to full faith and credit, more than as a case of the Court making inconsistent decisions about personal jurisdiction over a defendant based on whether enforcement was sought within the state or extraterritorially.<sup>61</sup>

---

the courts of Connecticut, he being at the time when the decree was rendered domiciled in that State.”).

57. *Atherton v. Atherton*, 181 U.S. 155, 172 (1901).

58. *See De la Montoya v. De la Montoya*, 44 P. 345, 348 (Cal. 1896) (“The idea that domicile determines jurisdiction in divorce rests upon the assumption that *status* depends on domicile, and is of interest there only. Judge Field could not have had this in mind in *Pennoyer v. Neff*. . . when he speaks of ‘absent defendants’ he cannot mean those not domiciled within the state, but must have meant simply those physically absent, and upon whom, therefore, personal service of process could not be made.”).

59. 201 U.S. at 575–77. *See also* Neal R. Feigenson, *Extraterritorial Recognition of Divorce Decrees in the Nineteenth Century*, 34 AM. J. LEGAL HIST. 119, 129–60 (1990) (detailing the history of the cases and history leading to *Haddock v. Haddock*).

60. *See* the domestic relations exception to subject matter jurisdiction, which “divests the federal courts of power to issue divorce, alimony, and child custody decrees.” *Ankenbrandt v. Richards*, 504 U.S. 689, 703 (1992).

61. *See* Feigenson, *supra* note 59, at 125–29.

### 3. Corporations and Consent Cases

The cases about service on an out-of-state corporation formed an important bridge between the territorial rigidity of the *Pennoyer* holding and the more abstract regime to come in *International Shoe*. As we shall see, expanding the viability of implied consent to jurisdiction was a major procedural innovation that ultimately culminated in the minimum contacts test, and the corporations cases were key in developing the doctrinal prerequisites to thinking broadly about the role of consent. Corporations were useful tools of jurisdictional expansion because if a business registered or otherwise affiliated itself with the forum state in a statutorily prescribed manner, “the state official was considered the corporation’s agent, [and] in-state service on the official was deemed valid service on the defendant, regardless of whether the official or the plaintiff made any attempt to notify the corporation itself.”<sup>62</sup>

In *St. Clair v. Cox*,<sup>63</sup> the Supreme Court held that a state could exercise personal jurisdiction over foreign corporations when jurisdiction was secured by service of process on designated agents of the corporation. Before *International Shoe* introduced minimum contacts as the constitutionally acceptable substitute for fictive corporate “presence” within a state, *St. Clair* stood for the proposition that when one serves a corporation’s authorized agent within the state, the corporation must also be doing business within the state.<sup>64</sup> Justice Field, who also penned *Pennoyer*, cautioned that this exercise of jurisdiction “must not . . . encroach upon that principle of natural justice which requires notice of a suit to a party before he can be bound by it”<sup>65</sup> and the notice itself “must be reasonable.”<sup>66</sup>

---

62. Rhodes, *supra* note 31, at 394.

63. 106 U.S. 350 (1882).

64. *See, e.g.*, *Frawley, Bundy & Wilcox v. Penn. Cas. Co.*, 124 F. 259, 263 (C.C.M.D. Pa. 1903) (citing *St. Clair v. Cox* for the proposition that “it is essential in every case in which personal jurisdiction over such a corporation is claimed that there shall have been an actual and substantial transacting of business by it within the state.”). *See also* *Riverside & Dan River Cotton Mills v. Menefee*, 237 U.S. 189, 193–94 (1915); *Hazeltine v. Miss. Valley Fire Ins. Co.*, 55 F. 743, 745 (C.C.W.D. Tenn. 1893); *United States v. Am. Bell Tel. Co.*, 29 F. 17, 35 (C.C.S.D. Ohio 1886); *Davidson v. Henry L. Doherty & Co.*, 241 N.W. 700, 703 (Iowa 1932) (upholding statute that allows substituted service on the agent of a foreign corporation as one that “meets every essential requirement of due process of law.”).

65. *St. Clair v. Cox*, 106 U.S. 350, 356 (1882).

66. *Id.* Some courts issued opinions that read like the forerunners of post-*International Shoe* jurisprudence, such as the Sixth Circuit’s decision in *Smith v. Farbenfabriken of Elberfeld Co.*, 203 F. 476 (6th Cir. 1913) which upheld service in patent infringement actions under a federal statute that allowed service on a business’s agent “conducting business” within the state. In finding both the statute

At times, the Court lapsed back into formalism, holding that strict compliance with a statute that allowed for substituted service of a corporation on a Secretary of State was sufficient, even when the state official did not provide any further notice of the pendency of an action to the defendant itself.<sup>67</sup> This pattern would repeat itself many times in the subsequent decades: having used notice as a functional due process crutch for creating a jurisdictional innovation, the Court would then treat the new rule as one with its own formal identity and justification, sometimes unmooring it from the original justification or connection to due process.<sup>68</sup>

*Hess v. Pawloski*<sup>69</sup> is often cited as a case that marks the beginning of the transition to the modern era of personal jurisdiction.<sup>70</sup> In *Hess*, the Court approved the use of non-resident motor vehicle statutes as a method of securing jurisdiction over out-of-state defendants on a theory of implied consent.<sup>71</sup> Much of the commentary on *Hess* focuses on how the Court stretched consent, which had always been a common law basis for exercising jurisdiction,<sup>72</sup> as a means to begin building a bridge between the strict territorial regime of *Pennoyer* to the permissibility of more modern long-arm

---

and its application valid, the court rehearsed a list of the business's contacts in the state of Michigan and found that service upon an agent pursuant to the federal statute fell unquestionably within the boundaries of *Pennoyer*.

67. *Washington ex rel. Bond & Goodwin & Tucker, Inc. v. Superior Court of Washington*, 289 U.S. 361 (1933).

68. *See infra* Part III.B.

69. 274 U.S. 352 (1927).

70. *See, e.g.*, Patrick J. Borchers, *Extending Federal Rule of Civil Procedure 4(k)(2): A Way to (Partially) Clean Up the Personal Jurisdiction Mess*, 67 AM. U. L. REV. 413, 440 (2017) (noting that non-resident motor vehicle statutes such as the Massachusetts statute in *Hess* “began pushing the common law’s jurisdictional bounds.”); Rhodes, *supra* note 31, at 393; Arthur T. von Mehren & Donald T. Trautman, *Jurisdiction to Adjudicate: A Suggested Analysis*, 79 HARV. L. REV. 1121, 1149 (1966) (citing *Hess* as a “seminal” case in the modern jurisdictional era). Prior to *Hess*, the Supreme Court and some lower courts had tied jurisdiction via consent to the fact that a party could consent to notice by substituted service such as an agent or by publication. *See, e.g.*, *Lafayette Ins. Co. v. French*, 59 U.S. 404, 408 (1855) (service on an agent in a state “foreign” to the “company’s creation” is permitted); *Michigan Tr. Co. v. Ferry*, 175 F. 667, 673 (8th Cir. 1910) (becoming an executor of an estate was an “office tendered on [the defendant] on the condition imposed by these statutes that the probate court should have the power to call him before it . . . without other warning than a notice published in a newspaper.”).

71. 274 U.S. at 357 (“in advance of the operation of a motor vehicle on its highway by a nonresident, the state may require him to appoint one of its officials as his agent on whom process may be served in proceedings growing out of such use.”).

72. *See Hill v. Mendenhall*, 88 U.S. 453 (1874) (pre-*Pennoyer* case holding that voluntary appearance in an action confers personal jurisdiction on a court).

statutes.<sup>73</sup> But, although consent formed an important doctrinal foundation, the value of notice was not far behind. The Court took care to note that under the Massachusetts statute, “[i]t is required that [the defendant] shall *actually receive* and receipt for notice of the service and a copy of the process.”<sup>74</sup> Echoing the “opportunity to be heard” aspect of notice, the Court noted with approval that the statute “contemplates such continuances as may be found necessary to give reasonable time and opportunity for defense.”<sup>75</sup> In other words, although the Court did not yet make explicit appeals to “fairness” or “reasonableness,” such equitable concerns were clearly at the forefront of the move to expand jurisdiction over absent defendants, and notice was a central value.<sup>76</sup>

#### 4. The End of the Journey to *International Shoe*

*Milliken v. Meyer*<sup>77</sup> is our last stop on the journey from *Pennoyer* to *International Shoe*. Milliken served Meyer, a Wyoming resident, with process in Colorado under a Wyoming statute that permitted out-of-state service on Wyoming residents. The Supreme Court upheld the constitutionality of this *in personam* out-of-state service on the theory that “[d]omicile in the state is alone sufficient to bring an absent defendant within the reach of the state’s jurisdiction for purposes of a personal judgment by means of appropriate substituted service.”<sup>78</sup> It is in this case, just before the watershed of the

---

73. See, e.g., Andrew D. Bradt & D. Theodore Rave, *Aggregation on Defendants’ Terms: Bristol-Myers Squibb and the Federalization of Mass-Tort Litigation*, 59 B.C. L. REV. 1251, 1269 (2018) (*Hess* was one of the “legal fictions to accommodate the *Pennoyer* regime to modern problems” which “the Supreme Court seemingly abandoned in 1945 with *International Shoe*.”); Rhodes, *supra* note 19, at 144–45 (*Hess* and other implied consent cases as a part of the evolution of jurisdiction over out of state defendants from *Pennoyer* to *International Shoe*); Verity Winship, *Jurisdiction Over Corporate Officers and the Incoherence of Implied Consent*, 2013 U. ILL. L. REV. 1171, 1187 (2013) (implied consent broadened jurisdiction after *Pennoyer* but was mostly abandoned as unnecessary after *International Shoe*).

74. *Hess*, 274 U.S. at 356.

75. *Id.*

76. The fact that a defendant had been notified pursuant to service on the secretary of state was also integral to the Supreme Court’s holding in *Wuchter v. Pizzutti*, 276 U.S. 13 (1928).

77. 311 U.S. 457 (1940).

78. *Id.* at 462. Although in hindsight it might seem like an obvious and foregone conclusion that domicile was a common law and thus *per se* constitutional basis of personal jurisdiction, this was not necessarily understood to courts or jurists pre-*Milliken*. For example in 1896, Justice Temple of the Supreme Court of California declared that “[d]omicile has never, so far as I am aware, been made the test of jurisdiction to render a personal judgment.” *De la Montoya v. De la Montoya*, 44 P. 345, 346 (Cal. 1896). See also *Raher v. Raher*, 129 N.W. 494, 499 (Iowa

minimum contacts era, that the Court set forth one of its strongest ties between notice and the due process basis for personal jurisdiction:

Its adequacy so far as due process is concerned is dependent on whether or not the form of substituted service provided for such cases and employed is reasonably calculated to give him actual notice of the proceedings and an opportunity to be heard. If it is, the traditional notions of fair play and substantial justice implicit in due process are satisfied.<sup>79</sup>

This passage is rather amazing in hindsight. Because it is *this very passage* that sets up the magic “fair play and substantial justice” words that formulate the minimum contacts test in *International Shoe*. Here, in *Milliken*, notice and opportunity to be heard forms the heart of the fairness argument that justifies the extension of *in personam* jurisdiction to the exercise of a long-arm statute. The appeal to due process fairness is not the only basis for the Court’s decision. Justice Douglas stressed that the “authority of a state over one of its citizens is not terminated by the mere fact of his absence from the state.”<sup>80</sup> But even this pronouncement is eventually capped by a reference back to notice. After noting that power over domiciliaries is “not dependent on continuous presence in the state,” Justice Douglas approved the use of out-of-state service “where the state has provided and employed a reasonable method for apprising such an absent party of the proceedings against him.”<sup>81</sup>

The cases I have summarized in this section show that notice played a quiet but relatively consistent role in pushing personal jurisdiction doctrine forward from the rigid formalism of *Pennoyer* to the modern functionalism of *International Shoe*. It remained the case, however, that the main animating theories behind *Pennoyer*—

---

1911) (finding as to cases involving residents served outside of the state that “in many cases . . . are to be found unqualified statements that the laws of the state have no extraterritorial force, and that process served outside the state will not give the court jurisdiction.”). The Supreme Court of Arizona characterized *Milliken* as a case in which the U.S. Supreme Court “receded from some of the implications of *Pennoyer v. Neff*.” *D.W. Onan & Sons v. Superior Court*, 179 P.2d 243, 262 (Ariz. 1947).

79. *Meyer*, 311 U.S. at 463. The Court went on to assure the reader that Meyer did, in fact, receive actual notice of the lawsuit.

80. *Id.* (“[T]he authority of a state over one of its citizens is not terminated by the mere fact of his absence from the state. The state which accords him privileges and affords protection to him and his property by virtue of his domicile may also exact reciprocal duties.”).

81. *Id.* at 464.

power, sovereignty, and territoriality dominated most of the analysis. This Article does not suggest otherwise. However, even cases that have come to stand for the strongest foundations of the power theory may not be so clear cut. Consider the 1917 Justice Holmes decision in *McDonald v. Mabee*.<sup>82</sup> The lawsuit at issue was filed in Texas, and the defendant was domiciled in that state at the outset of the suit. After drifting in and out of the state for a bit, the defendant finally established a new domicile in Missouri. The plaintiffs served the defendant under a Texas statute that permitted service on an absent defendant for four successive weeks in a newspaper. The Supreme Court had not yet affirmatively held that a state could exercise personal jurisdiction over an absent domiciliary,<sup>83</sup> so service was necessary as a predicate both for notice and for personal jurisdiction. The *McDonald* opinion is but five paragraphs long, but it contains all of the confusion of the past and future of personal jurisdiction regarding its theoretical bases. Justice Holmes declared that “[t]he foundation of jurisdiction is physical power,”<sup>84</sup> but the decision is riddled with concerns about notice. The Court found that, as far as the due process considerations of service were concerned, perhaps “a summons left at his last and usual place of abode would have been enough.”<sup>85</sup> This sentiment was preceded by a conspicuous notation that the defendant still had family in the state, thus implying that the defendant would be more likely to learn of the suit. Holmes quickly returned to tying power together with notice: “We repeat, also, that the ground for giving . . . effect to a judgment is that the court rendering it had acquired *power* to carry it out; and that it is going to the extreme to hold such power gained even by *service* at the last and usual place of abode.”<sup>86</sup>

Beyond emphasizing sovereignty and territoriality, the Court did not always treat notice and personal jurisdiction as identical or

---

82. 243 U.S. 90 (1917). Scholars frequently cite this case because of Holmes’ famous formulation that “the foundation of jurisdiction is physical power.” See, e.g., Rhodes, *supra* note 19, at 143; Arthur Taylor von Mehren, *Adjudicatory Jurisdiction: General Theories Compared and Evaluated*, 63 B.U. L. REV. 279, 283 (1983); Debra Windsor, *How Specific Can We Make General Jurisdiction: The Search for a Refined Set of Standards*, 44 BAYLOR L. REV. 593, 595 (1992).

83. That decision would not come until 1940 in *Milliken v. Meyer*. 311 U.S. 457 (1940).

84. *McDonald*, 243 U.S. at 91.

85. *Id.* As one scholar has noted, however, it is unclear exactly what Holmes meant to endorse here in terms of the specifics of service. See Arthur F. Greenbaum, *The Postman Never Rings Twice: The Constitutionality of Service of Process by Posting After Greene v. Lindsey*, 33 AM. U.L. REV. 601, 614 (1984).

86. *McDonald*, 243 U.S. at 91 (emphasis added).

interchangeable as a matter of due process and in the underlying appeals to natural justice. In *Baker v. Baker, Eccles & Co.*,<sup>87</sup> for example, the Court stressed the concept of notice as fundamental to the elevation of both personal jurisdiction and notice to constitutional due process values, noting that “[t]he fundamental requisite of due process in judicial proceedings is the opportunity to be heard. To hold one bound by the judgment who has not had such opportunity is contrary to the first principles of justice.”<sup>88</sup> However, the Court went on to clarify that personal jurisdiction did have a due process foundation independent of the notice and opportunity to be heard justification: “to assume that a party resident beyond the confines of a state is required to come within its borders and submit his personal controversy to its tribunals *upon receiving notice* . . . is a futile attempt to extend the authority and control of a state beyond its own territory.”<sup>89</sup> In *Baker*, it appears that the defendant in the original action—a bereaved mother who would later claim an entitlement to shares of her deceased son’s business—was not served with notice, and the lack of actual notice clearly vexed the Court. Nonetheless, it is significant that Justice Pitney took pains to note that, while “opportunity to be heard” appears to be at the heart of due process, the territorial concerns had their own jurisdictional merit.

As this Part has shown, the law of personal jurisdiction and notice went through a great deal of doctrinal development from the time prior to *Pennoyer* up through the era directly preceding *International Shoe*. Not only did the doctrines grow and change, but also there was a good deal of variation and inconsistency among the cases, given that the era was one of supposedly “strict” rigidity.

That being said, there is one generalization worth making about the law during this period about the development of notice doctrine. This was an era in which two eventually-distinct concepts of notice were merged: the concept of notice of suit and notice of jurisdiction. To the extent that courts were concerned at all with notice during this period,<sup>90</sup> they were primarily focused on actual or constructive notice of a pending lawsuit. Notice of jurisdiction was subsumed into the concept of notice of suit because of the joint

---

87. 242 U.S. 394 (1917).

88. *Id.* at 403.

89. *Id.*

90. See WASSERMAN, *supra* note 7 (noting that even cases of notice of suit were limited during this period because *Pennoyer*’s in-hand service requirement ensured that questions of notice only came up in the cases that fell into *Pennoyer*’s exceptions).

structure of personal jurisdiction and notice. The foundational rule was that, in most cases, personal jurisdiction was perfected by in-hand personal service within the territory of the forum state. It would not have occurred to people at that time to add an element of notice of jurisdiction to this scheme—they would have assumed that people understood that physical presence in the territory of the forum state was sufficient to subject them to the jurisdiction of that state in at least a limited fashion. Thus, additional “notice” of jurisdiction would have been redundant.

The same can be said of *in rem* jurisdiction—ownership of property within the forum state was itself notice that the state had jurisdiction over said property. This is why the primary question about jurisdiction over property was with notice of suit—the concern was that absent property owners might not learn of a pending action. The conclusion that property owners could or should be aware of the status of their property and thus be aware of any seizures or notices was almost always sufficient to satisfy the due process components of personal jurisdiction and notice.

The other possibilities for jurisdiction over out-of-state defendants were similarly constructed to include an element of notice of jurisdiction. The marriage and corporate status exceptions shared with *in rem* the conceptual foundations of adjudicative power and jurisdiction. The concepts of consent, both express and implied, have an even stronger link—that notice of jurisdiction is bound up with the act giving rise to consent. Personal jurisdiction over absent domiciliaries was analogous to in-hand service within the territory; it was simply assumed that a person domiciled within a state would understand that she was subject to its jurisdiction.

This, then, was the hierarchy of personal jurisdiction and notice prior to *International Shoe*. Personal jurisdiction took center stage, since the doctrine greatly restricted the availability of substituted service or service outside of the forum state. And notice doctrine was concerned almost entirely with notice of suit rather than notice of jurisdiction because personal jurisdiction itself was constructed so that notice of jurisdiction was nigh synonymous with its exercise. It is with this doctrinal backdrop in mind that I turn to the beginning of the modern era of personal jurisdiction and notice.

II.  
PERSONAL JURISDICTION IN *INTERNATIONAL  
SHOE AND MULLANE*

The decisions in *International Shoe* (1945)<sup>91</sup> and *Mullane v. Central Hannover Bank & Trust Co.* (1950)<sup>92</sup> came during a larger era of major procedural change in American jurisprudence.<sup>93</sup> *International Shoe* ushered in the modern era of personal jurisdiction jurisprudence by actively unchaining *in personam* jurisdiction from the rigid territorial sovereignty regime of *Pennoyer*.<sup>94</sup> *Mullane* marked the beginning of the modern era of notice jurisprudence by articulating that the due process right of “notice and opportunity to be heard” requires notice that is “reasonably calculated under all the circumstances, to apprise interested parties of the pendency of the action and afford them an opportunity to present their objections.”<sup>95</sup>

A. *International Shoe and the Alternative History of a Minimum  
Contacts Test for Notice*

*International Shoe* broke personal jurisdiction free from the *Pennoyer* framework where jurisdiction was tightly bound to notions of territoriality and sovereignty.<sup>96</sup> The Court held that Washington

91. *Intl. Shoe Co. v. Washington*, 326 U.S. 310 (1945).

92. 339 U.S. 306, 314 (1950)

93. 1938 saw the introduction of the Federal Rules of Civil Procedure and the famous *Erie* case. 1940 ushered in the era of modern class action jurisprudence with *Hansberry v. Lee*, 311 U.S. 32, (1940) and modern forum non conveniens doctrine was born in 1947 in *Gulf Oil Corp. v. Gilbert*, 330 U.S. 501 (1947), a case which led to the codification of transfer of venue within the federal court system.

94. See, e.g., Douglas D. McFarland, *Dictum Run Wild: How Long-Arm Statutes Extended the Limits of Due Process*, 84 B.U. L. REV. 491, 492–93 (2004) (*International Shoe* “transformed” personal jurisdiction analysis); William M. Richman, *Understanding Personal Jurisdiction*, 25 ARIZ. ST. L. J. 599, 599 (1993) (“in *International Shoe Co. v. Washington*, the Court rejected the rigid territorialism of *Pennoyer v. Neff*”); Jane Rutherford, *The Myth of Due Process*, 72 B.U. L. REV. 1, 37 (1992) (“While *Pennoyer* cast the due process question of jurisdiction in terms of territorial power by asking, ‘is the defendant there?’, *International Shoe* changed the question to, ‘is it fair?’”).

95. *Mullane*, 339 U.S. at 314.

96. See, e.g., Donald L. Doernberg, *Resolving International Shoe*, 2 TEX. A&M L. REV. 247, 260 (2014) (describing how *International Shoe* broke personal jurisdiction from the *Pennoyer* mold). But see Rhodes, *supra* note 31, at 390 (contesting “the familiar story . . . that *International Shoe* . . . wrought a fundamental change” in personal jurisdiction doctrine.); Andrew L. Strauss, *Where America Ends and the International Order Begins: Interpreting the Jurisdictional Reach of the U.S. Constitution in Light of a Proposed Hague Convention on Jurisdiction and Satisfaction of Judgments*, 61 ALB. L. REV. 1237, 1250-51 n.70 (1998) (“The Supreme Court formally brought an

State could exercise personal jurisdiction over the Missouri-headquartered Delaware corporation because “due process requires only that . . . if [the defendant] be not present within the territory of the forum, he have certain minimum contacts with it.”<sup>97</sup>

Justice Stone traced the history by noting that—

[h]istorically the jurisdiction of the courts to render judgment in personam is grounded on their de facto power over the defendant’s person. Hence his presence within the territorial jurisdiction of court was prerequisite to its rendition of a judgment personally binding him. But now that the *capias ad respondendum* has given way to personal service of summons or other form of notice, due process,” only minimum contacts with the forum state are required.<sup>98</sup>

This passage offers the first clue into the future detachment of notice from personal jurisdiction. The Court portrays personal service and “other forms” of notice as mechanisms that exist apart from due process, or certainly apart from the due process considerations of personal jurisdiction. One mechanism, the “*capias ad respondendum*,” simply gave way to new mechanisms. The focus of due process analysis would no longer be on the mechanism, but on presence and its newfound alternative: minimum contacts.

This is the very paragraph in which the Court introduces its enduring formulation of minimum contacts; that minimum contacts are “such that the maintenance of the suit does not offend ‘traditional notions of fair play and substantial justice.’”<sup>99</sup> But, we have already seen the origin of this language. It comes from *Milliken* where the Court’s assessment that jurisdiction did meet the “traditional notions of fair play and substantial justice” emerged from its explicit satisfaction that the defendant had actual notice of the lawsuit and that the Wyoming statute provided for adequate notice.

In *Milliken*, the Court wove its discussion of notice directly into the conclusion that Wyoming could exercise personal jurisdiction over a domiciliary served out of state.<sup>100</sup> In *International Shoe*, however, Justice Stone took up notice as a due process issue distinct from personal jurisdiction. Having finished the explanation of min-

---

end to the era of territorial jurisdiction when it explicitly articulated the new minimum contacts standard for asserting personal jurisdiction in the case of *International Shoe Co. v. Washington*.”).

97. *Intl. Shoe Co. v. Washington*, 326 U.S. 310, 316 (1945).

98. *Id.*

99. *Int’l Shoe Co.*, 326 U.S. at 316 (quoting *Milliken v. Meyer*, 341 U.S. 457, 463 (1940)).

100. See *supra* notes 77-81 and accompanying text.

imum contacts, he wrote that the Court is “*likewise* unable to conclude that the service of process . . . was not sufficient notice of the suit.”<sup>101</sup> If anything, it is personal jurisdiction that supports the conclusion that notice was sufficient under the due process clause, rather than relying on the fact of notice to support personal jurisdiction. In upholding service of process by registered mail, the Court opined that the lawsuit was sufficiently related to [the defendant’s] activities,” such that it “rendered International Shoe’s agent an appropriate “vehicle for communicating the notice.”<sup>102</sup>

Justice Stone further remarked that “[i]t is enough that appellant has established such contacts with the state that the particular form of substituted service adopted there gives reasonable assurance that the notice will be actual.”<sup>103</sup> This is, in retrospect, a remarkable sentence, for it contains the seeds of what might have been, in another world, the new test for notice.

One can imagine an alternate history in which *International Shoe* set out the minimum contacts test which would be used both for personal jurisdiction and for notice.<sup>104</sup> This would have offered some continuity with *Pennoyer* in which a single mechanism (in state personal service) fulfilled the due process requirements for both *in personam* jurisdiction and notice.

In this alternate world, courts would have operated under the assumption that due process sets the floor for personal jurisdiction and notice, and that both are governed by a minimum contacts test as set out in *International Shoe*. Much ink might have been spilled in parsing the difference for what “personal jurisdictional minimum contacts” means or requires versus “notice minimum contacts.” Of course, that world never unfolded. The parsing of minimum contacts has indeed been robust, but it is confined to personal jurisdiction. Notice remained, for the most part, disconnected from minimum contacts, and soon found its own test, and its own path, cemented five years later in *Mullane*.

For now, it is enough to see that *International Shoe* was a remarkable inflection point in the doctrinal journal of personal jurisdiction and notice. The Court spoke directly to the issue of notice of suit, passing favorably on that fact both in terms of the lawsuit at

---

101. *Int'l Shoe*, 326 U.S. at 320 (emphasis added).

102. *Id.*

103. *Id.*

104. At least one State Supreme Court, three years after *International Shoe*, characterized the historical development of jurisdiction from *Pennoyer* to *International Shoe* as one that was intimately bound up with the due process requirements of notice. See *Wein v. Crockett*, 195 P.2d 222 (Utah 1948).

hand and the Washington state statute that authorized out-of-state service. The Court also connected minimum contacts to the idea of notice of jurisdiction, thus foreshadowing how the due process focus on notice in personal jurisdiction would slowly shift from notice of suit to notice of jurisdiction. Finally, the Court wrote about notice as a concept that was integral to the due process inquiry of personal jurisdiction, but also characterized notice as a due process doctrine that is separate from personal jurisdiction. Although other pre-*International Shoe* cases show a similar ambivalence about the relationship, the separation in *International Shoe*, however casual, was a sign of the more formal split to come.

B. *Mullane and the New Trajectory of a Distinct Standard for Due Process in Notice*

Just five years after *International Shoe*, the Supreme Court issued its decision in *Mullane*,<sup>105</sup> the seminal notice case of the modern era. A trust company for a common trust fund brought an action for a judicial accounting mandated under New York banking law that provided for notice to the beneficiaries via publication for four successive weeks in a newspaper. The Supreme Court held that service by publication was insufficient for the known beneficiaries of the trust but was sufficient for the unknown beneficiaries who could not be found.<sup>106</sup>

*Mullane* was an ideal case for the Court to establish a reasonableness standard for notice because, in one fact pattern, it allowed the Court to compare and contrast differently situated parties affected by a legal proceeding and delineate the due process floor for each of them. Service by publication to all beneficiaries would not be “a reliable means of acquainting interested parties”<sup>107</sup> of the pendency of an action, but requiring personal service to all beneficiaries including the unknown beneficiaries “would place impossible or impractical obstacles”<sup>108</sup> to maintaining the lawsuit. Additionally, service by publication to the future or unknown beneficiaries was sufficient under these particular circumstances because

---

105. 339 U.S. 306 (1950).

106. *Id.* at 317. The Court here spoke both of beneficiaries who could not be found with reasonable due diligence, as well as beneficiaries with “conjectural or future” interests whose whereabouts might be required to be ascertained under other circumstances.

107. *Id.* at 315; *id.* at 318 (“Where the names and post office addresses of those affected by a proceeding are at hand, the reasons disappear for resort to a means less likely than the mails to apprise them of [the action’s] pendency.”).

108. *Id.* at 314.

the beneficiaries shared common interests, and thus it was reasonable to believe that the known beneficiaries could “safeguard the interests of all.”<sup>109</sup>

Even though *Mullane* is primarily a decision about notice, it is important to remember that the Court also passed on questions of personal jurisdiction because *Mullane* challenged the jurisdiction of the New York court to hear the action and issue a binding judgment as to all of the beneficiaries. Like the defendant in *Pennoyer*, *Mullane* made constitutional objections both to the exercise of power and to the mechanism of service. But unlike *Pennoyer*, the Court had little problem disposing of the jurisdictional issue. The Court acknowledged that there was some uncertainty over whether the action was *in rem* or *in personam*, but that, regardless of the classification, a judgment would bind known and unknown beneficiaries, many of whom were outside the state of New York and were not served within the state.<sup>110</sup> Justice Jackson breezily dispensed with the jurisdictional question, holding that states had an interest in providing for such accounting proceedings, and that it is “beyond doubt the right of its courts to determine the interests of all claimants, resident or nonresident, provided its procedure accords full opportunity to be heard.”<sup>111</sup> In other words, the power of the court wasn’t the *real* question, or a close call. The real question here was notice.

To appreciate *Mullane*’s place in the shared genealogy of personal jurisdiction and notice, it is useful to investigate the authority that Justice Jackson cites—and that which he omits—following the famous “reasonably calculated under the circumstances” language. He first cites *Milliken*, the very same case that lent the words “traditional notions of fair play and substantial justice” to *International Shoe*’s minimum contacts standard.<sup>112</sup> It seems unlikely that the *Milliken* Court itself thought of *Milliken* as a major notice case. The service of process in *Milliken* was quite ordinary—personal service on the defendant.<sup>113</sup> The only wrinkle was that the Wyoming resident was served outside of the borders of Wyoming.

*Milliken* was not a case about the adequacy of notice. No one was wondering in that case whether the defendant was actually ap-

---

109. *Id.* at 319.

110. *Id.* at 312–13.

111. *Id.* at 313.

112. *Id.* at 314.

113. *Milliken v. Meyer*, 311 U.S. 457, 459 (1940) (“*Meyer*, who was asserted to be a resident of Wyoming, was personally served with process in Colorado pursuant to the Wyoming statutes.”).

prised of the action. Both the facts of the case and the logical inferences about personal service made that clear. *Milliken* was about power, pure and simple, and about whether it was a violation of due process for the state to exercise jurisdiction over a domiciliary, even if service occurred outside of the state borders. But *Milliken* did not really add anything interesting to the law of notice itself. It does not seem that there was any confusion prior to *Milliken* vis-à-vis notice doctrine about whether in-hand service was an adequate mechanism for apprising a defendant of an action, nor did there appear to be concerns with whether personal service outside of a state was somehow less likely to apprise a party of an action than process served personally within a state's borders. Outside of its personal jurisdiction holding, *Milliken* simply stands for the proposition that notice and opportunity to be heard is a crucial part of due process. After *Milliken*, Justice Jackson supported the notice formulation with more traditional cases like *Grannis v. Ordean*<sup>114</sup> and *Roller v. Holly*,<sup>115</sup> cases that in turn grounded the due process notice right in *Lafayette Insurance Company v. French*, the same pre-*Pennoyer* case that located notice and jurisdiction in principles of "natural justice."<sup>116</sup>

Justice Jackson also cited *Hess v. Pawlowski*, as a favorable example of a method of service "that is in itself reasonably certain to inform those affected."<sup>117</sup> Recall that *Hess* occupied a more controversial space regarding notice than did *Milliken*, belonging to the group of cases in which the Supreme Court found and then reaffirmed that substituted service on a state officer, who would then attempt to find and serve the relevant defendant, did indeed give sufficient notice to defendants such that they did not lose the ability to appear and defend themselves in a proceeding.<sup>118</sup>

It is important to observe that the Court did *not* cite *International Shoe*. In fact, the Court actually cited no authority at all for its personal jurisdiction holding. It rejected the strict framework of *Pennoyer* without identifying *International Shoe* as the recent source of that rejection. Justice Jackson then asserted, citing no authority at all, that personal jurisdiction is primarily justified on the State's interest. The omission is all the more puzzling considering that the *International Shoe* Court nodded in the direction of recognizing the interests of the forum state via the observation that "sufficient contacts or ties with the state of the forum to make it reasonable and

---

114. 234 U.S. 385 (1914).

115. 176 U.S. 398 (1900).

116. See *supra* note 19 and accompanying text.

117. *Mullane*, 339 U.S. at 315.

118. See *supra* Part I.C.3.

just . . . to permit the state to enforce the obligations which appellant has incurred there.”<sup>119</sup> It would not have been a stretch to reinforce the importance of New York’s interest in adjudicating the rights of absent beneficiaries with a comparison to Washington state’s adjudicational interest in *International Shoe*, particularly because the *International Shoe* Court used minimum contacts to support its holding that Washington state could levy the unemployment tax in addition to its holding regarding personal jurisdiction.<sup>120</sup> *International Shoe*’s absence in *Mullane* is also notable considering that the *International Shoe* Court’s discussion of notice implied that minimum contacts might provide the constitutional standard for evaluating whether notice comports with due process.<sup>121</sup>

It is understandable, however, why the Court did not return to the minimum contacts concept here to provide a standard. It would have been quite a leap to go from the “systematic and continuous activities” of a large corporation selling shoes in the forum state to the contacts of beneficiaries, many of them unknown, to a common trust fund. The Court would need more sophisticated jurisdictional tools to give a fuller explanation of personal jurisdiction over absent claimants or beneficiaries, tools that would develop in tandem with the growth of class and other mass actions, and standards that have resurfaced as difficult and contested in the Court’s newest round of personal jurisdiction cases.<sup>122</sup>

*International Shoe* and *Mullane* were not only the launching pad for the modern era of personal jurisdiction and notice; they form an inflection point in the parallel development of the two doctrines. Despite the evident relationship of personal jurisdiction and notice in *Hess* and *Milliken*, the *International Shoe* Court took care to segregate its discussion of notice from that of personal jurisdiction. And *Mullane* took no note of *International Shoe* at all, despite the shared doctrinal history and the presence of a personal jurisdiction issue in that case. Having used notice to gently prod personal jurisdiction toward a place where courts could consider issues of fairness and convenience, the Court used *International Shoe* and *Mullane* as the occasion to begin the process of breaking apart the shared space of these two doctrines.

---

119. *Int’l Shoe Co. v. Washington*, 326 U.S. 310, 320 (1945).

120. *Id.* at 321 (“The activities which establish its ‘presence’ subject it alike to taxation by the state and to suit to recover the tax.”).

121. *See supra* notes 103-104 and accompanying text.

122. *See Bristol-Myers Squibb Co. v. Superior Court of California*, 137 S. Ct. 1773 (2017); *infra* at notes 246-254 and accompanying text.

III.  
THE SHIFTING AND FADING ROLE OF NOTICE IN  
PERSONAL JURISDICTION FROM  
*INTERNATIONAL SHOE* AND  
*MULLANE THROUGH ASAHI*

In the decades after *International Shoe*, the Supreme Court embarked on its long (and still unfinished) project of defining the constitutional limits on personal jurisdiction, a time marked by bursts of judicial activity followed by long periods of silence on the matter.<sup>123</sup> The minimum contacts test occupied the central, but certainly not the only, space in the due process analysis. Along the way, various theories rose and fell in prominence: the power and sovereignty theories;<sup>124</sup> the respective interests of the defendant (in particular, the convenience interests of the defendant),<sup>125</sup> the plaintiff,<sup>126</sup> and the forum state;<sup>127</sup> and general notions of fairness,<sup>128</sup> the question of whether any or all of these considerations form a part of the minimum contacts test or exist outside of it as an additional constitutional check.<sup>129</sup> Many of these ideas had already begun to gain traction in the pre-*International Shoe* era, and some of them even predated *Pennoyer*.

Today, the Supreme Court continues to grapple with all these theories and clearly favors some more than others. But even the disfavored theories and considerations are still part of the jurisdic-

---

123. The longest period of inactivity was the more than twenty-year gap between the *Asahi* and *Burnham* decisions of the late 1980s and early 1990s, and the renewed interest in personal jurisdiction kicked off by *J. McIntyre* and *Goodyear* in 2011. See Bradt & Rave, *supra* note 73, at 1272.

124. See, e.g., *Michigan Trust Co. v. Ferry*, 228 U.S. 346, 353 (1913) (“Ordinarily jurisdiction over a person is based on the power of the sovereign asserting it to seize that person and imprison him to await the sovereign’s pleasure.”).

125. See, Arthur Taylor von Mehran, *Adjudicatory Jurisdiction: General Theories Compared and Evaluated*, 63 B.U.L. REV. 279, 306-10 (1983).

126. See, e.g., *Gulf Oil Corp. v. Gilbert*, 330 U.S. 501 (1947).

127. Int’l Commercial Dispute Comm. Assn of the Bar of N.Y.C., *Lack of Jurisdiction and Forum Non Conveniens as Defenses to the Enforcement of Foreign Arbitral Awards*, 15 AM. REV. INT’L ARB. 407, 427 n.78 (2004) (“Convenience of the forum may be, in some instances, a factor in determining whether the assertion of jurisdiction comports with due process. See, e.g., *Asahi Metal Industry Co. v. Superior Court*, 480 U.S. 102 (1987).”).

128. von Mehran, *supra* note 125.

129. See, e.g., *Simpson v. Quality Oil Co.*, 723 F. Supp. 382, 388 (S.D. Ind. 1989) (“I believe that the question of ‘relatedness’ must ultimately turn upon a consideration of constitutional due process, and that the Constitution limits ‘relatedness’ to substantive relevance. Although ‘relatedness’ can be initially defined by state statute (just as ‘minimum contacts’ are now defined in state long-arm statutes), the Constitution is the final check on these state statutes.”).

tional discourse. Somehow, only notice has slipped away. This Section traces the role of notice in personal jurisdiction analysis in the post-*International Shoe* and *Mullane* era.

These two doctrines—personal jurisdiction and notice—had always been intertwined, joined by the mechanical device of service of process and a common due process constitutional ancestor in *Pennoyer*, and *Pennoyer*'s conceptual forerunners. But things began to change after *International Shoe* and *Mullane*. Courts encountering personal jurisdiction questions had once turned routinely to the fact and concept of notice, or lack thereof, to supplement their reasoning behind the grant or denial of personal jurisdiction. In the decades after *International Shoe* and *Mullane*, notice was still used as a doctrinal tool for jurisdictional innovation, but its role evolved and ultimately faded.

As for the due process analysis related to notice itself, once the Supreme Court set the constitutional floor for the sufficiency of notice at the very liberal *Mullane* level, due process challenges to the mechanics of service of process were rare,<sup>130</sup> although the advent of new technology and social media has generated some new questions about the constitutionality of service via electronic means.<sup>131</sup>

---

130. The Court has occasionally heard cases about service and actual notice. See *U.S. Aid Funds Inc. v. Espinosa*, 559 U.S. 260 (2010) (defect in service was not a basis upon which to void a bankruptcy court's judgment because the party received actual notice of the debtor's plan and failed to object); *Jones v. Flowers*, 547 U.S. 220, 220 (2006) (while actual notice was not required, the State was required to take additional steps when notice by certified mail returned unclaimed); *Dusenbery v. United States*, 534 U.S. 161, 172–73 (2002) (actual notice to prisoner in a forfeiture proceeding not required when process sent by certified mail); *Mennonite Bd. of Missions v. Adams*, 462 U.S. 791, 791 (1983) (publication, posting, and mailed notice to the property owner are insufficient means of informing a mortgagee of a tax sale); *Robinson v. Hanrahan*, 409 U.S. 38 (1972) (notice by mail sent to home address of property owner insufficient in forfeiture proceeding where State knew property owner was in jail); *Walker v. City of Hutchinson*, 352 U.S. 112 (1956) (notice by newspaper publication alone insufficient in condemnation proceeding where city knew property owner's name). These occasional forays back into the due process requirements of service pale in comparison to the number and detail of personal jurisdiction cases that the Court has heard since *International Shoe*. Instead these cases center primarily around statutory, rule, and treaty interpretation.

131. See, e.g., *Rio Props. v. Rio Int'l Interlink*, 284 F.3d 1007, 1017 (9th Cir. 2002) (authorizing service of process by email); MEMO ENDORSED ORDER granting 419 Motion to Serve Wikileaks by Twitter & Mail on 148 NOTICE of Motion. Democratic National Committee v. Russian Fed'n, No. 18-cv-3501-JGK) (S.D.N.Y. June 21, 2018), (granting motion to serve Wikileaks by Twitter); *FTC v. PCCare247 Inc.*, 2013 U.S. Dist. LEXIS 31969, at \*16–17 (S.D.N.Y. Mar. 7, 2013) (authorizing service by email and Facebook); *Qaza v. Alshalabi*, 43 N.Y.S.3d 713, 716 (Sup. Ct. 2016) (“[P]laintiff has not demonstrated that, under the facts presented

The locus of due process notice doctrine shifted to questions of which proceedings required notice at all<sup>132</sup> and what sort of a proceeding would satisfy the requirement of the “opportunity to be heard.”<sup>133</sup>

A. *Early Jurisdictional Expansion in Perkins and McGee: Continued Use of Notice as a Fairness and Due Process Crutch*

In the first years after *International Shoe*, the Court issued two expansive personal jurisdiction decisions. In each case, the Court’s use of notice echoed the reasoning in key pre-*International Shoe* decisions. The fact of actual notice allowed the Court to create a cushion of fairness and a reassurance of due process that aided jurisdictional innovation.

The Court’s first major move after *International Shoe* was to lay a capacious foundation for the exercise of general jurisdiction. In *Perkins v. Benguet Consolidated Mining Company*<sup>134</sup> the Court had to jus-

---

here, service by Facebook is reasonably calculated to apprise defendant of the matrimonial action.”); *St. Francis Assisi v. Kuwait Fin. House*, No. 17-cv-07203-PJH, 2016 U.S. Dist. LEXIS 136152 (N.D.Cal.) (authorizing service by Twitter to Kuwaiti national when plaintiff unable to determine his location).

132. *See, e.g., Kaley v. United States*, 571 U.S. 320 (2014) (notice not required for pre-trial restraining orders to preserve potentially forfeitable assets in criminal proceedings); *Hamdi v. Rumsfeld*, 542 U.S. 507 (2004) (“[A] citizen-detainee seeking to challenge his classification as an enemy combatant must receive notice of the factual basis for his classification, and a fair opportunity to rebut the Government’s factual assertions before a neutral decisionmaker.”); *United States v. James Daniel Good Real Prop.*, 510 U.S. 43 (1993) (“[T]he seizure of real property under § 881(a)(7) is not one of those extraordinary instances that justify the postponement of notice and hearing. Unless exigent circumstances are present, the Due Process Clause requires the Government to afford notice and a meaningful opportunity to be heard before seizing real property subject to civil forfeiture.”); *Mathews v. Eldridge*, 424 U.S. 319 (1976) (“[A]n evidentiary hearing [and, therefore, notice] is not required prior to the termination of [Social Security] disability benefits.”).

133. *See, e.g., Connecticut v. Doehr*, 501 U.S. 1 (1991) (procedure that allowed prejudgment attachment of real property without notice, hearing, or showing of extraordinary circumstances violates due process); *N. Ga. Finishing, Inc. v. Di-Chem, Inc.*, 419 U.S. 601 (1975) (finding Georgia procedure that allowed plaintiffs to secure a garnishment from a court clerk without involvement of a judge or an early hearing unconstitutional under the Fourteenth Amendment); *Fuentes v. Shevin*, 407 U.S. 67 (1972) (holding Florida and Pennsylvania prejudgment replevin provisions as violating due process “insofar as they deny the right to a prior opportunity to be heard before chattels are taken from their possessor.”); *Snidach v. Family Fin. Corp. of Bay View*, 395 U.S. 337, 339–40 (1969) (“[A]bsent notice and a prior hearing, this prejudgment garnishment procedure violates the fundamental principles of due process.”) (internal citations omitted).

134. 342 U.S. 437 (1952).

tify Ohio's exercise of personal jurisdiction over a Filipino corporation for a lawsuit that was unrelated to any of its activities in Ohio.<sup>135</sup> The Court held that Benguet's Ohio activities were sufficiently substantial in nature so as to constitute the sort of minimum contacts that could serve as the fictive "presence" contemplated in *International Shoe*.<sup>136</sup> *Perkins* was such a large factual and doctrinal leap that it has been largely criticized by later jurists and commentators, and courts have narrowed its holding by limiting the decision to its somewhat unique facts.<sup>137</sup> But the decision remains relevant for our story because of the prominence of notice in the Court's analysis.

Like *International Shoe* and *Milliken* before it, the *Perkins* Court used the language of fairness to justify its decision. Echoing these earlier cases, the Court folded discussions of notice and service of process into its appeal to principles of fairness. The Court even characterized the question presented as "whether the state courts of Ohio are open to a proceeding *in personam* against an *amply notified* foreign corporation."<sup>138</sup> While much of the fairness analysis concerned Benguet's contacts with Ohio,<sup>139</sup> the Court found "no unfairness" where the corporation was carrying on such "activities appropriate to *accepting service* or *receiving notice* on [the corporation's] behalf."<sup>140</sup> Moreover, the Court criticized the Ohio Supreme

---

135. *Id.* at 438 (holding that a Philippines corporation could be sued *in personam* in Ohio for suit that "did not arise in Ohio and [did] not relate to the corporation's activities there.").

136. *Id.* at 447–48.

137. See *BNSF Ry. v. Tyrrell*, 137 S. Ct. 1549, 1561–62 (2017) (Sotomayor, J., concurring in judgment); *L. D. Reeder Contractors v. Higgins Indus., Inc.*, 265 F.2d 768, 775 (9th Cir. 1959) ("We realize that [*Perkins*] is authority for the theory that the cause of action need not arise out of the activity of the nonresident within the forum state. But this was an earlier case than either *McGee* or *Hanson*, and rests upon its own peculiar facts."). Note, however, that *Perkins* is commonly invoked by commentators and even courts seeing to promote jurisdictional expansion. See, e.g., Taylor Simpson-Wood, *In the Aftermath of Goodyear Dunlop: Oyez! Oyez! Oyez! A Call for a Hybrid Approach to Personal Jurisdiction in International Products Liability Controversies*, 64 BAYLOR L. REV. 113, 156 (2012) ("[I]t can be hoped that somewhere in the future, a majority of the Court will, at a minimum, recognize a more expansive view of general jurisdiction than that of *Perkins*.").

138. 342 U.S. at 440 (emphasis added).

139. *Id.* at 447–48.

140. *Id.* at 445 (emphasis added). The Court also explicitly rejected the idea that the defendant could make a specific due process challenge based on a lack of notice, noting that "[a]ctual notice of the proceeding was given to the corporation . . . through regular summons upon its president while he was in Ohio acting in that capacity. Accordingly, there can be no jurisdictional objection based on a lack of notice to a responsible representative of the corporation." *Id.* at 439–40.

Court for relying on two older U.S. Supreme Court cases that, aside from predating *International Shoe*, were instances in which “no actual notice of the proceedings was received” by the foreign corporation or its representative.<sup>141</sup> Notice was a due process link that provided continuity between pre- and post-*International Shoe* cases. Where notice once bolstered expansive claims about a corporation’s literal presence in the forum state, it could now be deployed to rationalize the fairness of corporate contacts with the forum state.

The Supreme Court also addressed specific jurisdiction during this period in *McGee v. International Life Insurance Company*.<sup>142</sup> There, the Court held that a single contact with a forum state could be enough to support personal jurisdiction in a lawsuit arising from that contact.<sup>143</sup> The Court grounded its jurisdictional expansion by echoing *International Shoe*’s appeals to changes in the nationalizing economy<sup>144</sup> but also with appeals to fairness that loosely accounted for the interests and conveniences of the plaintiff, the defendant, and the forum state.<sup>145</sup> The Court punctuated its assessment of the fairness of exercising jurisdiction on the basis of a single contact by observing that “[t]here is no contention that respondent did not have *adequate notice* of the suit or sufficient time to prepare its defenses and appear.”<sup>146</sup>

There is a subtle difference in how the Court used notice in *McGee* compared with *Perkins*. The *Perkins* Court assessed the activities of the corporate defendant and drew a direct link between the fact that these activities were minimum contacts that mimicked the older need to establish territorial jurisdiction through presence in the forum, and the fact that some of these activities specifically enabled service and ensured notice. The *McGee* Court’s treatment of notice is more ambiguous. It is possible to read the Court’s sentence about notice as a cursory, pro forma statement that is ap-

---

141. *Id.* at 443–44.

142. 355 U.S. 220 (1957).

143. *Id.* at 221 (upholding a California long-arm statute that “subject[ed] foreign corporations to suit in California on insurance contracts with residents of that State.”). *See also* Rhodes, *supra* note 19, at 196.

144. *Id.* at 222 (focusing primarily on the insurer’s tight connections to the insured in California which it portrayed as a feature of the “fundamental transformation of our national economy over the years.”).

145. *Id.* at 224 (plaintiffs might be “at a severe disadvantage if they were forced to follow the insurance company to a distant State,” whereas the burden to the defendant might be an “inconvenience . . . but certainly nothing which amounts to a denial of due process.”).

146. *Id.* at 224 (emphasis added).

pendent to a discussion about personal jurisdiction, which is another topic entirely. The Court simply meant to dispense with any lingering questions that there might be *other* due process objections in the case *outside* of the jurisdictional challenge. On the other hand, the Court does little to signal that concerns about notice must be limited to separate and formal challenges to a due process notice deficiency. The observation about notice is part of the core paragraph in which the Court described due process limitations on personal jurisdiction. Given the history of connecting the fairness or even “natural justice” entitlements of notice and opportunity to be heard to personal jurisdiction, Justice Black might have assumed that readers would expect a nod to notice within personal jurisdiction even though litigants could make a separate challenge to a failure of notice. This would be no different than his emphasis on convenience despite the existence of separate judicial doctrines and remedies for inconveniently located adjudication.<sup>147</sup>

Curiously, the *McGee* sentence about notice is the only sentence in the due process paragraph that does not have a citation. Had the Court wanted to delineate notice as a separate due process doctrine or challenge to be made in a case like *McGee*, Justice Black might have punctuated this sentence with a cite such as “cf. *Mullane*,” reinforcing the idea that notice now had its own due process life aside and apart from personal jurisdiction. But the Court’s statement about notice is the only sentence in the paragraph unsupported by any authority. While it is possible that this was a deliberate attempt to obscure the relationship of personal jurisdiction and notice within due process, the more likely explanation is that it is indicative of the justices’ own muddled thinking about the two. As we have seen, the formal detachment of notice from personal jurisdiction in *Mullane* was not inevitable, nor was it entirely clear that a formal separation was what the Court meant to achieve in the *International Shoe* and *Mullane* sequence.<sup>148</sup>

The *McGee* Court did not center or emphasize notice to the same extent that the *Perkins* Court did, but the *McGee* opinion demonstrates some discomfort with the idea of jurisdictional innovation unsupported by an affirmative showing of notice and opportunity to be heard. *Perkins* and *McGee* together suggest that, although *International Shoe* and *Mullane* signaled a shift toward a new framework for evaluating the due process merits of personal jurisdiction and

---

147. See Maggie Gardner, *Retiring Forum Non Conveniens*, 92 N.Y.U. L. REV. 390, 429 (2017) (suggesting that forum non conveniens is redundant with several existing doctrines including personal jurisdiction).

148. See *supra* note 122 and accompanying text.

notice, the Court still considered notice to be a relevant touchstone of due process as it applied to personal jurisdiction.

*B. From Notice of Suit to Notice of Jurisdiction: Transforming Notice from a Tool of Jurisdictional Expansion to a Tool of Due Process Expansion*

Just one year after *McGee*, the Supreme Court began a slow shift in its use, and ultimately its disregard, of notice in personal jurisdiction cases. In the following decades, the Supreme Court took two concepts of notice, notice of suit and notice of jurisdiction, and separated them for doctrinal and analytic purposes. Prior to *International Shoe*, these concepts were merged because of the strict requirements of territoriality and in-hand service. But after the Court replaced the *Pennoyer* regime with minimum contacts in *International Shoe* and relaxed the constitutional requirements for notice in *Mullane*, notice of jurisdiction emerged as a problem. If a mélange of activities could subject an out-of-state defendant to jurisdiction in the forum state, and that defendant could be served by substituted service, how could these defendants assure themselves that they would or would not be subject to the jurisdiction of the forum state? And did such notice of jurisdiction matter? The answer that the Court would consistently give over the next half century was yes. Notice of jurisdiction is a constitutionally relevant concept, and the farther that personal jurisdiction analysis pulled away from the more traditional Fourteenth Amendment concept of notice of suit, the tighter the Court would cling to the importance of notice of jurisdiction. This shift was neither doctrinally consistent nor unproblematic, as the exploration of the following cases will show.

The Court first signaled this shift in *Hanson v. Denckla*,<sup>149</sup> in which it held that a Florida court could not exercise personal jurisdiction over a Delaware bank.<sup>150</sup> The Court took care to sever the issue of notice from personal jurisdiction, beginning its analysis by setting the issue of notice aside entirely, noting that “[t]here is no suggestion that the [Florida] court failed to employ a means of notice reasonably calculated to inform nonresident defendants of the pending proceedings, or denied them an opportunity to be heard in defense of their interests.”<sup>151</sup> By fronting the issue in a curt and

---

149. 357 U.S. 235 (1958).

150. *Id.* at 253–54 (holding that the Florida court did not have jurisdiction over the Delaware bank because the only contacts that the bank had with Florida were due to the unilateral actions of a third party (the settlor of the trust)).

151. *Id.* at 245.

cursory manner, Chief Justice Warren severed it from the analysis that would affect its view of jurisdiction. Contrast this with the approach from *McGee* where the observation about notice at the end of the personal jurisdiction discussion tied notice to the Court's holding and offered a reassurance of due process fairness and reasonability. Here, the up-front and summary dismissal of the notice issue signaled the Court's desire to turn to the "real" business of personal jurisdiction. This difference in approach from *Perkins* and *McGee* was not necessarily a new phenomenon. As I documented in Part I, courts in the pre-*International Shoe* era often treated notice as integral to personal jurisdiction analysis, but sometimes analyzed or mentioned it as a separate doctrine.<sup>152</sup> In this regard, the Court's rhetorical choice is not surprising. Since the Court held that the Florida court did not have personal jurisdiction, the Court's jurisdictional innovations in *Hanson* were not expansionist. The Court did not need to highlight the fact of notice to underscore fairness to the defendant or a solid foundation of due process. Instead, it needed to segregate notice as a due process issue which, when satisfied, did not preclude a separate finding of a personal jurisdiction due process violation.

Chief Justice Warren's citations confirm this rhetorical stroke. He cited *Mullane* along with two other cases to support his assertion that notice met due process standards, whereas Justice Black's *McGee* made no such citation.<sup>153</sup> Justice Black, for his part, dissented in *Hanson*, noting in his argument in favor of the Florida court's jurisdiction that the bank "had timely notice of the suit and an adequate opportunity to . . . appear." This statement had some of the same ambiguity that I noted in *McGee*; it is unclear whether this statement was tied to his subsequent discussion of fairness, convenience, and litigant and forum state interest, or if it was a prefatory comment meant to clear the way for the "real" personal jurisdiction analysis.<sup>154</sup> Thus the majority opinion and arguably the dissent in *Hanson* demonstrated the Court's movement towards a firmer delineation of due process notice from due process personal jurisdiction.

*Hanson* also contained the seeds of another transformation—the shift from emphasizing notice as a consideration linked to service of process and notification of the actual proceeding, to notice

---

152. See *supra* Part I.C. Sometimes notice was invoked to support a jurisdictional holding, sometimes it was treated as a separate issue, sometimes it was not mentioned at all.

153. 357 U.S. at 258 (Black, J., dissenting).

154. Interestingly, Justice Black cites *Mullane* here, but for its personal jurisdiction holding, and not for notice. *Id.* at 260–61.

as a more abstract idea of foreseeability. *Hanson* famously introduced the idea that unilateral activity by the plaintiff could not create contacts in the forum state which would be imputed to the defendant.<sup>155</sup> Justice Warren's discussion of unilateral activity in *Hanson* was not tied to notice or foreseeability, but to the idea that the defendant had not "purposefully availed" itself of the forum.<sup>156</sup> In the coming years, the Court would combine the new idea of purposeful availment with the older idea of notice to create a new and controversial factor in personal jurisdiction analysis: foreseeability, which would toggle between foreseeability of effect or harm, and foreseeability of jurisdiction.

An early case to presage this move came not from the U.S. Supreme Court, but from the Illinois Supreme Court in *Gray v. American Radiator & Sanitary Corp.*<sup>157</sup> The opinion followed in the *Hanson* model of formalizing different due process "tracks" for personal jurisdiction and notice. It delineated two questions, "first, whether [the defendant] has certain minimum contacts with the State . . . and second, whether there has been a reasonable method of notification,"<sup>158</sup> citing *International Shoe* separately for each of those propositions. The court made the perfunctory observation that the Illinois long-arm statute made adequate provisions for notice and that the plaintiff followed these provisions. Thus, the "real" work to be done was in assessing whether the defendant had minimum contacts with the state to justify the exercise of jurisdiction.<sup>159</sup>

Justice Klingbiel reiterated the division of the two due process doctrines by opining that "the trend in defining due process of law is away from the emphasis on territorial limitations and toward emphasis on providing adequate notice and opportunity to be heard: from the court with immediate power over the defendant, toward the court in which both parties can most conveniently settle their dispute."<sup>160</sup> In this view, the trajectory of due process jurisprudence was one in which rigid rules of territoriality and forms of service had given way to standards of reasonableness or fairness for both

---

155. *Hanson*, 357 U.S. at 253.

156. *Id.*

157. 176 N.E.2d 761 (Ill. 1961) (Illinois State Supreme Court case that entered the modern personal jurisdiction canon because of its early and formative treatment of the idea that a product manufactured or sold outside of the forum state and put in the "stream of commerce" could serve as a sufficient minimum contact).

158. *Id.* at 763.

159. *Id.*

160. *Id.* (citing *Smyth v. Twin State Improvement Corp.*, 80 A.2d 664 (Vt. 1951)).

personal jurisdiction and notice, and these standards operated independently and should be separately evaluated. Whereas earlier courts, such as the *Hess* and *Milliken* courts, had grounded the fairness of expanding jurisdiction at least in part on the fact of reasonable notice, courts pioneering the *Gray* model treated notice as a necessary co-requisite for due process, and emphasized how a manufacturer of a product sold in the forum has purposefully availed itself of that state. While Justice Klingbiel stopped short of using the term “foreseeable,” the stream-of-commerce groundwork had been laid and the due process division of notice and personal jurisdiction was further entrenched.

The Supreme Court echoed this bifurcated structure in *World-Wide Volkswagen v. Woodson*,<sup>161</sup> with its holding that “due process requires that the defendant be given adequate notice of the suit, and be subject to the personal jurisdiction of the court.” The Court quickly set aside notice because it was “not contended that notice was inadequate; the only question is whether these particular petitioners were subject to the jurisdiction of the Oklahoma courts.”<sup>162</sup> But where the Illinois Supreme Court cited *International Shoe* for both notice and personal jurisdiction, Justice White cited *Mullane* for notice and *International Shoe* for personal jurisdiction, further cementing the separate due process tracks for each. *World-Wide Volkswagen* is the Court’s first major introduction of foreseeability into the mélange of personal jurisdiction due process factors. The Court took care to distinguish two types of foreseeability. “[T]he foreseeability that is *critical to due process analysis* is not the mere likelihood that a product will find its way into the forum State. Rather, it is that the defendant’s conduct and connection with the forum State are such that he should reasonably anticipate being haled into court there.”<sup>163</sup>

---

161. 444 U.S. 286 (1980). The Court had also used the bifurcated structure two years earlier in *Kulko v. Superior Court of California*, 436 U.S. 84, 91 (1978) (“the presence of reasonable notice to the defendant that an action has been brought” and minimum contacts that ensure that it would be “fair to require defense of the action in the forum.”). The balance of *Kulko* examined personal jurisdiction from several angles: the nature of the contacts that the defendant had with California, fairness, reasonableness, the relevance of the interests of the plaintiff and defendant, and the interests of the forum. While the Court discussed whether the defendant could foresee that his actions would have consequences in the forum state, it notably omitted a discussion of whether the defendant could have foreseen jurisdiction itself.

162. *Id.* at 291.

163. *Id.* (emphasis added).

The real innovation in *World-Wide Volkswagen* was to harness the older doctrines of notice and purposeful availment to create the malleable and controversial personal jurisdiction factor of foreseeability of jurisdiction. Instead of pointing to the adequacy of notice of the pendency of the action at hand, the Court addressed a different sort of notice, notice of jurisdiction:

When a corporation ‘purposefully avails itself of the privilege of conducting activities within the forum State,’ it has clear notice that it is subject to suit there, and can act to alleviate the risk of burdensome litigation by procuring insurance, passing the expected costs on to customers, or, if the risks are too great, severing its connection with the State.<sup>164</sup>

The Court did not deploy the word “notice” here by accident. It reflects the longer tradition, outlined earlier, of using notice as a factor in the due process analysis of personal jurisdiction and that notice of jurisdiction had always been an implicit part of notice of suit because of the doctrinal and historical tie to personal jurisdiction via due process.

*World-Wide Volkswagen’s* introduction of notice of jurisdiction has been roundly (and rightly) criticized as an amorphous and circular standard.<sup>165</sup> Its status among the other personal jurisdiction factors is still unsettled as a matter of Supreme Court jurisprudence. Notice of jurisdiction also featured prominently in the later “stream of commerce” cases. Several years after *World-Wide Volkswagen*, the Supreme Court issued its split plurality decision in *Asahi*

---

164. *Id.* at 297 (citing *Hanson v. Denckla*, 357 U.S. 235, 253 (1958)).

165. *Id.* at 311 n.18 (Brennan, J. dissenting). See Borchers, *supra* note 8 at 94 (“the ‘jurisdictional surprise’ argument is circular. Any expectation that the defendant has of avoiding an out-of-state court is a function of the jurisdictional rules themselves”); Lea Brilmayer & Matthew Smith, *The (Theoretical) Future of Personal Jurisdiction: Issues Left Open By Goodyear Dunlop Tires v. Brown and J. McIntyre Machinery v. Nicastro*, 63 S.C. L. REV. 617, 623 (2012) (“[T]he problem with expectations is that it has too little semantic content and uniformly results in circularity.”); Allan Erbsen, *Personal Jurisdiction Based on the Local Effects of Intentional Misconduct*, 57 WM. & MARY L. REV. 385, 430 (2015) (“the anticipation test is circular”); Martin H. Redish, *Due Process, Federalism, and Personal Jurisdiction: A Theoretical Evaluation*, 75 NW. U. L. REV. 1112, 1134 (1981) (“in a sense, such an argument is circular, since a potential defendant can only have such an expectation because the law so provides.”); Stein, *supra* note 8 at 701 (“expectation is defined largely by the courts’ ruling on the subject” and “is therefore circular”). But see Linda Sandstrom Simard, *Hybrid Personal Jurisdiction: It’s Not General Jurisdiction, or Specific Jurisdiction, but Is It Constitutional?*, 48 CASE W. RES. L. REV. 559, 589–90 (1998) (addressing the circularity problem by distinguishing legal consequences and factual consequences of defendants’ actions).

*Metal Industries v. Superior Court*.<sup>166</sup> The question in *Asahi* was whether a California court could assert personal jurisdiction over a Japanese manufacturer of tire valves that it sold to a Taiwanese tire manufacturer, who in turn sold its tires world-wide. The plaintiff sued both manufacturers over a motorcycle accident in California, alleging defects in the tire and valves.<sup>167</sup> The Justices were unanimous in the judgment: California did not have personal jurisdiction over *Asahi*, but they split sharply over the reasoning. Justice Brennan argued that *Asahi* did have minimum contacts with California, but that it would be unreasonable as a matter of due process to exercise jurisdiction over the company.<sup>168</sup> Justice O'Connor led a four-justice plurality and argued that *Asahi* did not have the requisite minimum contacts with California.<sup>169</sup> These two positions have come to be known as the “stream of commerce” and “stream of commerce plus” theories, respectively. Although the Court has, as of late, indicated a preference for the “stream of commerce plus” theory,<sup>170</sup> the question of which test applies remains open.<sup>171</sup> Notable for our purposes is the extent to which the justices agreed on the role of foreseeability of jurisdiction.

Justice Brennan’s stream of commerce theory rested, in part, on his belief that litigation in California would be foreseeable to a manufacturer whose products arrive in a forum state as part of the

---

166. 480 U.S. 102 (1987).

167. *Id.* at 107–08.

168. *Id.* at 119–20.

169. *Id.* at 116. For extensive commentary on the *Asahi* opinion and stream of commerce, see Todd David Peterson, *The Timing of Minimum Contacts After Goodyear and McIntyre*, 80 GEO. WASH. L. REV. 202, 207–28 (2011); Allan Ides & Simona Grossi, *The Purposeful Availment Trap*, 7 FED. CTS. L. REV. 113, 118–19 (2014); Arthur R. Miller, *Simplified Pleading, Meaningful Days in Court, and Trials on the Merits: Reflections on the Deformation of Federal Procedure*, 88 N.Y.U. L. REV. 286, 348–49 (2013); Alan M. Trammell, *A Tale of Two Jurisdictions*, 68 VAND. L. REV. 501, 522 (2015).

170. See *J. McIntyre Machinery Ltd. v. Nicastro*, 564 U.S. 873, 883 (2011).

171. Roger W. Hughes, *Personal Jurisdiction: Selected Current Issues*, 80 THE ADVOC. (TEX.) 63, 67 (2017) (discussing the “unresolved stream-of-commerce dispute” and different jurisdictions’ varying approaches); Patrick J. Borchers, *How “International” Should a Third Conflicts Restatement be in Tort and Contract?*, 27 DUKE J. COMP. & INT’L L. 461, 467 (2017) (“After *Asahi*, predictable confusion reigned among lower courts as to which version of the stream-of-commerce test to apply - Justice O’Connor’s or Justice Brennan’s. Courts divided as to which to follow, and others hedged their bets by concluding that the result would be the same under either test.”); Howard M. Wasserman, *The Roberts Court and the Civil Procedure Revival*, 31 REV. LITIG. 313, 325 (2012) (“[A]fter waiting twenty years to hear a personal jurisdiction case and taking a case that squarely presented the stream-stream-plus divide, the Court still left the issue unresolved.”).

“regular and anticipated flow of products.”<sup>172</sup> In these circumstances, “[a]s long as a participant in this process is aware that the final product is being marketed in the forum State, the *possibility of a lawsuit* there cannot come as a surprise.”<sup>173</sup> To support this position, Brennan quoted extensively from *World-Wide Volkswagen* and *Gray*. For Brennan, the foreseeability of the results of the stream of commerce were sufficient to put the defendant on notice that it might be sued in the forum. Brennan viewed foreseeability of lawsuits in the forum state as a due process matter, a position he had already stated in *Burger King v. Rudzewicz*.<sup>174</sup> But while Brennan’s *Asahi* opinion stressed that this sort of foreseeability can count as a minimum contact per *International Shoe*, he was adamant that other due process considerations such as convenience, burdens to the defendant, and the relative interests of the plaintiff and the forum state can override minimum contacts and provide an alternative due process barrier to the exercise of jurisdiction.<sup>175</sup>

Justice O’Connor disagreed with Brennan, contending that products placed in the stream of commerce cannot alone constitute a minimum contact, in part, because the stream of commerce itself is insufficient as a mechanism for predicting the possibility of a lawsuit in the forum. The “stream of commerce plus” approach requires that the defendant engage in other activities that demonstrate that it was targeting the state. The emphasis on purposeful availment goes far beyond notice of potential lawsuits. Purposeful engagement with the forum meshes well with notions of consent, reciprocity, and even territoriality and sovereignty.<sup>176</sup> But notice undoubtedly motivated O’Connor’s formulation of stream of commerce plus. She favorably cites the *World-Wide Volkswagen* and *Hanson* notice of jurisdiction language, that purposeful availment is

---

172. 480 U.S. at 117.

173. *Id.* (emphasis added).

174. *Burger King Corp. v. Rudzewicz*, 471 U.S. 462, 472 (1985) (“By requiring that individuals have fair warning that a particular activity may subject (them) to the jurisdiction of a foreign sovereign, the Due Process Clause gives a degree of predictability to the legal system that allows potential defendants to structure their primary conduct with some minimum assurance as to where that conduct will and will not render them liable to suit.”).

175. *Asahi*, 480 U.S. at 116-21 (Brennan, J., concurring in part).

176. See William S. Dodge & Scott Dodson, *Personal Jurisdiction and Aliens*, 116 MICH. L. REV. 1205 (2018); Robin J. Effron, *Solving the Nonresident Alien Due Process Paradox in Personal Jurisdiction*, 116 MICH. L. REV. ONLINE 123 (2018); Austen L. Parrish, *Sovereignty, Not Due Process: Personal Jurisdiction Over Nonresident Alien Defendants*, 41 WAKE FOREST L. REV. 1 (2006).

evidence that the defendant “has clear notice that it is subject to suit there.”<sup>177</sup>

In other words, Brennan and O’Connor disagreed on a number of issues, including whether placing an article in the stream of commerce is sufficient to constructively notify the defendant of possible jurisdiction. But what is clear is that by the time *Asahi* was decided, all the Justices were operating under the assumption that one of the functions of minimum contacts was to provide constructive notice of jurisdiction to defendants. In other words, if minimum contacts were to serve as a workable framework, it had to do more than act as a modern substitute for physical presence in the forum state. It also had to provide the notice of jurisdiction that had always been implicit in the pre-*International Shoe* understanding of service of process and personal jurisdiction. The Justices simply disagreed about what sort of conduct could be reasonably understood to provide such advanced warning. As we shall see, this assumption would not last. By the time the Court heard *J. McIntyre v. Nicastro*, Justice Kennedy would opine that “foreseeability is inconsistent with the premises of lawful judicial power. This Court’s precedents make clear that it is the defendant’s actions, *not his expectations*, that empower a State’s courts to subject him to judgment.”<sup>178</sup>

A few years after *World-Wide Volkswagen*, the Supreme Court used notice of jurisdiction to support personal jurisdiction over defendants in two intentional tort cases, *Calder v. Jones*<sup>179</sup> and *Keeton v. Hustler*.<sup>180</sup> In *Calder*, the writer and the editor of the *National Enquirer*, both Florida residents, had only sporadic contacts with the state of California where the plaintiff filed her lawsuit. However, the Court upheld personal jurisdiction on the basis that the defendants’ “intentional, and allegedly tortious, actions were expressly aimed at California.”<sup>181</sup> Such purposeful conduct, combined with the defendants’ knowledge that the plaintiff lived in California where the paper had a large circulation, led the Court to conclude that the defendants “must ‘reasonably anticipate being haled into court [in California]’ to answer for the truth of the statements

---

177. *Asahi*, 480 U.S. at 110 (citing *Hanson v. Denckla*, 357 U.S. 235, 297 (1958)).

178. *J. McIntyre Mach., Ltd. v. Nicastro*, 564 U.S. 873, 883 (2011).

179. 465 U.S. 783 (1984).

180. 465 U.S. 770 (1984).

181. *Calder*, 465 U.S. at 789.

made in their article.”<sup>182</sup> In *Keeton*, the defendant was the magazine itself. Although the plaintiff had no connection to the forum state, the Court held that when the defendant “has continuously and deliberately exploited the New Hampshire market, it must reasonably anticipate being haled into court there in a libel action based on the contents of its magazine.”<sup>183</sup> Neither *Calder* nor *Keeton* mentioned notice of suit specifically, and the foreseeability point is more of a perfunctory observation rather than the key compelling factor driving the Court toward jurisdictional expansion. Comparing this to the more detailed (if also more circular and confusing) arguments in *World-Wide Volkswagen* and other stream of commerce cases, it appears that the Court was more interested in notice of jurisdiction discourse as a method of pressing forward due process expansion rather than jurisdictional expansion.

During the era of *World-Wide Volkswagen* and *Asahi*, the concept of notice of jurisdiction exerted a powerful influence over the due process analysis of personal jurisdiction. These cases, with their varying deployment of “foreseeability” show a refashioned attentiveness to the due process consequences of notice; notice of jurisdiction had become a “New Notice” doctrine for the post-*International Shoe* and *Mullane* era. By the time of *World-Wide Volkswagen* and subsequent cases, notice of suit was relatively easy to establish. Particularly with the advent of a more connected society and better modes of communications, giving actual, or even constructive, notice of a lawsuit just was not that difficult.

The relative ease of executing notice of suit provided the Court with a soothing due process cushion from which it expanded the reach of constitutionally permissible personal jurisdiction in the first half of the twentieth century. Halting the growth of personal jurisdiction, however, required an expansion of due process. Notice, in the form of notice of suit, had served as a useful due process hook throughout the early growth of constitutional personal jurisdiction doctrine. It was only natural that the Court would again rely on notice, this time in the form of notice of jurisdiction, as a tool for doctrinal expansion, but this time, expanding due process as a means of restricting personal jurisdiction.

This “new notice” was complex; a vague, circular standard with ever-changing goal posts of which contacts would suffice as “purposeful availment.” It provided the Court with a useful barrier to

---

182. *Id.* at 790 (quoting *World-Wide Volkswagen Corp. v. Woodson*, 444 U.S. 286, 297 (1980)).

183. *Keeton*, 465 U.S. at 781 (citing *World-Wide Volkswagen Corp.*, 444 U.S. at 297–98).

the exercise or expansion of state jurisdictional power. By linking purposeful availment and notice, the Court could maintain the connection between personal jurisdiction and an intuitive due process value—the personal liberty interest in receiving notice of a lawsuit. Shifting the locus of notice from the pendency of an extant suit to notice of jurisdiction ensured continuity of due process values in a doctrine that began its life on shaky due process grounds and was becoming increasingly distant from due process.

Notice of jurisdiction also differed significantly from notice of suit in that it made passive what had once been active. Notice of suit involved concrete delivery methods that the forum state and the plaintiff could together accomplish to perfect notice. The state was responsible for promulgating long-arm statutes with prescribed methods of personal or substituted service. These would be constitutional so long as they were designed to apprise the defendants of the pendency of the action. Plaintiffs were tasked with actually serving (or attempting to serve) the defendants pursuant to that statute. So long as the state rules were reasonably calculated under the circumstances to apprise defendants of the action and so long as the plaintiff followed those rules, courts were more or less comfortable with assuming that the due process requirements of notice under *Mullane* had been met. Generalized worries that defendants would not learn of lawsuits had long faded from the judicial horizon, punctuated only by isolated instances of problematic service of process on idiosyncratically situated parties.<sup>184</sup> In other words, notice of suit provided states and plaintiffs with relatively concrete rules and steps that could be taken to ensure that a lawsuit would not interfere with a defendant's due process liberty interest, and this supported the Court's earlier intuition that state jurisdictional power could be expanded beyond *Pennoyer's* rigid confines.

It is not surprising, then, that the due process tool the Court had once used to expand personal jurisdiction could be refash-

---

184. See, e.g., *Nelson v. Adams USA, Inc.*, 529 U.S. 460 (2000) (requiring notice when amending judgment to simultaneously add a third party defendant not previously named in lawsuit); *Volkswagenwerk A.G. v. Schlunk*, 486 U.S. 694 (1988) (reviewing the applicability of the Hague Service Convention to service of process on a foreign corporation's domestic subsidiary); *Eisen v. Carlisle & Jacqueline*, 417 U.S. 156 (1974) (considering notice requirements for absent class members in a Rule 23(b)(3) class action suit). But see Adrian Gottshall, *Solving Sewer Service: Fighting Fraud with Technology*, 70 ARK. L. REV. 813, 816 (2018) ("Unreliable and unfair service practices are a national problem."). As noted earlier, the bulk of constitutional notice doctrine shifted to the question of what sort of proceedings trigger a notice requirement at all, and what sort of hearing or process satisfies the constitutional requirement. See *supra* notes 132-133.

ioned into a limiting device. Notice of suit was, in a sense, too easy to satisfy. Since nearly every plaintiff with reasonable diligence and the right resources could meet the basic requirements, it no longer did enough “work” in the personal jurisdiction calculus. Problems with notice of suit no longer looked like jurisdictional issues, but as separate problems to be solved with reference to *Mullane* and perhaps interpretation of a state or federal statute.

But notice of jurisdiction was different. It clothed minimum contacts in familiar language that had a much more intuitive connection to a litigant’s personal liberty due process interest. And since notice of jurisdiction was conveniently circular, it was always up for redefinition by reference to other jurisdictional values, none of which had a particularly compelling connection to due process. This had consequences for both jurisdictional expansionists and jurisdictional contractionists. Jurisdictional expansionists advocated for a frank recognition that foreseeability was not about notice and advocated for its use as a broad, independently justified concept that concerned minimum contacts by the defendant with the forum state and the concomitant justification of the exercise of state power. Jurisdictional contractionists used notice of jurisdiction foreseeability as a bridge between old notice and purposeful availment to create the “stream of commerce plus” test.

C. *A New Use for Notice in a New World of  
In Rem Jurisdictional Problems*

*Shaffer v. Heitner*<sup>185</sup> was the first *in rem* jurisdictional case post-*International Shoe*. The Court used this occasion to hold that the Fourteenth Amendment applied to a much larger universe of cases than had been previously considered. *Shaffer* was a shareholder derivative action against Greyhound, a Delaware corporation, and several of its directors and officers. The plaintiff filed the lawsuit in Delaware where it seemed unlikely that the court would have *in personam* jurisdiction over 21 of the directors and officers. The plaintiff used Delaware’s sequestration statute to seize the defendants’ property—stock in the Greyhound corporation, the situs of which was Delaware, as a basis for *quasi in rem* jurisdiction.<sup>186</sup> The lower courts had been operating under the assumption that under *Pennoyer*, a forum state had power over all property within its borders, rendering a due process minimum contacts inquiry irrelevant

---

185. 433 U.S. 186 (1977).

186. *Id.* at 190–91.

to this exercise of jurisdiction.<sup>187</sup> The Supreme Court reversed, holding that the due process clause applies to exercises of *in rem* jurisdiction. Thus, post-*Shaffer*, out-of-state defendants must have the requisite minimum contacts with the forum state, regardless of whether the jurisdictional predicate is *in personam* or *in rem*.

Extending the minimum contacts test to the *in rem* jurisdiction cases required some significant justification by the Court, including a detailed effort to explain that the *Shaffer* decision would, in practice, only change the outcome in a small number of cases, primarily involving intangible property whose situs in a forum was dictated by statute.<sup>188</sup> Justice Marshall noted *Pennoyer's* focus on notice, but argued that this was clearly subordinate to the concerns for state power and sovereignty. While the Court could question the sufficiency of notice within a state's borders, and seizure of property was considered adequate in most cases,<sup>189</sup> due process in terms of notice of suit was simply irrelevant to *in rem* jurisdiction when the property owner was outside of the state.<sup>190</sup> Justice Marshall did not discuss the role of notice as a due process vehicle for advancing jurisdictional expansion between *Pennoyer* and *International Shoe*. Nevertheless, he used the concept of notice to expand doctrine, here the due process protection to the *in rems*, prefiguring what the Court would soon do with *World-Wide Volkswagen* and *Asahi*.

After summarizing the historical development of *in personam* jurisdiction into minimum contacts, he noted that “[n]o equally dramatic change has occurred in the law governing jurisdiction in rem.”<sup>191</sup> In building the bridge from *International Shoe* to the *in rem* cases, Marshall cited lower court and scholarly sources advancing arguments for applying minimum contacts. But his primary appeal to Supreme Court precedent was to note that “we have held that property cannot be subjected to a court’s judgment unless reasonable and appropriate efforts have been made to give the property owners *actual notice* of the action. . . . This conclusion recognizes, contrary to *Pennoyer*, that an adverse judgment *in rem* directly affects the property owner by divesting him of his rights in the property before the court.”<sup>192</sup> In other words, notice doctrine was one of the

---

187. *Id.* at 192.

188. *See id.* at 208–12.

189. *See supra* Part I.C.1.

190. 433 U.S. at 200 (“since a State’s process could not reach beyond its borders, this Court held after *Pennoyer* that due process did not require any effort to give a property owner personal notice that his property was involved in an *in rem* proceeding.”).

191. *Id.* at 205.

192. *Id.* at 206 (emphasis added).

doctrinal keys to unlocking an expansion of due process, just as notice had been a comfortable cushion to soften the reach of personal jurisdiction itself.<sup>193</sup>

Justice Marshall seemed concerned that if *in rem* actions were exempted from minimum contacts scrutiny, courts would be trapped in a *Pennoyer*-era world of rigid assumptions about notice. If notice matters as a due process consideration, courts should be able to account for notice (or its absence) functionally as a matter of minimum contacts. Marshall rejected the *Pennoyer*-era polarized thinking about notice, in which courts were to assume that attachment of property would almost always provide requisite notice for *in rem* actions, but anything short of personal service was likely deficient in notifying parties of *in personam* actions. What had changed between *Pennoyer* and *Shaffer* was a more relaxed sense of the sufficiency of substituted service, but also the increasing ubiquity of intangible property and the growth of interstate commerce.<sup>194</sup>

The opinions reflect a real division among the judges as to the role of notice in personal jurisdiction. Justice Stevens wrote a short concurring opinion to stress that notice should have been at the heart of the Court's concern about the Delaware statute, thus centering both notice of suit and notice of jurisdiction. He opened with the declaration that "[t]he Due Process Clause affords protection against 'judgments without notice,'"<sup>195</sup> and then emphasized the constitutional requirements for sufficient substituted service of process.<sup>196</sup> Justice Stevens then pivoted to notice of jurisdiction, opining that "[t]he requirement of fair notice also, I believe, includes fair warning that a particular activity may subject a person to the jurisdiction of a foreign sovereign,"<sup>197</sup> and detailed the sort of conduct that might have alerted Greyhound's directors and officers to the possibility that they could be subject to suit in Delaware. His argument is pure assertion—he cites no authority for this proposi-

---

193. Also interesting is that Justice Marshall turned to *Mullane* for its holding that 14th amendment rights cannot be divided based on a classification as *in personam* or *in rem*. *Id.*

194. *See also* *Rush v. Savchuk*, 444 U.S. 320 (1980) (no minimum contacts with a forum state when jurisdiction is based on *quasi in rem II* attachment of an automobile liability insurance policy in a state in which the defendant otherwise had no other contacts).

195. *Id.* at 217.

196. *Id.* at 217–18. Curiously, Stevens cited *McDonald v. Mabee* rather than *Mullane* for the due process standards of notice via service of process. This is perhaps because *McDonald* is a far stronger statement about the need for service that is likely to reach the defendant than the language in *Mullane*.

197. *Id.* at 218.

tion, although it does presage the more formal advent of notice of jurisdiction in *World-Wide Volkswagen* three years later.<sup>198</sup>

*D. Jurisdiction Over Plaintiffs and the Return to the Touchstone of Notice of Suit*

In 1985, the Supreme Court was presented with a new opportunity to engage in an extended treatment of notice and personal jurisdiction. *Phillips Petroleum Co. v. Shutts*<sup>199</sup> was a class action filed in Kansas state court by nearly 28,000 natural gas royalty owners who were residents of all 50 states, D.C., and several foreign countries. Phillips did not challenge whether it was subject to the personal jurisdiction of the Kansas court. Rather, it challenged the court's jurisdiction over the absent class members by linking an alleged problem with notice to a lack of jurisdiction. The notice to the absent class members required an affirmative response in order to opt out of the action, and Phillips argued that this was constitutionally deficient.<sup>200</sup>

The Supreme Court had not confronted a major substantive challenge to both notice and personal jurisdiction in the same case since *Mullane* and some of the pre-*International Shoe* cases such as *McDonald v. Mabee*.<sup>201</sup> Evaluating jurisdictional and notice challenges side-by-side forced the Court to reckon with the relationship between them in a more concrete way than it had done in decades. Although the Court did not announce an explicit holding or theory linking the two doctrines, the *Shutts* opinion shed some light about what role notice plays in post-*International Shoe* personal jurisdiction analysis.

As a general Fourteenth Amendment matter, the Court held that a claimant has a due process right in her claim which is a "chose in action."<sup>202</sup> Thus, it was appropriate for the Court to consider whether Kansas did, in fact, have personal jurisdiction over the absent class members and whether notice satisfied the requirements of due process.<sup>203</sup> On the substantive notice issue, the Court held that the opt-out notice satisfied the *Mullane* standard, both in its contents (which Justice Rehnquist described as "fully descrip-

---

198. See *supra* notes 161-165 and accompanying text.

199. 472 U.S. 797 (1985).

200. *Id.* at 799-802. Phillips also raised significant choice-of-law issues, and *Shutts* therefore stands for several important procedural issues in class action litigation.

201. See *supra* note 82 and accompanying text.

202. *Shutts* at 807.

203. *Id.* at 807-08.

tive”)<sup>204</sup> and its delivery. Grafting an “opt-in” requirement for class actions onto due process “would probably impede the prosecution of those class actions involving an aggregation of small individual claims.”<sup>205</sup>

Turning to personal jurisdiction, the Court openly acknowledged that the absent non-resident class members did not have minimum contacts with Kansas. A defendant in the same position as the claimant royalty owners would not be subject to personal jurisdiction in Kansas. This did not matter, however, because procedurally, “[a] class-action plaintiff . . . is in quite a different posture.”<sup>206</sup> Justice Rehnquist described the litigation burdens on defendants as far greater than those on absent plaintiffs. He stressed that the consequences of a default judgment for a defendant are more severe than the *res judicata* effects of a judgment on an absent plaintiff. He also detailed the many procedural protections that class action procedures put in place to protect absent plaintiffs—procedures that do not exist to protect the interests of no-show defendants.<sup>207</sup> The Court concluded with the following pronouncement:

If the forum State wishes to bind an absent plaintiff concerning a claim for money damages or similar relief at law, it must provide minimal procedural due process protection. The plaintiff must receive notice plus an opportunity to be heard and participate in the litigation, whether in person or through counsel. The notice must be the best practicable, ‘reasonably calculated, under all the circumstances, to apprise interested parties of the pendency of the action and afford them an opportunity to present their objections.’<sup>208</sup>

Observe what the Court is saying here: it is possible, for some litigants, that personal jurisdiction consists *entirely of notice* and does not require *any* of the other considerations that personal jurisdiction doctrine had accrued over the years.<sup>209</sup>

---

204. *Id.* at 812.

205. *Id.* at 813.

206. *Id.* at 808.

207. *Id.* at 808–13.

208. *Id.* at 811–12.

209. An alternative reading of this passage would be to say that absent class members do not have a specific due process right to resist personal jurisdiction, and may challenge due process on notice grounds. However, this would be contrary to the language of the Court earlier in the opinion which speaks directly to plaintiffs’ due process rights in personal jurisdiction. Moreover, subsequent cases and scholarship all treat *Shutts* as standing for the proposition that courts have personal jurisdiction over absent class members, and not simply that absent class members have notice rights. Even *Bristol-Myers Squibb*, which largely undid the

That is a rather astonishing conclusion hiding in plain sight, that personal jurisdiction could be supported entirely on the basis of constitutionally sufficient notice, just as it could be supported entirely on the basis of consent to jurisdiction.<sup>210</sup> Even property located within the jurisdiction had been subject to minimum contacts in *Shaffer*. But the *Shutts* Court placed class action plaintiffs on the periphery of the world where minimum contacts governs, holding that “a forum State may exercise jurisdiction over the claim of an absent class-action plaintiff, even though that plaintiff may not possess the minimum contacts with the forum which would support personal jurisdiction over the defendant.”<sup>211</sup>

*Shutts* opened the door to the idea that notice is not simply a due process fellow traveler with personal jurisdiction, nor is it just a consideration that can be invoked at a court’s convenience to justify jurisdictional expansion or jurisdictional contraction via due process expansion. It could be, in at least some instances, synonymous with personal jurisdiction itself.

Under this aggressive reading of *Shutts*, one could imagine a world in which a set of procedural protections for *defendants*, if crafted and implemented correctly, could satisfy all of Justice Rehnquist’s concerns. In this world, the minimum contacts test would not only provide the legal fictional substitute for presence in the forum state, but would also act as a proxy for additional procedural protections for the defendant. In this alternative world, the demand for minimum contacts would wane in relationship, to the extent that other procedural mechanisms provided the protections that minimum contacts would otherwise cover. For example, a court might point to the rules of venue, transfer, or forum non conveniens to show that minimum contacts are not needed to protect a defendant’s interest in avoiding litigation in an inconvenient or burdensome forum. The only personal jurisdiction concerns that could not be meaningfully addressed by adequate procedural protections for defendants would be sovereignty and territoriality. It is

---

holding of *Shutts*, did not disturb that premise. See *Bristol-Myers Squibb*, 137 S. Ct. 1773, 1782-83 (2017) (clarifying that the Court’s holding did not disturb the applicability of *Shutts* to the notice and personal jurisdiction requirements for absent class members).

210. See *Carnival Cruise Lines v. Shute*, 499 U.S. 585 (1991) (forum-selection clause); *Bremen v. Zapata Off-Shore Co.*, 407 U.S. 1 (1972) (forum-selection clause); *Hess v. Pawloski*, 274 U.S. 352 (1927) (non-resident motor vehicle statute). See generally Tanya J. Monestier, *Registration Statutes, General Jurisdiction, and the Fallacy of Consent*, 36 *CARDOZO L. REV.* 1343 (2015).

211. 472 U.S. at 811.

no surprise that they are also the concerns that are the most awkward fit with due process.<sup>212</sup>

We have already peeked into an alternative universe in which minimum contacts is a unified—if flexible—standard that serves both personal jurisdiction and notice. This was the alternate universe that I proposed as a plausible reading of the treatment of notice in *International Shoe* and the treatment of personal jurisdiction in *Mullane*.<sup>213</sup> As the subsequent cases have shown, the Court never really pursued that path, choosing instead to use the concept of notice, sometimes of suit, sometimes of jurisdiction, as a due process crutch rather than as an explicit and fully realized part of constitutional personal jurisdiction analysis. *Shutts* shows that this link was still lurking in the doctrinal background; that a more direct connection between these two due process doctrines might lend greater clarity to the due process contours of personal jurisdiction than the practice of justifying evolutions in the minimum contacts standards with passing reference to an ever-shifting concept of notice.

While this “additional procedural protections for defendants” approach is a tenable if bold inference one can draw from the logic of the *Shutts* opinion, it is clear that the Court did not approve of this approach, even at the time. Rules for venue, transfer, and forum non conveniens that could easily serve as “additional procedural protections” were just as alive and well in 1985 when the Rule 23 procedures that the Court cited as ensuring fair treatment of absent class members. Rehnquist contrasted absent class members to ordinary *in personam* defendants for whom the minimum contacts test still applied, thus showing a tacit belief that that rule or statute-based procedural protections for defendants are not to be trusted at a constitutional level, and the due process backstop ought to remain.

Two years after *Shutts*, the Court reiterated its reluctance to create or infer procedural protections for defendants in the absence of affirmative state or federal rules. In *Omni Capital International v. Rudolf Wolff & Co.*,<sup>214</sup> the Court considered the following puzzle: the plaintiffs sued English defendants in Louisiana federal court alleging violation of federal commodities laws and related

---

212. See Whitten (Part Two), *supra* note 37 at 808; Wendy Collins Perdue, *Sin, Scandal, and Substantive Due Process: Personal Jurisdiction and Pennoyer Reconsidered*, 62 WASH. L. REV. 479, 510-18 (examining and questioning the link between due process and several putative interests addressed by personal jurisdiction doctrine).

213. See *supra* note 104 and accompanying text.

214. *Omni Capital Int'l, Ltd. v. Rudolf Wolff & Co.*, 484 U.S. 97 (1987).

state law causes of action. The defendants did not fall within the terms of the Louisiana long-arm statute and the federal commodities law was silent as to service of process. The defendants almost certainly had minimum contacts with the United States as a whole.<sup>215</sup> The Supreme Court held that neither the Louisiana long-arm statute, nor Federal Rule of Civil Procedure 4 authorized service of process on the English defendants. The Court noted that it was “unclear . . . whether it is open to us to fashion a rule authorizing service of process,”<sup>216</sup> that it would be “unwise for a court to make its own rule authorizing service of summons” and that “federal courts cannot add to the scope of service of summons Congress has authorized.”<sup>217</sup> If the relevant sovereign has not designated a method for service of process on a particular type of party, then there is no procedure for ensuring that their due process right of notice *or* personal jurisdiction is met. And the Court will be parsimonious in fashioning such a rule.

Observe how *Omni Capital* allowed the Court to double down on “procedural protections” as a jurisdictional limiting device, rather than as an invitation to innovate due process solutions. The Court observed that the defendants could not be served with process, refused to fashion a rule that would permit service, and, finally, commented that the defendants were unreachable.<sup>218</sup> One could imagine that the upshot of situations like *Shutts* and *Omni Capital* would be to highlight the precise types of procedural protections that help ensure due process, and then demand that courts interpret and apply personal jurisdiction and notice doctrines in a manner that encourages procedural protections. The due process right, then, seems not to be the procedural protections themselves, but simply the right for parties to fortuitously be in a forum where such protections already exist.

*Omni Capital* simultaneously centers a party’s due process right to notice and then subordinates that right to the authority of the relevant sovereign to craft the mechanics for such notice. It is a reminder that notice is the concept that conveniently aligns with intuitive notions of the liberty interest that the due process clauses

---

215. *Id.* at 100–02. The Supreme Court accepted the finding of the lower courts that service was not authorized under the Louisiana long arm statute, and therefore did not fully analyze whether personal jurisdiction in Louisiana would be unconstitutional. One may infer from the opinion that the question of minimum contacts with Louisiana was, at best, a close call.

216. *Id.* at 108.

217. *Id.* at 109.

218. *Id.*

protect, but that the sovereign authority of the forum will almost always take precedence. Notice and the attendant mechanics of service process are as much about due process window dressing as they are about aggressively pursuing procedural protections for parties before a court.

Beyond the question of additional procedural protections for parties to an action, *Shutts* also sheds some light on the phenomenon of notice of jurisdiction described in the previous section. Circular as it is, notice of jurisdiction was a concept born of the intuition that personal jurisdiction and due process could provide procedural protections to absent defendants. Notice of jurisdiction is the *in personam* analog of an “opt-in” class. It posits that defendants should have to do things that affirmatively “choose” a forum, that such a choice will prepare a defendant for litigation, and that appearing in a forum will be neither surprising nor particularly burdensome. Notice of jurisdiction can be seen as central to Justice Scalia’s plurality opinion in *Burnham* in which he reasoned that “tag jurisdiction” is constitutional without regard to minimum contacts because in state personal service had always been an accepted basis for jurisdiction, and this long-standing historical fact is *itself* sufficient notice of a circumstance under which one might be haled into court to defend a lawsuit.<sup>219</sup> In other words, Justice Scalia advocated for another scenario in which notice (here, notice of jurisdiction) could itself be an “additional procedural protection” that obviated the need for minimum contacts.

None of this solves the circularity problem of foreseeability as notice of jurisdiction, and circularity is one of the biggest problems with using this instantiation of notice as a genuine procedural protection. However, the “additional procedural protection” justification rounds out the story of how the Supreme Court landed on this oddly circular formulation of the minimum contacts test. For decades, notice had played a quiet supporting role in the doctrinal development of personal jurisdiction. While it never stole the spotlight from other prominent considerations, such as state power and sovereignty, fairness and reasonableness, and convenience, it was undeniably in the mix of considerations and values that the Court turned to in defining the constitutional scope of personal jurisdiction.

---

219. *Burnham v. Superior Court of Cal.*, 495 U.S. 604, 618–19 (1990).

#### IV. WHITHER NOTICE?

From the pre-*Pennoyer* era through the first four decades after *International Shoe*, notice had been a fairly regular, if underappreciated, factor in personal jurisdiction analysis. Although notice did not feature in every single personal jurisdiction case, it reliably appeared at most of the key inflection points in the development of personal jurisdiction analysis, sometimes used to justify a jurisdictional expansion, and sometimes used to mold new doctrinal frameworks for due process expansion. But starting in the 1980s, notice began to fade away. By the post-*Asahi* era of Supreme Court jurisprudence, notice had vanished altogether.

##### A. *Notice Vanishes from General Jurisdiction.* . .

As detailed in the previous sections, several cases in the 1980s had a significant notice element, from the stream of commerce cases, to the *in rem* and class action cases, and even smaller mentions in cases like *Burger King*. But alongside these jurisdictional innovations, in 1984 the Court revisited general jurisdiction for the first time since *Perkins* in 1952. It is here that we see the first signs that notice would fade away.

The plaintiffs in *Helicopteros Nacionales de Colombia v. Hall* (*Helicol*),<sup>220</sup> filed a lawsuit in Texas against Helicol, a Colombian company, for damages arising out of a helicopter crash in Peru. Helicol had assorted business dealings with persons and companies in Texas, some of them related to purchase and operation of the ill-fated helicopter.<sup>221</sup> In finding the exercise of general jurisdiction under these facts unconstitutional, the Court stressed that the defendant's sporadic business dealings with Texas did not rise to the level of systematic and continuous contacts that would justify the exercise of jurisdiction over Helicol for any and all claims brought against it.<sup>222</sup> While the Court invoked the language of unilateral activity and purposeful availment, it did not discuss how or whether the defendants received actual notice of the lawsuit, nor did it discuss the relevance of foreseeability of lawsuits in the forum. It was enough for the Court to detail the defendant's sporadic contacts,

---

220. 466 U.S. 408 (1984).

221. Because the plaintiffs had sued under a general jurisdiction theory, the Court evaluated the contacts on that basis and did not consider the relationship of the contacts to the cause of action. *Id.* at 415 (“[A]ll parties . . . concede that respondents’ claims against Helicol did not ‘arise out of’ and are not related to Helicol’s activities within Texas”).

222. *Id.* at 414–18.

assess them against language of systematic and continuous activity, and draw conclusions about sufficient minimum contacts for personal jurisdiction.

Recall that in *International Shoe*, the Court twice connected the idea of “systematic and continuous” contacts with the concept of notice; first with notice of suit and then again with notice of jurisdiction.<sup>223</sup> The term “systematic and continuous” was not only about creating a new doctrine that would stand in for physical presence in the forum state, but was also about capturing other due process values inherent in personal jurisdiction. In *Helicol*, the Court did address the question of fairness, but the metric for fairness was almost entirely that of the degree of presence in the state. When Court addressed general jurisdiction in *Perkins*, the evaluation of systematic and continuous contacts included notice.<sup>224</sup> In *Helicol*, the Court aligned fairness almost exclusively with the extent of the defendant’s presence in the state. Disconnecting fairness in personal jurisdiction from notice would be the canary in the coal mine signaling that the Court would eventually abandon “systematic and continuous” almost altogether nearly thirty years later.

After more than two decades of silence, the Supreme Court returned to personal jurisdiction in 2011 addressing both general and specific jurisdiction. The unanimous *Goodyear v. Brown*<sup>225</sup> decision drastically narrowed and reformed general jurisdiction, discarding the old assumption that companies with large and consistent retail or commercial activity in a state were subject to general jurisdiction<sup>226</sup> in favor of a standard in which only defendants who are “essentially at home” in the forum state are subject to general jurisdiction.<sup>227</sup> The opinion makes no mention of notice, foreseeability, or even the general expectations of the parties. Pur-

---

223. See *supra* Part II.A.

224. See *supra* notes 134-141 and accompanying text.

225. *Goodyear Dunlop Tires Operations, S.A. v. Brown*, 564 U.S. 915 (2011).

226. See, e.g., Charles W. “Rocky” Rhodes & Cassandra Burke Robertson, *Toward a New Equilibrium in Personal Jurisdiction*, 48 U.C. DAVIS L. REV. 207, 217-18 (explaining that *Goodyear* and *Bauman* effectively eliminated general “doing business” jurisdiction, replacing it with the “essentially at home” test); Allan R. Stein, *The Meaning of “Essentially at Home” in Goodyear Dunlop*, 63 S.C. L. REV. 527, 532 (2012) (“stating that the “essentially at home” standard is new and “does not appear in any prior federal or state judicial decision.”).

227. *Goodyear Dunlop Tires Operations*, 564 U.S. at 919. Although the Court stopped short of saying that a company could only have one “jurisdictional home” analogous to the single “principal place of business” from diversity jurisdiction, the Court’s definition of “essentially at home” limited a defendant’s exposure to general jurisdiction to very few states. See Tanya J. Monestier, *Where is Home Depot at Home?: Daimler v. Bauman and the End of Doing Business Jurisdiction*, 66 HASTINGS

poseful availment is invoked only for its role in highlighting the benefits and privileges that a defendant gets from being “essentially at home” in a forum.<sup>228</sup> *Helicol* had already set the precedent for ignoring any role of notice in evaluating the constitutionality of general jurisdiction. *Goodyear* followed suit. While the absence of notice in *Helicol* provides some continuity with the similar absence in *Goodyear*, *Goodyear* is probably better grouped with the new round of cases in which notice is absent.

Notice, and even any mention of foreseeability or notice of jurisdiction, appears neither in the majority nor the dissenting opinions in *Daimler AG v. Bauman*<sup>229</sup> and *BNSF Railway Co. v. Tyrrell*,<sup>230</sup> both general jurisdiction cases. General jurisdiction was, thus, the first realm of personal jurisdiction analysis in which the Court extinguished notice as a part of jurisdictional discourse. Notice had once been a helpful due process concept that propped up the exercise of general jurisdiction. By *Daimler*, notice had become either superfluous or an actual barrier to jurisdictional contraction. In specific jurisdiction, the Court had harnessed the concept of notice, seizing on notice of jurisdiction as a useful due process justification for limiting personal jurisdiction. However, general jurisdiction had become problematic. Any defendant large enough to have even been considered a meaningful candidate for general jurisdiction within a forum state would be easy to find and easy to serve with process. Since the Court was not interested in jurisdictional expansion, the concept of notice was of little use. But drawing attention to notice would raise the inconvenient proposition that a broad scope of general jurisdiction really could exist comfortably within the bounds of due process to the extent that due process is bound up with notice.

Thus, the Court used *Daimler* to put even more distance between the “systematic and continuous” standard and the aspects of fairness and reasonableness that courts had previously found were central to procedural due process. According to Justice Ginsburg, the only “continuous and systematic” contacts that are strong enough to make a defendant subject to general jurisdiction are those that render the entity “essentially at home” in the forum state, thus collapsing it almost entirely back into a proxy for territorial presence.<sup>231</sup> Under the pre-*Goodyear* and *Daimler* regime, the pre-

---

L.J. 233, 235-36 (2014) (summarizing the change from general “doing business” jurisdiction to the “essentially at home” standard).

228. *Id.* at 924.

229. 571 U.S. 117 (2014).

230. 137 S. Ct. 1549 (2017).

231. 571 U.S. at 127.

vailing thought was that a large company might be subject to general jurisdiction in many states because in each relevant state, its activities were so thorough and pervasive that they were functionally indistinguishable from local businesses.

This analogy to local businesses should be about the *totality* of circumstances that make it constitutionally permissible to sue a defendant on a general jurisdiction theory.<sup>232</sup> Not only does a high volume of systematic and continuous activity ensure that a non-natural defendant has something approximating the physical presence in a jurisdiction that would justify the exercise of sovereign and territorial power, but it is a proxy for the other due process values as well, such as the relative burden of litigating in a distant jurisdiction, the relevance of reciprocal benefits and burdens of operating within the forum state, and, yes—the fact that a defendant with a strong local presence is easy to notify.

Thus, in hindsight, the little-noticed abandonment of notice in *Helicol* was an early sign that the Court would soon retreat from other due process values as well in general jurisdiction analysis, collapsing the whole enterprise back into a question foremost about the sort of presence that one associates with sovereignty and territoriality, and not the sort of presence that could, taken holistically, justify the exercise of personal jurisdiction under a number of due process justifications. And where general jurisdiction led, specific jurisdiction would soon follow.

*B. . . . And Then Notice Disappears from Specific Jurisdiction*

Considering the steady presence of notice, both notice of suit and notice of jurisdiction, in specific jurisdiction jurisprudence through *Asahi*, the absence of notice discourse in the Supreme Court's latest round of specific jurisdiction cases is startling. The plaintiff in *J. McIntyre Machinery, Ltd. v. Nicaastro*<sup>233</sup> was injured at work in his home state of New Jersey by a metal shearing machine that a British manufacturer sold through an American distrib-

---

232. Cf. Lea Brilmayer et. al, *A General Look at General Jurisdiction*, 66 TEX. L. REV. 721 (1988) (discussing unique and non-unique forum affiliations that form the basis for general jurisdiction). Justice Ginsburg cites this article in *Daimler* for the proposition that a defendant must be like a local business in order to be subject to general jurisdiction and uses this idea to support her “essentially at home” test. *Daimler*, 571 U.S. at 151-52. But the Brilmayer et al. article says no such thing – although the authors cite the “unique” affiliations of domicile, principal place of business, and state of incorporation as the “paradigm” bases of general jurisdiction, they explicitly identify “nonunique” bases of jurisdiction as bases that courts “have properly relied on . . . for general jurisdiction.” *Id.* at 735.

233. 564 U.S. 873 (2011).

uter.<sup>234</sup> The litigants and judges appeared to agree that J. McIntyre could foresee that its product would end up in New Jersey and cause injury there. The company used its distributor to sell its products in the U.S. market, attended trade shows in several states, and knew that possible buyers and users existed across America including New Jersey. The case presented the possibility of resolving the persistent “stream of commerce” debate that *Asahi* left unresolved. Is “mere foreseeability” that a product would reach the forum state a sufficient minimum contact, albeit one still subject to due process considerations of fairness and reasonableness? Or do minimum contacts require additional purposeful or targeted activity in the forum state beyond the foreseeability of a regular stream of commerce?

The Court was unable to generate a majority opinion that settled the “stream of commerce” issue, instead granting judgment in favor of J. McIntyre and issuing two plurality opinions and one dissent. Justice Kennedy wrote for a four-justice plurality. His *J. McIntyre* opinion is famous for its full-throated defense of the power theory of jurisdiction,<sup>235</sup> stressing throughout the opinion that any exercise of personal jurisdiction by New Jersey under these facts would violate jurisdictional principles of sovereignty and territoriality.<sup>236</sup> Justice Breyer, joined by Justice Alito, declined to sign on to Justice Kennedy’s reasoning, but agreed that New Jersey could not exercise jurisdiction over the manufacturer. Justice Breyer argued that the Court need not break new doctrinal ground to do so<sup>237</sup> and expressed a general frustration that the case was not about the Internet, the realm in which he believed doctrinal innovation should take place.<sup>238</sup> Justice Ginsburg’s dissent stressed that New Jersey was part of a national market that J. McIntyre targeted in its sales and marketing efforts, criticized Kennedy’s heavy conceptual reliance on sovereignty, presence, and consent, and emphasized the ways in which subjecting the manufacturer to suit in New Jersey

---

234. *Id.* at 878-79.

235. *See, e.g.*, Patrick Borchers, *J. McIntyre Machinery, Goodyear, and the Incoherence of the Minimum Contacts Test*, 44 CREIGHTON L. REV. 1245, 1255-56 (2011) (analyzing the plurality’s emphasis on power and sovereignty); Glenn S. Koppel, *The Functional and Dysfunctional Role of Formalism in Federalism: Shady Grove Versus Nicaastro*, 16 LEWIS & CLARK L. REV. 905, 961 (2012) (describing the plurality’s “strict adherence to state forum-focused sovereignty.”); Stewart E. Sterk, *Personal Jurisdiction and Choice of Law*, 98 IOWA L. REV. 1163, 1198-1200 (2013) (exploring Justice Kennedy’s “focus on power.”).

236. *J. McIntyre*, 564 U.S. at 879-82.

237. *Id.* at 887-89 (Breyer, J., concurring).

238. *Id.* at 890 (Breyer, J., concurring).

would not be unconstitutionally burdensome, inconvenient, or unfair.<sup>239</sup> None of the opinions made even a glancing reference to notice.

And yet, few if any commentators have remarked upon the absence of notice as a due process consideration in *J. McIntyre*.<sup>240</sup> Notice had long been more or less absent from personal jurisdiction commentary, but as we have seen, notice itself persisted for over 100 years after *Pennoyer* as part of personal jurisdiction jurisprudence. Consider the role that notice played in the long doctrinal journey from *Pennoyer*, where notice and service of process shaped Justice Field's conceptions of sovereignty and territoriality, to the use of notice at nearly every doctrinal inflection point to offer reassurances of due process compliance, or to push for due process expansion. Notice had once been closely tied to the notions of sovereignty and territoriality upon which Kennedy centered his opinion, and notice had once been a part of the conceptual foundation that fueled the development of stream of commerce doctrine in the first place. But by *J. McIntyre*, the Court seemed entirely uninterested in implications of whether the defendant had been duly notified of the lawsuit, nor in the question of whether the reasonableness aspect of due process should account for the expectations of parties, that is, whether J. McIntyre's conduct could or should have put it on notice that it might be sued in any American jurisdiction where its products were sold.

The disinterest continued in the new spate of Supreme Court personal jurisdiction cases. *Walden v. Fiore*,<sup>241</sup> the Court's latest statement on personal jurisdiction over parties alleged to have committed intentional torts, did not mention notice directly. However, Justice Thomas, writing for a unanimous Court, did discuss the relevance of whether a defendant could foresee harm in the forum state. Unlike the Court's reasoning in *Calder* and *Keeton*, Thomas did not extend foreseeability of harm in the forum state to notice of jurisdiction in terms of amenability to suit in the forum. In one way, dropping notice of jurisdiction from personal jurisdiction analysis is a step forward in making more sense out of personal jurisdiction doctrine. A party's ability to foresee being haled into court in the

---

239. *Id.* at 893–910 (Ginsburg, J., dissenting).

240. I count myself among those who paid little attention to this phenomenon, having already written three times about personal jurisdiction since *J. McIntyre* and *Goodyear*. Robin J. Effron, *Letting the Perfect Become the Enemy of the Good: The Relatedness Problem in Personal Jurisdiction*, 16 LEWIS & CLARK L. REV. 867 (2012); Effron, *supra* note 176, at 123.

241. 571 U.S. 277 (2014).

forum state had always suffered from a circularity problem, undermining the very predictability it was meant to create.<sup>242</sup>

Reducing foreseeability to foreseeability of harm only, and not notice of jurisdiction, exposes the whole concept of foreseeability as irrelevant to the question of jurisdiction. Much of the *Walden* opinion is devoted to the idea that the defendant's contacts with the forum must amount to more than mere knowledge or awareness of the *plaintiff's* contacts with the forum.<sup>243</sup> If the defendant's conduct is directed at the plaintiff, then contacts with the forum state are incidental to that targeting. *Walden* demands that the defendant establish minimum contacts with the forum state that operate independently of the plaintiff's fortuitous, and even predictable, presence in the state. This argument, though, reveals the work that notice of jurisdiction had been doing in foreseeability analysis. Recall that notice of jurisdiction grew out of the tradition of using notice to offer reassurances of the presence of due process, or to question whether an assertion of jurisdiction met the requirements of due process. Notice of jurisdiction was *itself* a contact with the forum state. It was the connection between the happenstance of harm to a plaintiff and the location of the plaintiff herself. Notice is as much about the forum as it is about the plaintiff because, while *harm* to the plaintiff might be incidental to the plaintiff's location, the prospect of a legal proceeding is always tied to a particular and theoretically foreseeable forum. With close connection between notice and service of process, notice had always included an important element of state power alongside the age-old intuitions about fairness and due process. The innovators of notice of jurisdiction harnessed this connection as a conceptual basis for foreseeability.

Notice of jurisdiction and foreseeability had always been doomed as a viable conceptual basis for minimum contacts because of the incoherence borne of the circularity problem. But taking away notice of jurisdiction might have doomed foreseeability in its entirety because it robbed foreseeability of the connection to the forum upon which the concept was once based. *Walden* provides a good encapsulation of the story of notice in personal jurisdiction. In the post-*International Shoe* era, the presence of notice helped to create a vague and unstable doctrine: foreseeability of litigation in the forum. In the post-*Asahi* era, the absence of notice only muddled things further.

---

242. See *supra* note 165 and accompanying text.

243. *Walden*, 571 U.S. at 283–85.

The last case in the doctrinal odyssey from *Pennoyer* to the present is *Bristol-Myers Squibb v. Superior Court*.<sup>244</sup> As of this writing, it is the latest word on personal jurisdiction. It is also, perhaps, the final nail in the coffin of the relationship between personal jurisdiction and notice. The plaintiffs in *Bristol-Myers Squibb* (BMS) were a group of over 600 persons who sued BMS in California state court alleging injuries from BMS's drug Plavix. Eighty-six plaintiffs were California residents, and the rest were residents of 33 other states. BMS could not argue that California lacked jurisdiction over the plaintiffs—each plaintiff had sued individually, thus consenting to the jurisdiction of the court. Moreover, *Shutts* all but foreclosed the idea that courts could meaningfully question personal jurisdiction over aggregated claimants.<sup>245</sup> So instead, BMS argued that California lacked jurisdiction over BMS for the claims brought by non-residents. The Supreme Court agreed, holding that California lacked personal jurisdiction over the claims of plaintiffs who “did not allege that they obtained Plavix through California physicians or from any other California source; nor did they claim that they were injured by Plavix or were treated for their injuries in California.”<sup>246</sup>

Justice Alito wrote for himself and seven other justices; only Justice Sotomayor dissented. Other commentators have already begun the scholarly debate over the merits of the opinion and the holding itself.<sup>247</sup> For our purposes, what is striking about the opinion is the absence of any mention of notice in a case where notice once might have figured in the due process analysis. The opinion builds on Justice Kennedy's rhetoric from *J. McIntyre* and firmly centers federalism, sovereignty, and power as the doctrinal north star for personal jurisdiction analysis.<sup>248</sup> But *Pennoyer* itself and the decisions in the decades following routinely discussed notice alongside the application of the power and sovereignty principles at the heart of the *Pennoyer* framework. As we have seen, notice was one of the crucial links between the structural doctrines of sovereignty and territoriality, and the justice-based due process ground that *Pennoyer*

---

244. 137 S. Ct. 1773 (2017).

245. See *supra*, note 203 and accompanying text.

246. *Bristol-Myers Squibb Co. v. Superior Court of California*, 137 S. Ct. 1773, 1778 (2017).

247. See, e.g., Bradt & Rave, *supra* note 73; Michael H. Hoffheimer, *The Stealth Revolution in Personal Jurisdiction*, 70 FLA. L. REV. 499 (2018); Scott Dodson, *Plaintiff Personal Jurisdiction and Venue Transfer*, 117 MICH. L. REV. (forthcoming 2019), available at <https://ssrn.com/abstract=3228023>.

248. *Bristol-Myers Squibb*, 137 S. Ct. at 1780 (“restrictions on personal jurisdiction are . . . a consequence of territorial limitations on the power of the respective states.”). See also, Hoffheimer, *supra* note 247.

claimed as the primary constitutional grounding for the right to resist personal jurisdiction.

Although the *Bristol-Myers Squibb* decision sets forth a forceful recapitulation of sovereignty and territoriality, the Court did not jettison due process.<sup>249</sup> Justice Alito's opinion, however, shows just how awkward the fit between personal jurisdiction and due process had become. This is most evident in Part II.B of the opinion in which Justice Alito addresses the "variety of interests"<sup>250</sup> that a court must consider in assessing whether the exercise of personal jurisdiction is unconstitutional. He begins with a mention of the interests of the forum state and the plaintiff, but dismisses these without any analysis before turning to "the primary concern" which is "the burden on the defendant."<sup>251</sup> He acknowledges that this burden "obviously requires a court to consider the practical problems resulting from litigating in the forum"<sup>252</sup> but does not examine what such problems might be in this case. And this is probably because there is little that he could say—BMS was *already* in the forum state defending *nearly identical claims*, so that the usual jurisdictional bugaboos about distance, convenience, familiarity with the legal system, and other practical burdens would have made little sense in this situation.

Thus, Justice Alito refashioned the burden to "encompass[ ] the more abstract matter of submitting to the coercive power of a State that may have little legitimate interest in the claims in question."<sup>253</sup> Stated differently, Justice Alito added a new "burden" that a defendant might bear—the psychic burden of an alleged violation of state sovereignty. This possibility had lurked in the shadows of personal jurisdiction analysis for decades, and he quoted *World-Wide Volkswagen's* admonition that "the Due Process Clause, acting as an instrument of interstate federalism, may sometimes act to divest the State of its power to render a valid judgment"<sup>254</sup> to support the claim that state sovereignty can be folded into considerations of the burden on the defendant in order to shoehorn a federalism argument into a due process framework.

This part of the Court's opinion is the natural result of the Court's insistence on doctrinal formalism regardless of whether the

---

249. 137 S. Ct. at 1779 ("It has long been established that the Fourteenth Amendment limits the personal jurisdiction of state courts.").

250. *Id.* at 1780.

251. *Id.*

252. *Id.*

253. *Id.*

254. *Id.* (quoting *World-Wide Volkswagen*, 444 U.S. at 294).

results of that formalism had any relationship to how plaintiffs and defendants actually encounter the world or the forum state. Sotomayor's dissent is a plea to apply common sense to the *International Shoe* framework of traditional notions of fair play and substantial justice. For decades, notice had provided the connection between due process formalism, the intuitions of fair play and substantial justice, and the exercise of authority over out-of-state defendants. Stripped of notice, the Sotomayor dissent in *Bristol-Myers Squibb* and the Ginsburg dissent in *J. McIntyre* stand as simple pleas for the constitutional law of personal jurisdiction to conform to modern realities of commerce and contact. With one of the natural pillars of due process gone, the remaining generalized appeals to considerations of reasonableness in light of the relative interests of the defendant, plaintiff, and forum state fell on increasingly deaf ears. The due process innovations of *International Shoe* had finally collapsed back into the world of *Pennoyer*, and the only remaining task was to sort out the formal boundaries of minimum contacts. The task of relating such an inquiry to other due process values had been set aside.

## V.

### NOTICE RESURRECTED

The primary focus of this Article has been devoted to the explanatory project of demonstrating that the history of constitutional personal jurisdiction doctrine is incomplete without a thorough account of how notice has aided the development, maintenance, and evolution of personal jurisdiction as a due process doctrine. The story illuminates a history that has been hiding in plain sight. But beyond that, it gives a more satisfying answer to the question of why the Court has been able to so deftly drop core due process concerns from personal jurisdiction analysis by neatly redefining each in terms of the older *Pennoyer*-style justifications. The evolution of notice of jurisdiction from notice of suit had provided a template for refashioning individual liberty concerns into sovereignty and territoriality arguments. And the eventual absence of notice at all provided the warning signal that the Court was willing to drop some due process concerns altogether.

In this Part, I sketch a path forward for reviving notice as a due process concern in personal jurisdiction analysis. Resurrecting notice would help to construct a broad personal jurisdiction doctrine. Court access for plaintiffs could be re-elevated as a core justice and due process concern. It is easier to embrace a personal jurisdiction doctrine that includes plaintiff or forum-focused analyses if courts

can reassure themselves that a core component of the defendant's personal jurisdiction due process rights has been satisfied.

That being said, personal jurisdiction cannot completely collapse back into notice. The two doctrines were never synonymous nor interchangeable. It would be strange to suggest that a return to notice would entail having notice swallow personal jurisdiction in its entirety. Rather, notice should be a meaningful factor in personal jurisdiction analysis. Such a resurrection would be tricky. For one thing, doing so is hardly a guarantee of clarity or consistency in personal jurisdiction doctrine. Constitutional personal jurisdiction doctrine has been plagued by such problems since *Pennoyer*, meaning that there is no magical point in time to which we could set back the clock and discover the "perfect" use of notice in an elegantly logical personal jurisdiction doctrine.

How, then, should notice be marshaled in service of a better and broader personal jurisdiction doctrine? *International Shoe* provides the first clue. Notice would allow the Court to resurrect a broader general jurisdiction theory. The mistake that the Supreme Court made in the post-2011 cases was to focus almost exclusively on how minimum contacts are a proxy for presence at the expense of other core due process values. The problem is that *International Shoe's* minimum contacts language was never *just* about presence and the power that a forum may exercise over non-natural persons. The requirement of systematic and continuous contacts ensures that a forum will not exercise jurisdiction over a defendant who is unlikely to anticipate jurisdiction or hear of a pending lawsuit. It is possible to repackage both of these notice-related concerns in service of a new, notice-inclusive approach.

Briefly stated, the notice-inclusive approach to personal jurisdiction has four components: (1) Establishing comfort with a broadly available and easily-satisfied jurisdictional standard; (2) Re-committing to a deeper constitutional examination of notice and service of process; (3) Using specific jurisdiction as a doctrine that provides constitutionally required "additional procedural protections" for defendants who are not subject to general jurisdiction, but nevertheless have a connection with the forum in which that personal jurisdiction may be constitutionally appropriate; (4) Refashioning general jurisdiction on systematic and continuous contacts that represent presence and notice while limiting its scope by reference to constitutional concerns such as burden and convenience. I shall address each of these in turn.

A. *Establishing Comfort with Easily-Satisfied Due Process Criteria*

Scholars have routinely tied the transformation of personal jurisdiction doctrine in *International Shoe* to the growth of a larger and more complex national economy from the late Nineteenth Century through World War II.<sup>255</sup> These changes in commerce certainly necessitated a more capacious personal jurisdiction doctrine that accounted for an economy that operated seamlessly across state borders. But this era brought another change as well. As the economy grew and became more interconnected, so too did the modes of communication and transportation that eased the burdens and uncertainties of service of process. The due process entitlement to be notified of a pending action was one that courts could not assume would always be easily satisfied. Thus, constructing rules that ensured actual or constructive notice of a lawsuit was a crucial feature of the exercise of state power over a defendant. Notice was not taken for granted, and many of the personal jurisdiction decisions of the pre-*Pennoyer* and pre-*International Shoe* era reflect this concern. Notice, along with other due process and sovereignty considerations, was part of a system of actively policing the boundaries of state territorial power.

In erecting and maintaining personal jurisdiction barriers, there seems to be an unspoken norm that personal jurisdiction should be *hard*. Perhaps it was this subconscious realization that caused the Court to slowly abandon notice in personal jurisdiction analysis. An admission that a core due process value of personal jurisdiction, notice, is quite easily satisfied would make personal jurisdiction itself too easy. Rather than make the self-congratulatory admission that modern commerce and communications have relegated the centuries-old concerns about notice somewhat obsolete, the Court quietly turned to the other concerns. It is as if the criteria for centering a personal jurisdiction rationale focus on whether that doctrine will be hard to satisfy, and not whether the doctrine itself is a good fit for either due process or common-sense boundaries on jurisdiction. Thus, the Court slowly let go of a due process

---

255. See, e.g., Damon C. Andrews & John M. Newton, *Personal Jurisdiction and Choice of Law in the Cloud*, 73 MD. L. REV. 313, 336 (2013) (“The Supreme Court’s adoption of the ‘minimum contacts’ standard in *International Shoe* was a reaction to the evolving methods by which business was conducted in the twentieth century.”); Douglas D. McFarland, *Drop the Shoe: A Law of Personal Jurisdiction*, 68 MO. L. REV. 753, 753 (2003) (After *Pennoyer*, “courts struggled with application of [*Pennoyer*’s] rigid principle to an expanding and increasingly mobile economy, and to a new type of defendant, the corporation.”).

rationale that could have resulted in far more defendants subject to suit in many more states.

It is time to let go of this chain of logic. A recognition that most defendants are easily reachable by in-hand service either in or out of state, by reliable public and private mail delivery services, or by reliable electronic means should be an opportunity for the Court to acknowledge that the baseline of acceptable jurisdiction can move accordingly. If it is true that many more defendants can be notified of a lawsuit in a manner consistent with due process, courts should accept this happy reality, rather than constantly recalibrating personal jurisdiction doctrine so that it continues to be “hard enough” to exclude some vaguely unspecified quantum of out-of-state defendants.

This recognition should extend to due process values beyond notice. For example, the acknowledgement that defending a lawsuit in a geographically distant state is, for many defendants, not terribly burdensome or inconvenient in the modern economy should encourage greater comfort with a broader scope of personal jurisdiction. It should not prompt the Court to discard burdens and inconvenience as a meaningful factor in personal jurisdiction analysis or to completely retrofit the concept so as to equate it with a “harder” standard like territoriality as the Court did in *Bristol-Myers Squibb*.<sup>256</sup> Instead, the Court should take up the invitation to *recalibrate* due process concerns rather than discard them in favor of searching for the criteria that will be the most limiting of personal jurisdiction.

*B. Pressing for Deeper Constitutional Scrutiny of Notice and Service of Process Practices*

A willingness to accept easily-satisfied due process criteria such as notice as meaningful indicia of constitutionality should not encompass an unexamined acquiescence to current constitutional notice doctrine. Sustained constitutional examination of notice doctrine as it relates to notice and service of process itself has been thin and sporadic in the decades since *Mullane*.<sup>257</sup> A more searching approach to notice might bring some added due process heft to the role that notice might play in personal jurisdiction.

While it is beyond the scope of this Article to develop and suggest comprehensive changes in constitutional notice doctrine, it is worth sketching a few avenues for exploration that would fit com-

---

256. See *supra* notes 250-254 and accompanying text.

257. See *supra* notes 130-133 and accompanying text.

fortably within *Mullane*'s "reasonably calculated under the circumstances" standard.<sup>258</sup>

Defendants from vulnerable or underrepresented populations might benefit from more robust constitutional notice protections. These would be, by and large, individual persons or small, local non-natural defendants who lack access to retained or in-house counsel that might provide them with internal systems for accepting service of process and making sense of a summons and complaint. Because most individuals are at least theoretically reachable by a process server, first class mail, or reliable electronic means,<sup>259</sup> it is unlikely that the Court would need to make serious changes to the delivery aspect of service of process. That being said, there could be room for taking seriously the service problems when vulnerable populations are involved, such as persons who lack a steady, fixed address or whose residential or work environments make receipt of service of process less of a certainty, even in the modern economy.<sup>260</sup>

Beyond the mechanics of service of process itself, there is a real opportunity for examination of the constitutional sufficiency of the content of summonses, complaints, and other notices such as the notices sent to absent class members. Constitutionally sufficient notices often involve documents with small print<sup>261</sup> or written in legalese or other inscrutable language.<sup>262</sup> While this poses little problem to well-heeled defendants with reliable access to legal counsel, it can present a real barrier to a natural lay person defen-

---

258. *Mullane v. Cent. Hanover Bank & Tr. Co.*, 339 U.S. 306, 314 (1950).

259. See Christine P. Bartholomew, *E-Notice*, 68 Duke L.J. 217 (2018) (empirical study of electronic notice in class actions); Philip P. Ehrlich, *Comment: A Balancing Equation for Social Media Notice*, 83 U. CHI. L. REV. 2163 (2016) (discussing current issues with electronic forms of notice).

260. See Gottshall, *supra* not 184, at 814 (arguing that "[t]raditional methods of service, which lack reliable verifications, are not reasonably calculated to provide constitutionally adequate notice. The technological advancements that have occurred in the decades following *Mullane*, provide new and better circumstances under which notice must be provided.").

261. Shannon R. Wheatman & Terry R. LeClercq, *Majority of Class Action Publication Notices Fail to Satisfy Rule 23 Requirements*, 30 REV. LIT. 53, 58 (finding in a study of securities class action notices that "over 60% of notices were written in less than an 8-point font.").

262. Todd B. Hilsee, Shannon R. Wheatman, & Gina M. Intrepido, *Do You Really Want to Me to Know My Rights? The Ethics Behind Due Process in Class Action Notice Is More Than Just Plain Language*, 18 Geo. J. L. Ethics 1359 (2005) (describing problems with plain language usage in class action notices).

dant who might need extra time and resources to decode a summons or class action notice.<sup>263</sup>

In the end, a focus on aspects of notice might provide a means to address an important distinction among defendants that has vexed some members of the Court in recent years. Consider the Appalachian potter that caused Justice Breyer such concern in *McIntyre*:

What might appear fair in the case of a large manufacturer which specifically seeks, or expects, an equal-sized distributor to sell its product in a distant State might seem unfair in the case of a small manufacturer (say, an Appalachian potter) who sells his product (cups and saucers) exclusively to a large distributor, who resells a single item (a coffee mug) to a buyer from a distant State (Hawaii).<sup>264</sup>

Justice Breyer's example is focused on the purposefulness of the seller-defendant. This is unsurprising given the long doctrinal emphasis on purposeful availing and targeting of the forum that has driven much of personal jurisdiction analysis since *World-Wide Volkswagen*. But perhaps the obsessive search for the line between targeting an in-state market versus a region versus the nation as a whole obscures other distinctions that are just as relevant from a due process perspective. We might be just as concerned with the fact that a sole proprietor artisan might be significantly less equipped to respond to a legal notice than a large corporation, just as the ability of such a defendant to retain and manage local counsel in a far-flung jurisdiction is quite different from the ability to do so by a big company with its own legal department.

Beefing up constitutional notice doctrine is unlikely to radically redefine personal jurisdiction analysis. The *Mullane* bar is rela-

---

263. See Debra Lyn Bassett, *Just Go Away: Representation, Due Process, and Preclusion in Class Actions*, 2009 B.Y.U. L. REV. 1079, 1115 (“[T]he nature of the claims in many class actions often renders notice by publication necessary, despite its notorious ineffectiveness.”); Susan P. Koniak, *How Like a Winter? The Plight of Absent Class Members Denied Adequate Representation*, 79 Notre Dame L. Rev. 1787, 1811-12, 1815-17, 1823 (2004) (describing problems with the content of class action notices); Tobias Barrington Wolff, *Federal Jurisdiction and Due Process in the Era of the Nationwide Class Action*, 156 U. Pa. L. Rev. 2035, 2075 (2008) (“Absent class members . . . generally do not receive any form of notice regarding the proceeding [denying certification], and certainly not the individual notice and opportunity to opt out that a 23(b)(3) action would require if certification were granted.”). See generally Debra Lyn Bassett, *Class Action Silence*, 94 B.U. L. REV. 1781 (2014) (describing the relationship between problems with class action notice, class members' responses to notice, and the presumption of consent to personal jurisdiction).

264. *J. McIntyre Machinery, Ltd. v. Nicastro*, 564 U.S. 873, 891-92 (2011).

tively low, and while there are good reasons to reinvigorate parts of that doctrine, there is great utility in keeping the due process notice requirement relatively easy to satisfy. It allows courts and legislatures to fashion manageable means of allowing lawsuits and other proceedings to go forward without cutting off a plaintiff's ability to pursue a remedy, or making service of process so onerous that it creates a serious access to justice issue. Nevertheless, treating notice doctrine as a constitutionally significant and live issue ensures that it will not be an immutably easy constitutional hurdle. It is easier to demand comfort with a relatively low constitutional bar when that bar is periodically recalibrated to reflect the underlying due process concerns at hand.

C. *Sharpening Specific Jurisdiction with an "Additional Procedural Protections" Approach*

The Supreme Court has restricted the scope of specific jurisdiction over the past decade. It has accomplished this contraction by focusing heavily on the sovereignty and territoriality concerns. Other due process values have either been discarded, devalued, or refashioned as concerns that are subsumed by sovereignty and territoriality. A notice-inclusive approach to personal jurisdiction could restore a broader due process basis to personal jurisdiction analysis, thus broadening specific jurisdiction's scope.

The minimum contacts standard was never meant to be an exclusive proxy for presence. A close reading of *International Shoe* demonstrates that minimum contacts are meant to provide an assurance that other due process values like notice are protected.<sup>265</sup> The key to unlocking a broader scope of specific jurisdiction is to recognize when a core due process value to personal jurisdiction has been satisfied. If it has not, a court might then look to see if additional procedural protections can or would make up for this deficit. The greater assurance a court has that a defendant has actual or constructive notice of a lawsuit, the more diminished the need to grasp tightly to fictive presence as a value that supersedes all others.

A notice-inclusive approach is, in some senses, a version of the dreaded "sliding scale" of contacts and relatedness that the Supreme Court rejected in *Bristol-Myers Squibb*.<sup>266</sup> Notice is a due pro-

---

265. See *supra* notes 100-102 and accompanying text.

266. *Bristol-Myers Squibb Co. v. Superior Court*, 137 S. Ct. 1773, 1778-89 (2017) (rejecting California's sliding scale approach to specific jurisdiction in which a wider range of general contacts with the forum state compensate for fewer dispute-related contacts with the forum state).

cess concern that provides a meaningful link between wide-spread forum contacts that fall short of general jurisdiction and dispute-related contacts that fall short of the tight relationship between plaintiff, claim, and defendant required by current specific jurisdiction doctrine. A notice-inclusive approach to personal jurisdiction anchors this type of holistic analysis in the due process heart of *International Shoe's* minimum contacts standard. A few examples illustrate this approach.

### 1. Specific Jurisdiction in Mass Actions and Class Actions

In the wake of *Bristol-Myers Squibb*, some commentators have worried that the Supreme Court has opened up personal jurisdiction as yet another tool for defendants to break apart mass actions and possibly class actions as well. Plaintiffs must now show that a court has jurisdiction over a defendant in each plaintiff's individual claim in a class action. Lower courts are split on whether and how this applies to class actions, with some courts holding that a court must have personal jurisdiction over the claims of all absent class members<sup>267</sup> and others holding that only the named representatives are considered for personal jurisdiction purposes.<sup>268</sup>

*Bristol-Myers Squibb* is the logical end of a personal jurisdiction journey in which the Court has been quietly dropping and diluting core due process values from personal jurisdiction analysis, leaving only a skeleton of purposeful availment, sovereignty, and territoriality.<sup>269</sup> Even purposeful availment has lost its early robust dimension. The point of purposeful availment seems only to be a means of strengthening the end conclusions regarding sovereignty and territoriality, rather than standing for larger due process values.

The mass action context has revealed the awkward due process architecture that the Court had constructed over the past two de-

---

267. See, e.g., *Practice Management Support Services, Inc. v. Cirque Du Soleil, Inc.*, 301 F. Supp. 3d 840, 861 (N.D. Ill. 2018) ("It [is] not clear how Practice Management can distinguish the Supreme Court's basic holding in *Bristol-Myers* simply because this is a class action."); *Wenokur v. AXA Equitable Life Ins. Co.*, 2017 WL 4357916 at \*4 n.4 (D. Ariz. 2017) ("The Court also notes that it lacks personal jurisdiction over the claims of putative class members with no connection to Arizona and therefore would not be able to certify a nationwide class.").

268. See *Knotts v. Nissan North America, Inc.*, 2018 WL 4922360 (D. Minn. 2018) (summarizing district court decisions in "California, Louisiana, Florida, Georgia, Virginia, Texas, the District of Columbia, and even Illinois [that] have concluded that there are valid reasons for limiting BMS to named parties—particularly due to the material distinctions between mass tort actions and class actions.").

269. See *supra* notes 246-255 and accompanying text.

cares. While Justice Alito's eight-judge majority opinion was arguably a better fit with the most recent specific jurisdiction jurisprudence, Justice Sotomayor's lone dissent rightly called attention to the lack of a common-sense framework that the Court's decisions up to and including *Bristol-Myers Squibb* had produced.<sup>270</sup>

Notice, among other core due process concerns, might play a role in reviving a meaningful fairness inquiry to specific jurisdiction analysis in mass action cases. In fact, notice might be the missing link between the majority's strict formalism in its "relatedness" requirement for the non-residents' claims and the strained logic of concluding that it would really be contrary to the standards of "fair play and substantial justice" to require the defendant to defend "materially identical"<sup>271</sup> claims to those it is already defending in the forum state.

Observe the impasse here between Justice Alito and Justice Sotomayor: Justice Alito sees a bevy of claims that might look superficially as if they belong in California, but in fact the individual disputes have nothing to do with the state. Justice Sotomayor sees materially identical claims which all arise out of conduct that does relate to the "same essential acts" that included relevant forum conduct in California.<sup>272</sup> Much of Justice Sotomayor's argument hinges on a conviction that a nationwide course of conduct that includes the forum state should be sufficient to show purposeful availment of the forum state. That is a powerful argument in and of itself and one that I (among others) have defended on grounds unrelated to notice.<sup>273</sup>

But mass actions close the loop between purposeful availment and relatedness in an even more concrete way. The nationwide (or, in some cases, regional) course of action shows purposeful availment of the forum. In *Bristol-Myers Squibb*, for example, the defendant marketed and sold the drug at issue directly into California.<sup>274</sup> The existence of the California claims themselves provides an important type of notice. The defendant knows that it is being sued for injuries allegedly caused by Plavix. It knows that it is being sued in California. It is, thus, especially well-positioned vis-à-vis notice of jurisdiction (i.e., notice that a lawsuit about conduct related to this drug might be brought in a forum where it marketed the drug).

---

270. 137 S.Ct. at 1787 (Sotomayor, J. dissenting) ("[O]ur precedents do not require this result, and common sense says that it cannot be correct.").

271. *Id.* at 1785.

272. *Id.*

273. See Dodge & Dodson, *supra* note 176; Effron, *supra* note 176.

274. 127 S.Ct. at 1778.

Such a defendant is also well-positioned vis-à-vis notice of suit. If it is constitutionally permissible to serve Bristol-Myers Squibb with a summons and complaint detailing allegations of Plavix injuries suffered by a patient in California, then the defendant has not just notice, but a sophisticated preview of the dimensions of the claims generally. The device of the mass action itself ties the generalized purposeful availment regarding forum-related conduct to the specificity of actual non-resident claims.

Aggregated litigation already has a host of additional procedural protections that are meant to buffer against due process problems for both plaintiffs and defendants in these actions. Class actions in particular have such structures, and this is what many lower courts have stressed when declining to extend *Bristol-Myers Squibb* to class actions.<sup>275</sup> There are vigorous debates about the scope of these protections in state and federal courts for both class actions and mass actions. A better personal jurisdiction debate would take seriously these additional procedural protections. Jurists should inquire whether these protections enhance a claim to personal jurisdiction, or, alternatively, whether they fall short of the Fourteenth Amendment protections that are core to personal jurisdiction in particular, and not due process generally. Due process problems with aggregation should be addressed directly, and not hidden beneath another due process doctrine.

## 2. Using Registration Statutes to Broaden the Availability of Specific Jurisdiction

States require out-of-state corporations to appoint an agent for service of process upon registration.<sup>276</sup> The jurisdictional effect of registration statutes is constitutionally uncertain, as the Court has not addressed them in the post-*International Shoe* era.<sup>277</sup> Thus, the relationship between registration, minimum contacts, and consent to jurisdiction is unclear. For many decades, registration statutes lurked in the background of personal jurisdiction doctrine because the broader availability of general jurisdiction over large companies made the use of such statutes mostly unnecessary. However, inter-

---

<sup>275.</sup> See *supra* notes 267-268 and accompanying text.

<sup>276.</sup> Matthew Kipp, *Inferring Express Consent: The Paradox of Permitting Registration Statutes to Confer General Jurisdiction*, 9 REV. LITIG. 1, 1-2 (1990) (describing registration statutes and citing examples of state registration statutes).

<sup>277.</sup> *Id.* (describing the doctrinal puzzle of the pre-*International Shoe* case of *Pennsylvania Fire Insurance Co. v. Gold Issue Mining Co.*, 243 U.S. 93 (1917), that appears to authorize personal jurisdiction on a registration statute alone, and *International Shoe* which imposed the minimum contacts test); Monestier, *supra* note 210, at 1361-62 (describing the academic discourse surrounding *Pennsylvania Fire*).

est in using registration statutes as a path to general jurisdiction has surged in the wake of the *Goodyear* and *Daimler* restriction.<sup>278</sup>

Most courts have continued to hold that registration statutes do not confer general jurisdiction on the forum state.<sup>279</sup> But general jurisdiction is not the only possibility. Courts in some jurisdictions have held that registration statutes are a form of consent to the specific jurisdiction of that forum “for causes of action arising from the business that it actually conducts in the state.”<sup>280</sup> Registration statutes could be key to filling the gaps between a narrow general jurisdiction doctrine and the overly restrictive minimum contacts standard set in the post-2011 specific jurisdiction cases. Courts should use notice as part of the doctrinal justification for this move.

The cases in which registration statutes bolstered by notice doctrine might fill a void are those in which the defendant conducts business in the forum state that is related to the wrongdoing in a plaintiff’s cause of action but might not be specifically related enough to that plaintiff’s cause of action to support specific jurisdiction under *J. McIntyre* and *Bristol-Myers Squibb*. This would cover many large businesses with national sales, marketing, or employment schemes. A well-written registration statute might extend a forum’s jurisdiction over all actions that are *related to* the corporation’s business actions within the forum state. These claims need not *arise out of* those specific activities, so long as the claims are related to the same business activities that the defendant has conducted within the state.

Notice is the due process doctrine that closes the loop between the “arise out of” and “related to” concepts that Justice Alito broke apart in *Bristol-Myers Squibb*. If a defendant has consented to jurisdiction for claims *relating to* its business activities, it has notice of jurisdiction that it can be sued for precisely these activities. The additional element of purposeful availment connected to each and every plaintiff is less necessary. Moreover, registration statutes have the added bonus of ensuring an extra layer of comfort with regard

---

278. See Monestier, *supra* note 210, at 1358 (“Now that plaintiffs will have a much harder time establishing general jurisdiction over defendants in all but the most obvious of cases, a different ground of jurisdiction will most certainly take center stage: that of corporate registration.”).

279. See e.g., *Brown v. Lockheed Martin Corp.*, 814 F.3d 619, 625-27 (2d Cir. 2016); *Deleon v. BNSF Rwy. Co.*, 426 P.3d 1 (Mont. 2018); *Amelius v. Grand Imperial, LLC*, 64 N.Y.S.3d 855 (2018) (New York’s registration statute does not confer general jurisdiction over foreign corporate defendant).

280. Monestier, *supra* note 210, at 1370.

to notice of suit via the mechanism of appointing an agent for service of process under the statutes.

If registration statutes are to be used in this manner, it would be part of a larger project of reinvigorating the core due process concerns of personal jurisdiction. For example, courts might use registration statutes to afford jurisdictions the ability to start with a *presumption* of notice in lawsuits related to business activities conducted within the state. But the presumption that due process has been satisfied largely through notice and consent could be overcome by a showing that other concerns are particularly present in a given case. Perhaps a small corporation has registered to do business in all 50 states in the hopes that it will one day be a national company, but only acts regionally. It has scattered business activities in a forum state where a plaintiff chooses to sue for similar conduct in another state. This might be a case in which the court looks to burdens, convenience, and reasonableness to find that specific jurisdiction is unavailable notwithstanding the corporate registration.

*Philips Petroleum v. Shutts* is the doctrinal ancestor of this approach. Recall that a plausible reading of this case was that, as a matter of due process, personal jurisdiction could be justified as almost completely synonymous with notice because of the architecture of the “additional procedural protections” that surround absent class members under Rule 23.<sup>281</sup> Registration statutes should be the analogue of sound class action procedures. They marshal the doctrinal due process resources of notice and associated procedures to enhance forum-related conduct that, in other contexts, might fall short of due process.

#### *D. Restoring Notice to a Broader General Jurisdiction Doctrine*

The collapse of a broadly available general jurisdiction doctrine was a seismic change in personal jurisdiction. Limiting the general jurisdiction of domestic defendants to just one or two states drastically changed the presumed access to courts that plaintiffs previously enjoyed against large companies with a hefty business presence in many or even all states. An explicit reincorporation of the concept of notice into general jurisdiction doctrine would allow the Court to return to a doctrine in which “systematic and continuous” contact with a forum state is sufficient for general jurisdiction without the straightjacket of a doctrinal category in which the only systematic and continuous contacts that matter are the decision to

---

281. See *supra* notes 202-213 and accompanying text.

organize under the laws of a state and the decision to locate headquarters in a forum.

To see how notice can nudge the doctrine back towards a more capacious definition of “systematic and continuous,” we need only to return to the source of that phrase itself, *International Shoe*. Recall that notice had a dual significance to the systematic and continuous concept that *International Shoe* introduced. If a company like International Shoe had systematic and continuous contact with the forum, then the company would justifiably anticipate a lawsuit in the forum<sup>282</sup>—this was the genesis of the emphasis on notice of jurisdiction that would blossom in the late 1950’s through 1980’s. But more critical (and, perplexingly, mostly forgotten) was the Court’s observation that “[i]t is enough that [the defendant] has established such contacts with the state that the particular form of substituted service adopted there gives *reasonable assurance that the notice will be actual*.”<sup>283</sup>

Due process notice doctrine connected concerns about the constitutional viability of substituted service with concerns about asserting jurisdiction over entities that lacked the traditional indicia of physical presence within the territory of a forum state. The absence of notice from latter day personal jurisdiction analysis, then, is at least in part responsible for the constricted standard from *Goodyear* and *Daimler*. Notice had disappeared from personal jurisdiction analysis and, along with it, a central justification for valuing a “systematic and continuous” presence in the forum. This left the Court with precious few justifications for general jurisdiction aside from the power theories of sovereignty and territoriality. From this angle, it’s easy to see how the Court came to view pre-*Goodyear* general jurisdiction as unwieldy and unfair. If territoriality were to mean anything, it had to mean something more than the idea that a powerful presence in a state could be sufficient for general jurisdiction, even if that strong presence were replicated across many or all states within the United States. For the post-*Goodyear* Court, the requisite minimum contacts for general jurisdiction must be such that only a few states could “claim” that entity as “belonging” to the jurisdiction.

But imagine if the Court were to bring back notice as a central due process value in personal jurisdiction. The question of power over a defendant was present in 1945 and it is still present today.

---

282. See *supra* note 103 and accompanying text.

283. *Int’l Shoe Co. v. Washington*, 326 U.S. 310, 316 (1945) (emphasis added).

Notice was never a substitute for concerns about the exercise of state power. But to ignore notice is to take away a central due process justification for constitutional limits on the exercise of jurisdiction; a connection that is arguably one of the best fits with the concept of due process in the first place. Taking pressure off the sovereignty/territoriality justification allows for a world in which it is utterly sensible and fair to understand that some entities are so big that they are, in fact, effectively “present” in a multitude of jurisdictions.

If, once upon a time, the fear had been that large and diffuse presence meant that the Court should be worried about the effectiveness of substituted service, this concern has largely disappeared. In 1945, it probably was still necessary for Justice Stone to explicitly reassure his readers that substituted service was reliable and that minimum contacts could be an indicium of that reliability. Today, the idea that a large corporation might not receive notice of a lawsuit is quaint.<sup>284</sup> So quaint, in fact, that one might balk at the idea of considering that to be a factor in personal jurisdiction analysis. After all, notice is almost too easily satisfied in most cases. In an odd reversal from the concerns at the forefront of the *Pennoyer* era, it is now *natural* persons and not non-natural persons over whom most of the concern about actual notice is focused. This is the role that minimum contacts would play in a notice-inclusive personal jurisdiction analysis: the stronger the defendant’s connection is with the forum, the less one worries that the defendant cannot be found or adequately notified.

Under the notice-inclusive approach described above, the scope of general jurisdiction would be very broad. So broad, in fact, that it might exceed the reach of pre-*Goodyear* general jurisdiction, which was already criticized for its breadth.<sup>285</sup> But this would only be the case in a notice-centered approach, in which the fact of notice and service would be synonymous with personal jurisdiction. As has

---

284. When, from time to time, an administrative snafu leads to a defect in actual notice, it is a startling and headline-catching event. See, e.g., *Joyce v. PepsiCo, Inc.*, 813 N.W.2d 247 (Wis. App. 2012) (an administrative assistant at Pepsi mishandled service forwarded from Pepsi’s registered agent for service of process, Pepsi did not appear in the action and the trial court entered a \$1.26 billion default judgment that it later set aside).

285. See Patrick J. Borchers, *The Problem with General Jurisdiction*, 2001 U. CHI. LEGAL F. 119 (2001) (criticizing general jurisdiction doctrine under *Perkins* and *Helicor* as a doctrine that “do[es] not give clear legal rules for contacts-based general jurisdiction.”); Mary Twitchell, *The Myth of General Jurisdiction*, 101 HARV. L. REV. 610, 629 (1988) (“[C]ourts have no clear concept of what general jurisdiction is or how it relates to specific jurisdiction.”).

already been noted, this was never the case in American jurisprudence, nor should it be. A notice-inclusive approach to general jurisdiction would be part of a larger project of reinjecting other core due process concerns into a general jurisdiction minimum contacts analysis. Some of these due process concerns, such as considering the burden and inconvenience to the defendant, have only fallen away in the post-2011 Supreme Court jurisprudence. Other values, such as taking seriously the interests of the plaintiff and the interests of the forum state, have been somewhat dormant for much longer.

One need not advocate for a broad general jurisdiction doctrine, or even a return to the pre-*Goodyear* and *Daimler* stasis, in order to take advantage of what notice has to offer. Justice Ginsburg, leaving small spaces in which to advance more capacious definitions of “essentially at home,” stopped short of categorical definitions or categories of entity defendants in both *Goodyear* and *Daimler*. These are the spaces in which core due process concerns like notice might provide the content to broaden the scope of general jurisdiction in limited situations, such as when a company announces the formal existence of a second or even third location for corporate headquarters.<sup>286</sup>

## CONCLUSION

As we have seen, notice had been a long-time procedural law traveler with personal jurisdiction, the two tied together by the mechanics of service of process and by their common and simultaneous elevation to due process doctrines under the Fourteenth (and later Fifth) Amendments to the U.S. Constitution.<sup>287</sup> Despite the role that notice played in shaping personal jurisdiction doctrine and bolstering the Supreme Court’s analysis, it remained an under-recognized and under-theorized aspect of personal jurisdiction doctrine. At one level, this is not surprising. Nothing in this narrative should be mistaken for an argument that commentators have gotten personal jurisdiction jurisprudence “wrong” for the past century; that commentators have somehow overstated the role of forum contacts, purposeful availment and other purposeful con-

---

286. Laura Stevens, Keiki Morris, & Katie Honan, *Amazon Picks New York City, Northern Virginia for its HQ2 Locations*, WALL ST. J. (Nov. 13, 2018, 12:16 AM), <https://www.wsj.com/articles/amazon-chooses-new-york-city-and-northern-virginia-for-additional-headquarters-1542075336>.

287. Andrew D. Bradt, *The Long Arm of Multidistrict Litigation*, 59 WM. & MARY L. REV. 1191-92 (2018) (discussing the relationship between the 5th and 14th Amendments in personal jurisdiction due process analysis).

duct, reasonableness, fairness, convenience, the interests of the plaintiffs, and, of course, the interests of the forum state, most notably via sovereignty and territoriality. I am not arguing that these theories were all a mirage and that I am revealing a hidden-yet-unified grand theory of personal jurisdiction. Rather, what I hope to have shown is that notice has been crucial in doctrinal innovation, yet continually underappreciated by both courts and scholars. A notice-inclusive approach to personal jurisdiction could broaden the doctrine and ground personal jurisdiction in Fourteenth Amendment roots that fit better with the individual liberty core of due process itself.

## AN ORAL HISTORY OF RULE 23\*

AN INTERVIEW WITH PROFESSOR ARTHUR MILLER<sup>±</sup>  
BY PROFESSOR SAMUEL ISSACHAROFF,<sup>±±</sup>  
WITH INTRODUCTION BY PETER ZIMROTH<sup>±±±</sup>



**CENTER ON CIVIL JUSTICE**  
at NYU School of Law

*Professor Arthur Miller was present at the creation and drafting of Rule 23 in 1966. Professor Samuel Issacharoff conducted an interview, during a conference at NYU School of Law commemorating the 50th anniversary of Rule 23 hosted by the Center on Civil Justice, illuminating this history and the ethical issues that were anticipated at the time of drafting.*

---

\* This interview was conducted as part of the Center on Civil Justice at NYU Law School's 2016 Fall Conference titled Rule 23 @ 50: The 50th Anniversary of Rule 23. The Center on Civil Justice at New York University School of Law is dedicated to the study of the civil justice system in the United States and how it can continue to fulfill its purposes. The Center draws on the unmatched strengths of the NYU Law faculty in the fields of procedure and complex litigation, as well as on a Board of Advisers consisting of leading practitioners and judges, to identify the problems that most deserve further investigation and engagement, and to fill a void in scholarly and policy analysis.

<sup>±</sup> Arthur R. Miller is this nation's leading scholar in the field of civil procedure and is co-author with the late Charles Wright of *Federal Practice and Procedure*, the legendary treatise in the field. He is also one of the nation's most distinguished legal scholars in the areas of civil litigation, copyright and unfair competition, and privacy. He is the recipient of numerous awards, including five honorary doctorates, three American Bar Association Gavel Awards, a Special Recognition Gavel Award for promoting public understanding of the law, an appointment by Queen Elizabeth II as the Commander of the Order of the British Empire, and an Emmy for his work on *The Constitution: That Delicate Balance*.

<sup>±±</sup> Samuel Issacharoff is the Reiss Professor of Constitutional Law at New York University School of Law. His research addresses the law of the political process and constitutional law, as well as issues in civil procedure (especially complex litigation and class actions). He is a co-author of the seminal *Law of Democracy* casebook and recently served as the Reporter for the Project on Aggregate Litigation of the American Law Institute. Professor Issacharoff is a Fellow of the American Academy of Arts and Sciences.

<sup>±±±</sup> Peter Zimroth is the director of the Center on Civil Justice and teaches as an adjunct professor at the law school. He serves as the court appointed independent monitor for the NYPD in the stop and frisk matter. He previously served as Corporation Counsel for the City of New York and is a retired partner at Arnold & Porter.

PETER ZIMROTH: Good morning again, everybody. Nice to see everyone back. We have a special treat this morning, with Sam Issacharoff interviewing Arthur Miller. Arthur, as you all know, was there at the creation of Rule 23. Everybody in this room knows Sam and his prowess in the area of civil procedure, so I'm not going to say anything about that. I will just put a plug in that Sam is a scholar of very wide berth, not just in civil procedure, but also in law and democracy. You all know about Sam's work in that subject. He's also a fabulous teacher; he's won teaching awards at every institution at which he's taught, and I can attest personally that he is a very powerful lawyer. I first met Sam when we were on opposite sides of the diet drug litigation, until the case settled, and then we were on the same side. And I can tell you, it's a lot better having Sam on your side than on the other side. And what is there to say about Arthur? Arthur is a legend. Sam, what is there to say about Arthur?

PROFESSOR ARTHUR R. MILLER: [Moving his chair] I'm in the furniture business.

[Laughter]

ZIMROTH: I have a very fond connection with Arthur. We went to the same high school in Brooklyn. There's another Arthur Miller who went there. Isn't that right, Arthur?

MILLER: Yes, 10 years before I arrived, *Death of a Salesman* opened on Broadway, and I discovered we shared a name.

ZIMROTH: Right, but the important Arthur Miller from Abraham Lincoln High School is this one, right here. He is not only a giant in the field, but beloved. And he's won more awards than I can count, including an Emmy for his PBS series. I'm going to stop here because otherwise there won't be any time for the interview. Thanks.

PROFESSOR SAMUEL ISSACHAROFF: So it's me and you again.

MILLER: It's like watching a bad movie twice.

ISSACHAROFF: I think we've done this before.

MILLER: Not this, but others.

ISSACHAROFF: So I want to take us back to 1966, but first, let me take us back just a little bit earlier. One of the things that we tell our students is when you're trying to figure out where a rule, law, or constitutional provision comes from, you should first try to figure out what was wrong with what already existed. This is known as the Rule of Heydon's Case. It's Blackstone. In 1966, there was a package of reforms of all the joinder rules, not just Rule 23, al-

though we will focus on that. It was only 25 years since the Federal Rules had been adopted in 1938. What was the perceived problem?

MILLER: You actually have to go back to '61, which is when the committee made the decision to look at Rules 17 through 25, the joinder rules. There was a sense that the text of the rules at that time was murky, indefinite, unclear, obscure, whatever you want to call it. And that judges and lawyers were having difficulty and creating inconsistencies in the application of some of those rules. Those rules were underused, and it was thought time to rationalize them, to tie them together better than they had been tied in the 30s, and to clarify the text to capture the 25 years of experience, and to insert that experience under the Federal Rules of Civil Procedure into the rules.

ISSACHAROFF: And with Rule 23 in particular, what was the perceived problem? You had this very elaborate language, "spurious," "true," and all these things that sound just bizarre today, but what was the perceived problem in the application?

MILLER: Well, that was it, Sam. Those three categories, "true," "hybrid," and "spurious," just didn't mean anything to most people. It was metaphysical. It was, again, indefinite, unclear. Rule 23 was underused. There were basic questions about how it worked that had produced inconsistency in the few cases that had been decided under Rule 23. So the rulemakers were saying, "Let's put this in plain English." The operative word was "functional." Let's make this functional, and let's capture what we have learned from '38 to the early '60s and insert it, which is what you find in 23(c) and 23(d).

ISSACHAROFF: I want to turn to the animating problem with Rule 23 and its relationship to the civil rights era. But before we get to that, just on a personal note: there were rule revisions and there was a difficult problem of procedural law, so today it seems obvious—you get Arthur Miller, of course. What else are you going to do? But, if I'm not mistaken, you were 28 years old or something, so I'm not even sure you were Arthur Miller back then. How did you get involved? Why you?

MILLER: I was very lucky to have Benjamin Kaplan as my procedure teacher. I fell in love with Ben just watching him, because his face was so expressive and his linguistics were so powerful; he could paint pictures with words. At the beginning of my second year at law school, Ben asked me to be his research assistant that following summer, something no current law student would ever do. To this day I remember, I'm on the phone with Ben, and I'm saying to myself, "I've got 60 years to practice law, one summer to

work for Ben Kaplan. It's a no-brainer." So I go to work for Ben on copyright, which was his other love and became, actually, my first love. I go off to Cleary Gottlieb in New York. Get a chance to teach at Columbia and work on an international procedure project and some related Federal Rules of Civil Procedure.

ISSACHAROFF: Under Jack Weinstein.

MILLER: Under Jack Weinstein, my other mentor, and I think there was a lot of Machiavellian behind-the-scenes stuff. The key was they knew I was close to Ben. Ben treated me like a son, in effect, the summer I worked for him and in my third year. Ben was then the Reporter for the Federal Rules Advisory Committee. And if the Columbia Project was going to make federal rule proposals, what better cannon fodder would there be than sending me up to Cambridge to try and con Ben into bringing our proposals before the Committee—which is exactly what attracted me to the job. Ben was the carrot, plus the possibility of trying to teach.

And, sure enough, that was my mission. I went up to Cambridge with rule revisions on 4, 28, 44, 44.1, and all that stuff that nobody ever thinks about—at least not back then. And a pact with the devil was made. Ben would present these international procedure rules to the Committee if I would help him with the party rules. In other words, I never left my status as a research assistant for Ben. And that enabled me to go to all the Rules Advisory Committee meetings, where they treated me as sort of an assistant to the reporter.

ISSACHAROFF: You were considerably younger than everybody else there.

MILLER: Yes. The closest person in age to me was Charles Alan Wright. I think Charlie was less than 10 years older than I was. I was 27 or 28, I think.

ISSACHAROFF: So, again, before we get to the substance, just to mark the way of marking the change in eras, are we talking about 10 guys in the back room?

MILLER: Well, it wasn't the back room.

[Laughter]

MILLER: The Chair of the Committee was Dean Acheson, a former Secretary of State of renown and the senior partner at Covington & Burling. So it wasn't the back room; it was an elegant conference room at Covington & Burling. But it was closed.

ISSACHAROFF: But it was closed. And there were about 10 of you, right?

MILLER: I'd say 15 or 16.

ISSACHAROFF: 15 or 16.

MILLER: No one else. Sometimes we'd meet at the Supreme Court building, but again in a closed room.

ISSACHAROFF: All men, just to belabor the obvious. It was a different period.

MILLER: Well, you can belabor the obvious even more. It was all white men.

ISSACHAROFF: So you started there. I have heard you say, and David Marcus has written on this, that the first and most central concern on the Rule 23 side was, in some sense, legitimating the court role in the civil rights revolution, that *Brown*<sup>1</sup> was in the room, as it were, for everyone.

MILLER: Absolutely right. It became clear that in the work on Rules 17 through 25, the centerpiece became Rule 23. That got the most attention. And within that centerpiece, the centerpiece was civil rights. Even though *Brown v. Board of Education*, as we know, was not a formal class action, class actions were being employed in the desegregation context—certainly by '62, when the Committee really started to focus on Rule 23—so it was the banner motivation for the revision.

ISSACHAROFF: Again, let's go back to what was the perceived problem, because this is something that we've talked about. And it is something that our students ask every year, when we teach together. Why do you need a (b)(2) class action if you can already get injunctive relief, especially in a world where you have *Parklane Hosiery*<sup>2</sup> and preclusion laws? You can get everything you want out of an individual case. What were you trying to achieve by creating this (b)(2) class?

MILLER: The purpose of the (b)(2) class was simply, as was true of the remainder of the rule, to create a usable vehicle. The (b)(2) was thought necessary because there was no confidence, at that point, that simply because Mrs. Brown got an injunction against discriminatory conduct that a school board or a venal employer would apply that decree to every other member of the affected group. You didn't have *Parklane Hosiery* back then; you had mutuality of estoppel. So, to make sure that there really would be relief to the affected group and to forestall game playing in terms of extending the application of the decree, the (b)(2) was thought necessary. And there also was a recognition that if you put one par-

---

1. See *Brown v. Bd. of Educ.*, 347 U.S. 483 (1954).

2. See *Parklane Hosiery Co. v. Shore*, 439 U.S. 322 (1979).

ent up against the school board, or an organized political entity, it wouldn't work because of the disparity in resources.

ISSACHAROFF: So this is the pick-off problem?

MILLER: It's the pick-off-the-plaintiff problem. It's the uneven playing field. And they saw it in a very preliminary way in the employment context. Remember, this is before the Civil Rights Acts. But you could feel that what happened in *Brown* with regard to school desegregation had legs, and it was going to osmose into other contexts. And you had to give the civil rights group a mobilizing capability, and the vehicle for doing that was the class action.

ISSACHAROFF: That's the story that I've heard over the years, and I want to press about whether that really wraps it up. You were in the shadow, let's say, of *Brown II*. "All deliberate speed."<sup>3</sup> There was a sense that things weren't working. The courts had declared segregated schools unconstitutional, but all schools were still segregated, and the Civil Rights Acts had been stalled in Congress.

One of the questions that I've wondered about is whether you had a sense that this would give more legitimacy to the courts' taking over institutions or to the courts' playing a heavier hand than just simply the declaration. Because it's interesting that even the language of (b)(2) says "declaratory relief and injunctive relief." You had the sense that there were two functions that the courts might be called upon to play.

MILLER: The Committee conversations did not deal directly with that. It came in when (b)(3) became controversial.

ISSACHAROFF: We'll come back to that.

MILLER: There was conversation about the legitimacy of judicial intervention, whether it be in desegregation cases, or employment cases, or in the tort field. Some of the people who did not want the existing rule extended in any sense were arguing that (b)(2) and (b)(3), particularly (b)(3), might create illegitimacy with regard to perceptions about the judiciary. So it was sort of your point flipped 180 degrees.

ISSACHAROFF: Well, there was a discussion previously at this conference about why, then, couldn't you have done this through the standard rules of equity? We had class actions as an equitable device going back, under Steve Yeazell's view, to at least the 15th century, and why not just carry that forward?<sup>4</sup> I've never understood what you were really thinking was going to be the end game.

---

3. *Brown v. Bd. of Educ.*, 349 U.S. 294, 301 (1955).

4. See Stephen C. Yeazell, *Group Litigation and Social Context: Toward a History of*

MILLER: [Sigh] You sound like some of the members of the committee—

[Laughter]

ISSACHAROFF: [Interposing] Old, old.

MILLER: —who are no longer with us.

ISSACHAROFF: Even older.

MILLER: The English Chancery rules were judicial in character and rather vague. Courts in different parts of the country gave them different status and application. So formal equity rules were thought desirable. While not denominated as such, the first rule for class actions was in the equity rules of 1822. There was judicial development prior to that, including the new sets of rules in 1842 and 1912, and the Federal Rules in 1938. In other words, the class action form was there, and people like Ben Kaplan wanted to make it as usable as possible by formal recognition in the Federal Rules.<sup>5</sup>

Not simply for civil rights cases, Ben and others saw it as usable in antitrust and securities contexts as well. Those were the two most obvious purely private law areas, green goods-type cases, but Ben and a few of the others also saw it as valuable in terms of small-claim-large-group cases, the economically unviable cases. Remember, the Warren Court is in the background, that's a "put up or shut up" era in terms of people beginning to demand access to the courts and equal treatment by the justice system. If as a society we really mean what we say about fairness and access to the courts, we damn well better put it in the rules.

ISSACHAROFF: This may be hindsight bias, just reading the past into a state of inevitability, but it's striking that this new rule goes into effect in 1966, and within a year or two the whole dynamic of the Court changes. We have *Swann*<sup>6</sup>; we have *Milliken*.<sup>7</sup> You moved to Boston, to Harvard, in the early '70s, just in time for the Boston anti-busing riots. You have the aggressive desegregation decrees and the rise of the institutional injunction. Can it be that one was just an un-thought-of byproduct of the other? You must have been thinking about what you wanted the courts to be doing. . .

MILLER: There were senses that things in society were developing and changing in different ways, that mass phenomena were increasing in number and increasing in character and dimension.

---

*the Class Action*, 77 COLUM. L. REV. 866, 868-71 (1977); STEPHEN C. YEAZELL, FROM MEDIEVAL GROUP LITIGATION TO THE MODERN CLASS ACTION (1987).

5. See Benjamin Kaplan, *Continuing Work of the Civil Committee: 1966 Amendments of the Federal Rules of Civil Procedure (I)*, 81 HARV. L. REV. 356, 364-66 (1967).

6. See *Swann v. Charlotte-Mecklenburg Bd. of Educ.*, 402 U.S. 1 (1971).

7. See *Milliken v. Bradley*, 418 U.S. 717 (1974).

And at the same time, there was a sense that the Warren Court was ending, and more conservative judicial forces were going to be at play.

But remember that the drafting of Rule 23 starts late in '61 and the Rule gets locked by '63, '64. So this focus on the date of '66 is a little misleading because suddenly the Civil Rights Acts are coming in, and the entrepreneurial bar is beginning to think, "Hey, if it works for civil rights, it will work for us." It's just one of those historical mismatches. If they had drafted the rule in '76 rather than '66 it might've been quite different. The massive growth of class actions was not intended. It was a product of the unforeseen social forces and doctrinal shift that went well beyond civil rights.

ISSACHAROFF: One of the ways of looking back is to say, "How has it become institutionalized, and what would the people who are putting forth the reforms have thought?" Let's take a case that we teach together, perhaps the most important class action case of the last decade, *Brown v. Plata*, which is a case about the deinstitutionalization of many convicted criminals in the California penal system.<sup>8</sup>

What's striking to me is that, while it goes to the Supreme Court as a class action of all prisoners in California, and the Court splits 5–4 on whether the conditions of confinement are cruel and unusual, no member of the Court is the least bit perturbed by the fact that this is a class action. That strikes me as the complete acceptance of the vehicle as the way of organizing this kind of structural relief. Would a case of the scale of *Brown v. Plata*, the entire California penal system, have come as a surprise to the drafters?

MILLER: It would have come as a momentous surprise. I always say Rule 23 is analogous to the amoeba that ate New York. When drafted, it had a modest dimension. There was a sense that in application it would have a limited application. It has proven to have a dimension many times the size of anything conceived of by the people in that room, as bright as they were.

The best proceduralists in the United States were on that Committee. We had every significant academic proceduralist and some of the best district judges. We were constantly invaded at meetings by Judge Albert Maris, professors like Charlie Wright and James William Moore, and others. They were on the Standing Committee, and they would come into our meetings periodically. And we knew the rule had to go through the Standing Committee and the Judicial Conference, so if the Advisory Committee couldn't draft

---

8. See *Brown v. Plata*, 563 U.S. 493 (2011).

something that would get through those entities, it would be a fool's errand.

It was good that they came in with their perceptions, because those three were probably a third of the Standing Committee. But I'm not sure, somebody would have to tell me historically, whether by the early '60s they would have conceived of school bussing class actions, as opposed to just straight desegregation orders, pursuant to the *Brown* mandate, "with all deliberate speed."

I think they would have thought affirmative decrees were possible because certainly the academic members of the Committee, Judge Wyzanski, and Judge Roszel Thomsen of Maryland were old equity people. And, let's face it, the federal courts had been running the meat packing industry since 1920, and were running the music rights industry since the 1950s. So the notion of continuing jurisdiction over a structural decree was not unknown. But I think *Plata* really is a good illustration of what hath God wrought.

ISSACHAROFF: I think we're hearing from you that the Committee members may not have anticipated it, but it wouldn't seem so far beyond what they were unleashing as to cause them great consternation today.

MILLER: In the (b)(2) category.

ISSACHAROFF: Yes.

MILLER: I remember visiting Ben, who became a Justice on the Massachusetts Supreme Court around 1974. I lost him as a colleague four or five years after I arrived at Harvard. Every time I visited him, he was always saying, "What's going on with our Rule 23?" Sometimes he was rather astonished when I told him about some of these rather elaborate applications of the Rule, but he never dissented from the wisdom of the (b)(2). After all, what we have seen are all just extensions of the historic equity jurisdiction of the English and federal courts.

ISSACHAROFF: Okay, so let's turn now to something that has not quite eased so smoothly into our repertoire: the (b)(3). And there are two things that are striking about it. First, it was such a clear innovation. And second, the language of it is rather extravagant. It has these new terms of "manageability," "predominance," and the "desirability" of "concentration," things that have not perhaps panned out all that well in the case law. There are a lot of cases on them, but they have not really acquired a meaning. So let's start with the origin of all this: Where did this ridiculous language come from?

MILLER: [Pause]

ISSACHAROFF: You're going to have to tell it, so you might as well do it now.

[Laughter]

MILLER: Well, only a piece of it. Before I respond to that, is our friend Bob Klonoff here today?

ISSACHAROFF: Yes.

MILLER: Hi, Bob. This came up in the conference yesterday. The Committee in the '60s was cut from the same cloth, so to speak, as the Committee of the '30s led by Charlie Clark, then of Yale Law School, who was the chief architect of the Rules.

ISSACHAROFF: And whose assistant Charles Alan Wright was your co-author. So there's continuity in this.

MILLER: Yes, yes, it's all connected, somehow. One of the basic drafting philosophies of the Committee of the '30s and the Committee of the early '60s was open textured rules: simple language, short provisions, general statements, leave it to the judges, they can fill in the gaps. If you ever take out a copy of the '38 Rules, and even a copy of the rules following the '66 revision, and lay it next to the current Rules, it's like looking at the Constitution to the right and then looking at the tax code to the left.

[Laughter]

MILLER: The style of drafting was entirely different. Back then, words were chosen that carried out this notion of generality, because, as they would frequently say, we can't see what's around the corner.

ISSACHAROFF: But it's striking that if you compare (b)(2) to (b)(3), in (b)(2) you have a very clean rule that just says for purposes of declaratory relief or injunctions; and (b)(3), even in the '66 version, starts looking a lot more elaborate. I've known you for decades, and I know your preference for clean language. I don't recognize in (b)(3) your authorship, except for the fact that you did it, which we'll get to in a second.

MILLER: Even in (b)(2), you get generality. Look at some of the passages: "Course of conduct." "Affecting the members of the class." "Giving rise to final or corresponding declaratory relief." Even in (b)(2), you have a sense that they are not writing the tax code. The original text of (b)(3) was very general as well.

ISSACHAROFF: And so let's talk about the drafting, and then I want to talk about why you had to write so many pieces into it. So you were starting, there was some sense, as you said a minute ago, that the small value economic claims were going to come in some-

how. You were going to draft (b)(3) to accommodate that. Tell us the story of the actual text drafting.

MILLER: Oh, the text drafting went over a couple of years.

ISSACHAROFF: Right, but the final—

MILLER: [Interposing] I have a feeling you're referring to a particular piece.

ISSACHAROFF: You know what I'm referring to, and you're going to make me ask it 10 times.

MILLER: No, no, no.

ISSACHAROFF: As you always do in class—

[Laughter]

ISSACHAROFF: —but I'm going to get it out of you as I always do in class.

MILLER: It's just that my story is so idiosyncratic, and in its own way, in my head, it sounds too damn egotistical. You have to understand, when the Committee started on the text of (b)(3), at least Ben and some of the others conceived of it as an elaborate joinder device. Remember, the rule does not say who was bound by a class action. It doesn't, and it couldn't. That would be a violation of the Rules Enabling Act,<sup>9</sup> they would say, and I would agree.

ISSACHAROFF: Really?

MILLER: Yeah.

ISSACHAROFF: OK.

MILLER: I don't think a federal rule can provide preclusion, but leave that to one side. It's a matter of substance. It's not—

ISSACHAROFF: [Interposing] I've seen a lot of court opinions—

[Laughter]

MILLER: Look, I don't have to teach Procedure until Monday, so leave me alone.

[Laughter]

MILLER: It was conceived of as an elaborate joinder device that would aid in the fraud cases, the antitrust cases, and the negative value cases, and there's almost a touch of consumerism in that. I don't think they thought about environment. It was just a way of aggregating people who believed they had been injured by common conduct. Then a verbal fistfight broke out within the Committee. Some people thought that (b)(3) was very controversial. Remember, this is 1961-63, no one is thinking about tobacco or asbestos or pharmaceuticals. And some of the members of the

---

9. 28 U.S.C. § 2072 (1982).

committee, one in particular quite stridently and inexhaustibly, wanted no (b)(3). Indeed, he didn't even want a (b)(2).

ISSACHAROFF: This would be John Frank, right?

MILLER: You said it.

ISSACHAROFF: Actually, Steve Burbank did, but. . .

MILLER: OK. John was a wonderful lawyer. He started his life on the Yale faculty, migrated because of some health and faculty politics issues to Phoenix, and represented some of the most significant companies in the South.

ISSACHAROFF: And became one of the first heads of the Civil Rights Division. He had an extraordinary career.

MILLER: But he sure as hell didn't want a (b)(3). He gave up fighting the (b)(2), but at the beginning of the Committee's deliberations on Rule 23 he thought everything could be covered by (b)(1). And people like Ben and Judge Wyzanski, Charlie Wright was a little ambivalent initially, but Charlie Joiner and Dave Louisell, they wanted something that had the malleability to advance the class action as needed. And Ben would say that without (b)(3), in particular, the rule would be retrogressive. It would abate aggregate litigation, and the class action would not do the job that was necessary to permit small claims to be brought.

The ultimate compromise, triggered by a Wyzanski comment, was to build safeguards into (b)(3), because John Frank would always say that he was afraid of misbehavior in the (b)(3) context. He was afraid of misbehavior that would have an effect on absentees. There is a notion that is still prevalent in many nations that before you affect an individual, the individual has to be a participant in the action. It's his or her natural right to be there, and a court cannot foreclose that right.

ISSACHAROFF: Something like what the European courts have held, that there is a fundamental human right of not being bound without your affirmative consent.

MILLER: Exactly, which gives American class action judgments difficulties in many parts of the world. The most evil thing I could say about John is that I don't think John left his clients out of the discussion. John kept masking it, arguing that renegade plaintiff lawyers will start these class actions, get judgments, and hundreds of people would be bound without their participation or even their knowledge. And that's the way the debate between Ben and John got crystallized, and then the disagreement got worse and worse over meetings with a lot of back channeling.

And Judge Wyzanski just said, "Let's protect those absentees." So you end up with predominance, in other words, you've got to get

a lot of judicial bang for the buck before you certify under (b)(3). This has got to be a true efficiency economy win before you bind people with the (b)(3). It's got to be superior. God knows what that means, "superior." But it was understood to be protective.

ISSACHAROFF: I never understood it either. But I'm glad to know—

MILLER: Words like "predominance" and "superiority" were like silly putty that could be molded in any way by a judge in a particular context. And then the key became the (c)(2) notice—giving class members' individual notice. And I remember saying to Ben, who had taught me the *Mullane* case,<sup>10</sup> "Why don't we just write *Mullane* into the rule," which we did. And the key was thought to be the class members opt-out right because that would guarantee that each class member had a choice to advance his or her right on an individual basis.

ISSACHAROFF: Right, but where'd that come from? Because that wasn't present in any rule before.

MILLER: No, that was a construct. I forget who proposed it within the committee. . .

ISSACHAROFF: Wyzanski.

MILLER: I was going to guess that it was Wyzanski. He was incredibly smart and wise. So those four things have to be written in. That encumbered the text of (b)(3) and other parts of the Rule in and of itself, and then there was a feeling that in using predominance and superiority, you have to give that some texture. And that's where the Four Horsemen of the Apocalypse are laid out in the Rule as factors to be considered by the court. The most significant in retrospect has been "manageability." Even "manageability" is a soft word that means whatever a judge might want it to mean.

ISSACHAROFF: But it's a concept which we take for granted, particularly after the '83 reforms of the federal rules, for which you were then the Reporter to the Advisory Committee. But the word comes out of nowhere at the time in terms of the Anglo-American conception of the judge.

MILLER: That was the brilliance of Ben, being able to find words to meet the situation, without over-crystallizing. It's no different than the decision made by the original Rules Committee not to use words like "fact," "conclusion," or "cause of action." They canonically banned those words because they had too much baggage on them. The 1960's Committee, in doing the (b)(3) thing, was grasping for words that didn't have baggage.

---

10. See *Mullane v. Cent. Hanover Bank & Tr. Co.*, 339 U.S. 306 (1950).

ISSACHAROFF: Right, but “manageability” not only doesn’t have baggage, the concept of creating an administrative constraint on courts is recognition of an inquisitorial court-like power that is not easily found within American jurisprudence at that point.

MILLER: It simply reflected what the Committee saw, and that is the growth of the bigness of litigation. Not bigness as we now know, but bigness as they perceived it at that time. Let’s face it, a Section 10b-5 stock case, or a price fixing case, can be a big case. And they also saw the birth of judicial management. That’s what they saw, and they understood that the court should have discretion to say, “I can’t manage that” for hundreds or possibly thousands of claimants.

ISSACHAROFF: They didn’t see Judge Barbier trying to resolve 450,000 claims in a two-year period in the Deepwater Horizon litigation.

MILLER: They did not. So the drafting of (b)(3) took more time than anything else and led to the inclusion of a lot of verbiage.

ISSACHAROFF: So let’s talk about the drafting. As I recall, you were a young conscript or something.

MILLER: Well, conscript is the right word. I was on active duty with the United States Army. The big Committee meeting was coming up, and, I’ll use the word, Ben was “panicked.” Ben had a certain fragility to him which always amused me, but it was quite human. He said, “You’ve got to be at the meeting; you’ve got to be!” I said, “I’m on active duty with the 77th Infantry Division in Camp Drum, New York.” He goes to Secretary of State Acheson—

ISSACHAROFF: Acheson knew somebody in the military from his days as Secretary of State, I take it?

[Laughter]

MILLER: No, he had a better idea. Acheson knew Chief Justice Earl Warren. The Chief Justice writes a letter to the commanding General of the 77th Infantry Division. The letter actually says, “This man is needed on the nation’s business.”

[Laughter]

MILLER: I kid you not.

[Laughter]

MILLER: So they give me a pass. Ben and his wife, Felicia, had a wonderful home on Martha’s Vineyard. So it’s decided that they’ll pick me up at the Boston airport, and we’ll go to Martha’s Vineyard and work on Rule 23. So we’re now in the bowels of the Martha’s Vineyard Ferry. Now for all you hotshot technologists, I’m sitting in the back seat with a portable typewriter. It’s not even elec-

tric. Ben, who didn't drive, is in the front; Felicia is driving. And we're going back and forth on the drafting, and that's literally when some of these words get incorporated in what we now see in (b) (3).

I'm clacking away in the back seat, and we're going back and forth on words. In the middle of this, the woman in the car next to us rolls down her window, reaches over, and taps Felicia's window. Felicia rolls the window down, and the voice from the other car said, "Are we sinking? Do you hear that sound?" And it's the clacking of the typewriter. Felicia just points at me, and the woman is very relieved.

[Laughter]

MILLER: For me, that is an indelible memory.

ISSACHAROFF: So if the seas had been calmer, we might have gotten a better rule that—

[Laughter]

ISSACHAROFF: —that day?

MILLER: Yes, or a shorter rule.

[Laughter]

ISSACHAROFF: A shorter rule.

[Laughter]

MILLER: Yeah, I mean (b) (3) was the clumsiest, biggest, most textualized provision.

ISSACHAROFF: Right, so that's the part that strikes me, because I've known your writing for years and years, and we've taught together. It's not like you; it's got too many moving pieces, subordinate clauses. It's not clean. You don't see the flow from beginning to end. So it's written defensively, which is not your style.

MILLER: I think that's right, Sam. I think the fate of the rule was so indefinite at that point, because it wasn't just John Frank; even Charlie Wright initially had doubts about (b) (3), and a couple of the practicing lawyers had doubts. There was a concern that industrial forces would block the rule, unless somehow it was made acceptable. So I think (b) (3) is written recognizing that we had to put procedural safeguards in there, and some texture in there, and some committee members wanted more rather than less. That's what happens when you deal with Committees.

ISSACHAROFF: So the compromise was you had a lot of pieces in there that at least would force judges through a checklist of multiple safeguards.

MILLER: Yes, as nasty as that sounds, Rule 23(b) (3) and the other provisions relating to it are in effect a checklist, which was not characteristic of the Rules up to that point.

ISSACHAROFF: Right, or of you.

MILLER: Forgive me, Bob [Klonoff, who is on the current Committee], the damn rules are now cluttered with checklists, so I'm an unindicted co-conspirator in the development of that phenomenon.

ISSACHAROFF: Well, we have Elizabeth Cabraser here, also from the current Rules Committee, so we can—

MILLER: [Interposing] Oh, I'd rather pick on Bob.

[Laughter]

ISSACHAROFF: So this is an innovation. It's a controversial one, so you package it with *Mullane*, you package it with this new opt-out right. What was the paradigmatic case? What did you have in mind that you really wanted to facilitate? What did Ben think that he really wanted to facilitate?

MILLER: Well, civil rights was a given.

ISSACHAROFF: Yes, that's on the (b)(2) side. You haven't seen the damages cases yet for civil rights.

MILLER: That is right, and, boy, if you go back to the text of (b)(2), there's a lot of truth in what Justice Scalia said.<sup>11</sup> I hate to admit it, but the difference between a negative and an affirmative injunction is just not in there. And the (b)(3), they saw the antitrust and the securities cases, and they clearly saw, or at least Ben clearly saw, the negative value cases. They did not see environmentalism, or product safety, or privacy. They didn't really see consumerism. I think anyone who has read any of transcripts of the meetings realizes that they never talk about subject matter jurisdiction issues and that they never deal with the question of the diversity case and whether you could aggregate small claims to meet the more than \$75,000 requirement. But, of course, that is beyond the rulemaking power.

ISSACHAROFF: Or \$10,000 at the time.

MILLER: \$10,000 at the time, right. Beyond that, odds and ends.

ISSACHAROFF: Let's take the antitrust case as the example. What was the concern? Because in an antitrust case you have attorney's fees in the statute; you have treble damages; you have aggregation facilitation devices already built into it.

MILLER: There, again, they were concerned about small claim antitrust cases, wholesalers, for example, or even small retailers who individually could not proceed against a mega corp. And, let's face

---

11. See *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 360-61 (2011).

it, they were proceeding under the assumption that whatever they put into Rule 23 would in fact have bilateral binding effect, although they never say so in the rule. So they saw the use of the class action as gaining that type of efficiency and economy for the system and for defendants as well as plaintiffs. The rule represents an affirmative validation of multiparty joinder. But Johnny Frank saw it as an attractive nuisance that might encourage too many cases.

ISSACHAROFF: So could we use some modern language and say that the primary concern was effective vindication of the substantive law in areas like antitrust, with the small players there?

MILLER: It's '61 to '64, effectively, that we are talking about. The bar is not polarized the way it is today. You go through the list of Committee Members. You don't see a plaintiffs' lawyer in there, because nobody thought about the dichotomy between plaintiffs' lawyers and defense lawyers. There were large-firm practicing lawyers; there were judges; there was Archie Cox; there was Abe Chayes representing other branches of the government. They wanted it to be effective and final, leaving finality to the growth of preclusion law. But it was not polarized in the plaintiff/defendant sense. That wasn't one of the dynamics in that room at that time. And very few comments came in from the outside world.

ISSACHAROFF: But it's interesting—if you think about *Italian Colors*,<sup>12</sup> for example, and Scalia's opinion there, that opinion is so directly on point to what you're describing as the central concern at the time, and now the substitution of individual arbitration really is in direct tension with the objectives in 1966.

MILLER: Absolutely. Arbitration was not on their minds. If you asked me, I think some of them would be horrified by *Concepcion*<sup>13</sup> and *Italian Colors*<sup>14</sup> as a fundamental imposition on the judicial system and a transmogrification of the Arbitration Act that was not meant for employment or consumer cases.

ISSACHAROFF: And, as we know, we live in a different world now, where there is an established plaintiffs' bar. There are securities lawyers or antitrust lawyers. There are, increasingly, after CAFA,<sup>15</sup> consumer class actions being brought. What would have shocked them, besides the sheer entrepreneurialism of it all? What would have shocked the Committee?

---

12. See *Am. Express Co. v. Italian Colors Rest.*, 570 U.S. 228 (2013).

13. See *AT&T Mobility LLC v. Concepcion*, 563 U.S. 333 (2011).

14. See *American Express v. Italian Colors*, 570 U.S. 228.

15. Class Action Fairness Act of 2005, Pub. L. No. 109-2 (2005).

MILLER: Scale. There was a secondary verbal fistfight about mass accident cases. They would be shocked by the scale of today's cases. The notion of the *Castano* national tobacco case<sup>16</sup> would have boggled their minds. *Wal-Mart* would have been beyond their conception in those days.<sup>17</sup> Many people in that room didn't think they were writing a rule that really had significant application in the tort environment, yet they talked about fraud actions, which is ironic because the key element of fraud is reliance which is individualized. But they talked as if fraud cases would be natural (b)(3)s. But that was the outer perimeter of their vision in terms of the rule's scope.

ISSACHAROFF: Right, so let's turn to that, because I think that there is a tendency to read from the present debates into the past. We've put a lot of freight on the predominance issue, and so the reason that mass torts don't fall into class actions easily is generally thought to be that they can't get over the predominance hurdle because there are too many individual issues. As you say, fraud, until we have presumptions of reliance develop as a substantive matter, would also not fall into easy predominance of a common inquiry. But they had the idea of mass torts, and so there's this curious line in the Advisory Committee notes, "of course, nothing herein shall affect tort cases."

MILLER: No, that's not the line.

ISSACHAROFF: OK. I'm paraphrasing, but go ahead.

MILLER: It's just mass accidents not usually being appropriate for certification, something like that. There was a controversy about the language of that passage in the Advisory Committee notes.

ISSACHAROFF: So why don't you tell us about that, because it comes out of nowhere, and there's no other part of the Committee notes that seems to say, "Oh yeah, we're holding off on this particular substantive area here."

MILLER: Even though the meetings were held in closed rooms, elements of what was going on would get out. And there was concern about opposition to the rule, any rule, higher up the rule-making channel. It was assumed that various influences might play on the Standing Committee or Judicial Conference, let alone the Congress. John Frank constantly threatened or voiced the opinion that unless mass torts, or mass accidents as they were called, were excluded from the rule, there would be opposition to the pro-

---

16. See *Castano v. Am. Tobacco Co.*, 84 F.3d 734 (5th Cir. 1996).

17. See *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338 (2011).

mulgation of the rule from the insurance industry and other corporate interests. That was not said in open meeting; it was said several times over lunch or dinner or the telephone.

And there was a fear that that could kill everything they were doing. There was even a representation that unless mass accidents were excluded, certain members of the Supreme Court would oppose the rule. The Committee quite properly refused to put anything in the rule. There was a stronger sense of the trans-substantivity of the Federal Rules then than now. So they decided to put it in the Advisory Committee Notes. I do not know, to this day, and no one will ever know I guess, whether these representations about opposition were true or false, or whether conversations were ever held between John Frank and a couple of those he mentioned. I do not know, but it's quite clear those representations had an impact on the Committee.

ISSACHAROFF: So let's go back to that period. We did not have the mass torts that we have now, but you had airplane crashes back then, more than now. You have, in the same period, coincidence that long before '66 the rule was locked in, but a year later you have the *Tashire v. State Farm* case<sup>18</sup> where the Court intriguingly at the end of the opinion says, "Yes, maybe this should be under interpleader, but we think not. We've got to do something about mass torts, but we don't think interpleader is the right vehicle." So people were thinking about where these cases were supposed to end up, and what was the thinking? Where should they be? If they shouldn't be Rule 23, according to some, because you bind, what was the other view? Where should they be? Clearly with an airplane crash, you have to resolve a common issue, right?

MILLER: Yes, but the air crash or the mass accident impacts almost by definition human beings in a direct, physical way. That was the strongest core of the right of individuals to pursue their grievances on their own. They shouldn't be co-opted into a nameless, faceless mob of people. That was very powerful, even though that notion was starting to tail off. Also important was the notion that those cases can all be handled by contingent fee, that they were simple and would take care of themselves. Or, if you wanted to be cruel, there may have been people in the room who really just didn't give a damn about those cases.

Most of the lawyers on the Committee, I'd have to go back and look at the list, were commercial lawyers. They were litigators, but

---

18. See *State Farm Fire & Cas. Co. v. Tashire*, 386 U.S. 523 (1967).

primarily in the commercial context. But the mass tort issue sure as hell scared people off for a while.

In another anecdote, I got a call one day from the judge handling the Kansas City Hyatt skywalk collapse. He calls me up, and he says, "I've got a fight between some East Coast lawyers who want to pursue it on a class basis and a bunch of Missouri and Kansas lawyers who want to pursue it on an individualized basis. I've read the Committee note, but this sounds like a class action to me." I opined that it was a perfect class action. And if you stop and think about it, you've got a lock on predominance and the other class action requirements. In a way it is very much like a crash case. So I contributed to that inappropriate *ex parte* communication and encouraged his certifying it as a class action. He did, only to be reversed on a questionable basis.

ISSACHAROFF: Right, but there was that, and there was the Puerto Rico hotel fire. The thing that strikes me about the airplane crash or the skywalk case is the commonality of the issues. In the skywalk case, the question was whether the bolts were put in this way or that way, and that issue doesn't depend on who the plaintiff is. That's a fact that has to be determined. And in the airplane crashes, was it pilot error or a malfunction of the system? Was it a design defect? And courts were confronting that problem of these kinds of common issues.

*Tashire* talks about that, that the ultimate issue is who is negligent, the pickup truck driver or Greyhound?<sup>19</sup> And there's nothing that an individual is going to add to that inquiry. There may be damages down the road, but those central issues are common across plaintiffs. So what was going to happen with those cases? They were just going to work themselves out?

MILLER: They'd work themselves out. The contingent fee bar, particularly in the air context—that was a defined, powerful bar—they could take care of themselves. And although there were no plaintiffs' firms of the style of Elizabeth Cabraser's firm back then, there were very powerful individuals and small plaintiffs' groups, so they could take care of themselves. These cases were not negative value cases. I have a feeling that—I can't remember—the Advisory Committee Note was phrased as cautionary and not in absolute terms, and it left some room for a judge to certify a single-event mass accident case.

ISSACHAROFF: Let's jump to today, and the question is what would be most shocking. Let's take three of the cases that are float-

---

19. *Id.* at 533-37.

ing around now as class actions. You have *BP*, which is scale beyond anybody's wildest imagination. You have *Volkswagen* also, but *BP* is tort-like activity of an economic consequence primarily. *VW* is tort-like activity, but a purely economic and environmental consequence. And then you have *NFL*, which is tort-like activity of a tort-like consequence. Which of these would be unacceptable? Acceptable? What would be shocking? Or they all would be?

MILLER: Well, leave scale to one side because the scale of all three cases would shock them, all of them. But they'd get over that very quickly, I think, because it's just scale. Throw another body in the class—who cares?

ISSACHAROFF: Unless it's your body.

[Laughter]

MILLER: I wouldn't mind having my body thrown into a class—but I think *VW* is the easiest case, since it's purely economic.

ISSACHAROFF: And it's a government agency. All those complications you could extrapolate up and scale up to that?

MILLER: You could. That's a scale question, nothing more, because the case is a single issue. Liability is acknowledged, so it is only about damages. I think *BP* is next in the scale. There are tremendous variations following the disaster that occur over time. The spill lasted, I think I heard Judge Barbier say, 87 days or something like that. So you're moving to that predominance/no predominance line in that case. So I think that one might go either way.

*NFL*, I think, is the hardest. I don't think you would have gotten the class certified. Indeed that's why I wrote a couple of op-eds saying the settlement should be accepted by the class members. In that case the circumstances of each individual varied; it's physical injury; it's happening over time; it's a nascent science with tremendous uncertainty reminiscent of the Agent Orange case. And that's why Jack Weinstein granted summary judgment against the opt-outs in Agent Orange following the settlement.

So that's the way I would scale those three. I think the committee really would have had trouble with the *NFL* case.

ISSACHAROFF: So let me push on that a little bit, because you mentioned the opt-out was kind of added on as a notion of individual protection. And one of the things that we're seeing is many more media of participation by class members. So it doesn't look so much like you're bound in abstention. You now get active participation through Facebook, through Twitter accounts, out of everything that is done in these cases. Would they anticipate that the

class action could be a participatory event, as it were? As opposed to simply the idea that there's a lead plaintiff and a lead lawyer, and everybody else is passive and is just sitting back being represented? I mean, they couldn't have anticipated the technology, obviously. . .

MILLER: The rule does contemplate a right of intervention. Whether they actually thought people would intervene, I have my doubts. I doubt that they saw the class action as a series of town meetings. But they wanted to create the possibility of participation.

ISSACHAROFF: Let me close out by asking about three discreet topics that are of significance today and are developments as the rule has been implemented. The first is, we've talked about before, you were the ALI reporter for the Complex Litigation Project, which mostly tried to get cases better teed up across federal-state divides for trial. And then I was a reporter for the ALI on the Aggregate Litigation Project. The big change between the two was that you were heavily trial-focused, and our project was largely designed to legitimate and push further on the settlement process, acknowledging what class actions had become: a very big bill of peace and the mechanism for the resolution of very complicated, large-harm cases. And so what was settlement? Was there any discussion of a "settlement possibility"? A settlement class? This institutionalized bill of peace?

MILLER: God, no, no, no, no. I mean, one might find references to settlement with the judge's power to settle. But that was not part of the discussion. I think they were conceiving the rule as a trial-ready rule. Remember, the word "settlement" did not exist in the Federal Rules of Civil Procedure until 1983, when as the Reporter to the Advisory Committee I put it in the redraft of Rule 16. Settlement back in the early '60s was thought as being a non-judicial function. Judges simply did not involve themselves with settlement. Indeed, as late as the selling of the '83 changes, I would wander around to judicial conferences and meet lots of judges who would say, "I will never sully my hands by being involved in settlement." So that was not in the thinking of the Advisory Committee in the early 1960s.

ISSACHAROFF: And now they're sullied.

MILLER: That's one of the biggest changes in the latter half of the 20th century, the shift from adjudication to judicial management.

ISSACHAROFF: Right, and the class action has been almost the poster child for this, because few class actions go to trial. Few of anything go to trial. But here you have to create a decree at the end, because the preclusive effect is binding on absent people, and

so the judicial involvement is critical, and the judicial role in settlement is subject to exacting appellate scrutiny the way it isn't in other areas of law. So this is a real change in the role of the courts as a result of the (b)(3). The opt-out rights, and now we have the whole 23(e) settlement apparatus—

MILLER: [Interposing] No, but they did put it into 23(e).

ISSACHAROFF: Right. And so, what were they thinking about that? Because there's language there, right from the very beginning, that's a placeholder for this, if you will.

MILLER: No, they didn't perceive that in this context judicial involvement was necessary. I don't know how many people read Judith [Resnik]'s draft, but she puts a lot of weight on the *Mullane* case. And there's a touch of judicial involvement in the *Mullane* case in the appointment of the representatives for the principal beneficiaries and the interest beneficiaries. And 23(e), at the time of the 1966 revision, speaks of dismissal or compromise, not settlement.

ISSACHAROFF: So let me ask you about another drafting quirk that appeared from the beginning. In the last five years or so you had a number of appellate courts that have promoted the use of issue classes. And you've had this provision in 23(c)(4) that sat there for decades, as best I can tell, never being used. And now, all of a sudden, it has sprung to life. It just lay dormant, as it were. What were you thinking when that went in? And why was it in (c) and not a form of class action in (b)? What was it supposed to do?

MILLER: I hate to say this—

[Laughter]

MILLER: —but although I absolutely applaud what's been done in the Second and Seventh, and a little bit in the Ninth Circuit—

ISSACHAROFF: [Interposing] And the Sixth, and a little in the Fifth, and I think Third a little bit. It's spreading.

MILLER: I think the so-called issue class is wonderful. I don't think it would ever have passed a Justice Scalia inquiry, though, because it is in (c). It's in (c). Let me back up a bit. I don't think the provision ever was really discussed in open meetings.

ISSACHAROFF: Where'd it come from?

MILLER: It came from the notion, I believe, this is my best recollection, that they wanted to give the judge power to certify less than the full class. Remember subdivision (c) is a garbage can filled with procedural devices empowering the district judge to do various things.

ISSACHAROFF: But all rooted in some equitable notions.

PROFESSOR STEPHEN BURBANK: [From off stage] That's an interesting view of your field.

[Laughter]

MILLER: Well, that's what my field of sports law is—it's a garbage can. You find interesting things in garbage cans, Steve.

[Laughter]

MILLER: You should try some rummaging.

[Laughter]

MILLER: No, I think what they had in mind was a case certified under (b), with a decision by the judge to treat less than all of it, maybe a single issue of it, on a class wide basis, and then issue a decree or a judgment that applied class wide on that issue. I think (c)(4) was in early drafts, and it remained there before all of the bric-a-brac for (b)(3) cases was put in.

ISSACHAROFF: But where did this come from? It doesn't seem to correspond to anything that came before, except vague notions of equity.

MILLER: Yes. Yes. It's what I said almost at the beginning of this conversation. There was tremendous respect among Committee members for the equitable powers of a federal judge and the flexibility of those powers. And (c) is designed to encapsulate some of them and provide a bit of a procedural road map.

ISSACHAROFF: Let me ask you—we're at the end now—I want to ask you about something that came up, and we had an extraordinary last panel yesterday, a discussion of the way the law is being pushed in the mass tort context. There's a term that's used now: these "epidemiological cases," that is cases that you can't prove on the individual level; you need to go into bigger aggregations. I recall that when I started the Aggregate Litigation Project, at the very first advisers' meeting, which you attended, Jack Weinstein in his inimitable way gets up and says, "You're all wasting time. This is all useless. The real question is how we're going to get the procedural rules to accept the nature of proof, that the proof nowadays is all epidemiological, all statistical." This came up in the *Tyson Foods* case in Justice Kennedy's opinion.<sup>20</sup> Was this on the horizon, now that you were unleashing many more of these cases in different styles? Was the sense that the proof itself, the nature of liability, was going to start turning on the aggregation proper?

---

20. See *Tyson Foods, Inc. v. Bouaphakeo*, 136 S. Ct. 1036 (2016).

MILLER: No, no. Thinking about the people in that room, I don't think even Ben, who was among the most avantgarde types in the room, was thinking about that. Maybe Judge Wyzanski might've been thinking about that, but a Charlie Wright would not have been thinking about that, a Roszel Thomsen would not have been thinking about that. You've got to go back to '61 to '64. That would have been revolutionary. These were people truly committed to the adversary system, and if you go through 23 as promulgated in '66, it reflects a commitment to the adversary process. It does not reflect a commitment to judicial management, and it does not reflect commitment to science or epidemiology. I don't think that was on the mind of anybody in that room.

ISSACHAROFF: It's just fascinating, Arthur, because I understand that you can't project the future back to the past, but you're sitting there in the early '60s, and realistically all hell is about to break loose. You're going to get the injunction cases. You're going to get federal judges being demonized in community after community. You're going to have to get institutional responses from the judiciary. You're going to get Supreme Court cases deciding whether injunctive powers can allow you to replace the elected officials of Yonkers.

MILLER: Everybody was pretty dumb, huh?

ISSACHAROFF: No, it wasn't dumb, it was just—

MILLER: [Interposing] Shortsighted, huh?

ISSACHAROFF: Eh, maybe, maybe.

MILLER: Lack of vision, like most of us the night before the presidential election.

[Laughter]

ISSACHAROFF: Yes, perhaps. But it's a snapshot on a moment where the world was changing, and you unleashed powerful reform impulses, and it was a rather conservative group of people that did this at the end of the day. As you said, Dean Acheson is about as American establishment as you can get in that period.

MILLER: Oh, yes, yes, by any contemporary standard, the Committee members were Stonehenge types. And Acheson gave Ben tremendous freedom. And a lot of this was Ben. A lot of it was Ben supported in significant part by Wyzanski. Look, there was talk yesterday, and I think Steve Burbank led a little bit of discussion, about the fact that there was this gang in Chicago working on what became the MDL process. And, God, the biggest non-sight was MDL and the fact that the *Manual for Complex Litigation* is in the early stages of being put together. Why? It's '63 or '64, and the

Committee pressured Ben and Al Sacks to go talk to Phil Neal and the Chicago gang. They didn't want to go, but they felt that there was enough sentiment in the room that they were obliged to go. But that was a nonstarter. That was '63. The MDL statute doesn't come in until '68. So when Ben and Al come back from the meeting, it's obvious. Who knows what's going to happen with that Chicago thing or when. Will it ever be enacted? It sounds like a teeny weeny, and we have our own agenda and timetable.

BURBANK: Yeah, but they did put superiority in the rule as a result of that meeting.

MILLER: Yeah, and that's it. Because the Committee had an agenda, and it had a timetable, and it couldn't sit around, hold up all of the joinder rule revisions, and wait to see whether or not the MDL would be enacted by Congress. In fact, it wasn't for four more years after the rule was handed off by the Committee.

ISSACHAROFF: So, Arthur, we're out of time, and I just want to finish on one question, which is: We both made our careers as teachers and/or students. Part of this that keeps coming up again and again is the extraordinary shaping influence on your entire career of the fortuity of your relationship with Ben Kaplan. It really, at every level, is your introduction into this world as a student, and then as a research assistant, and then into the rules process and legal reform. It's a career-shaping experience. Is it possible to imagine your career differently?

MILLER: Yeah, I could still be in the Army.

[Laughter]

ISSACHAROFF: Actually, I'm sure that's not true. But anyway. . .

[Laughter]

MILLER: I always say that Ben was my mentor, as I would say Jack Weinstein also was a mentor. And Ben would always say, "I claim no responsibility for you." No, look, I spent almost my entire teaching career in civil procedure and copyright, which is exactly what Ben taught. He was an Advisory Committee Reporter; I was a Reporter. We both had a love of process. He probably did not like the fact that I went into television. He probably thought I should have ended up as a judge. But Ben and Jack shaped my life. And I'm very, very, very grateful to both of them.

ISSACHAROFF: Well, on that note, Arthur, as always, it's been a pleasure.

[Applause]

[END]

# PREDICTIVE POLICING: THE ARGUMENT FOR PUBLIC TRANSPARENCY

ERIK BAKKE\*

Introduction .....	132
I. Predictive Policing.....	134
A. Summary of Predictive Policing Technology .....	134
B. Benefits of Predictive Policing .....	137
C. Hazards of Predictive Policing .....	139
D. How Law Enforcement Obtains the Technology .	140
II. Transparency: The Current Arguments .....	142
A. Routes to Information .....	142
B. Arguments for Transparency.....	144
III. Additional Arguments for Transparency Specific to Predictive Policing.....	147
A. Transparency as an Additional Deterrent .....	147
1. Improved Effectiveness through Broader Trust in Law Enforcement .....	147
2. Improved Effectiveness through a Deterrence Effect.....	149
B. Transparency as an Accountability Gap-Filler ....	151
C. Transparency as Guarantee of Fairness .....	153
IV. Restrictions on Public Access to Predictive Policing Information .....	155
A. Open Records Law .....	155
B. The Business Information Exception.....	158
1. The Structure and Applications of the Exception .....	158
2. The Rationale for Extending the Business Information Exception to Predictive Policing .....	160
C. The Law Enforcement Exception .....	163
1. The Exception in General .....	163
2. The Exception's Applicability to Predictive Policing .....	165
3. Reasonably Segregable Exception .....	168

---

\* J.D. 2018, New York University School of Law. Special thanks to Professors Barry Friedman and Maria Ponomarenko, as well as the Policing Project. Additional thanks to Natalie C. Gow and the entire editorial staff at the *Annual Survey of American Law* for all of their input and efforts preparing this Note for publication.

D. The Unlikelihood of Accessing Predictive Policing Information ..... 169  
 Conclusion ..... 171

INTRODUCTION

At 4 p.m. on a Tuesday in Santa Cruz, the officer’s squad car radio reports a robbery. The officer puts his car in gear and radios back his pursuit of the robber.<sup>1</sup> After making a few turns, he catches up to the suspect running from the scene just moments after receiving the initial report. The officer’s success is not due to luck, but to the Santa Cruz Police Department’s (SCPD) predictive policing system. Santa Cruz receives a significant amount of tourism, and with that tourism comes a substantial amount of street crime. Fast response times are crucial for tracking down suspects. Looking to improve response times, the SCPD contracted with PredPol, a private company, to purchase predictive policing technology.

PredPol’s technology, like competing predictive policing tools, uses computing and analytics to make predictions about where and when crimes are most likely to occur. The PredPol system estimates the likelihood of four crimes (auto theft, vehicle burglary, burglary, and gang-related activity) each day for fifteen zones. Each zone covers a precise area of only 500 by 500 feet. The SCPD uses those predictions to determine where to station its officers.

Police departments using predictive technologies hope they will improve their departments’ effectiveness. PredPol allows the SCPD to pursue that goal without any additional hires, which its budget does not allow. The system is not the only new measure Santa Cruz is implementing to try to improve policing. The SCPD has also instructed officers to engage more with the community and forge more relationships with locals as they patrol, but PredPol remains a substantial part of the city’s plans for improving police tactics.

PredPol is just one of many predictive policing algorithms available for purchase, and Santa Cruz is just one of many police departments to have adopted the technology. Predictive policing devices and technologies are growing in popularity among law enforcement departments, with thirty-eight percent of police depart-

---

1. The introductory story and application in Santa Cruz come from an article by Eyragon Eidam, *The Role of Data Analytics in Predictive Policing*, GOV’T TECH. (Sep. 2016), <http://www.govtech.com/data/role-of-data-analytics-in-predictive-policing.html>.

ments currently using predictive policing and seventy percent expecting to implement it in the next two to five years.<sup>2</sup> But as their popularity is increasing, so too are concerns about the potential downsides of the technology. Law enforcement and the public are not entirely aware of every hazard that will accompany predictive policing technologies. Due to the highly technical nature of predictive policing, public transparency should be a priority. While real-time predictions of crime locations must be withheld for the technology to provide any real benefit, police should at the very least reveal algorithm inputs, algorithms, and obsolete predictions whenever possible if employing predictive policing.

In Section I, I describe what predictive policing is, its benefits, and its drawbacks. I also explain how law enforcement typically purchases the technology from private developers and why private development raises concerns about information asymmetry.

In Section II, I describe the most common arguments for and against predictive policing. In Section III, I add my own arguments. First, arguing the importance of transparency specific to predictive policing, I contend transparency provides a critical democratic check for predictive policing. Second, I argue that transparency can increase law enforcement effectiveness in jurisdictions with predictive policing through increased legitimacy and deterrence. Finally, I propose that transparency allows the public to demand change in cases where flawed algorithms arbitrarily distribute the costs of enforcing the law.

In Section IV, I describe the current apparatus through which the public can request transparency from the government. Open records laws, at the federal and state level, require the government to provide answers upon request from citizens, but they also contain a number of exceptions. Here, the business information and law enforcement exceptions both provide potential avenues for denying public information requests.

---

2. OFFICE OF CMTY. ORIENTED POLICING SERVS., U.S. DEP'T OF JUSTICE, *FUTURE TRENDS IN POLICING* (2014). As a whole, the predictive analytics industry is expected to grow from USD \$ 3.49 billion to \$10.95 billion by 2022. Zion Market Research, *Trends in Predictive Analytics Market Size & Share will Reach \$10.95 Billion by 2022*, GLOBENEWSWIRE (Mar. 2, 2018), <https://globenewswire.com/news-release/2018/03/02/1414176/0/en/Trends-in-Predictive-Analytics-Market-Size-Share-will-Reach-10-95-Billion-by-2022.html> [https://perma.cc/99RP-LMJC]. Over 60 U.S. police departments currently use some form of predictive policing. ANDREW GUTHRIE FERGUSON, *THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT* (2018).

## I. PREDICTIVE POLICING

Predictive policing is the application of analytical techniques to identify where crime is likely to occur and who is likely to commit crime.<sup>3</sup> It refers to both traditional methods of distributing law enforcement resources where departments believe they will most be needed and modern predictive algorithms and codes, the latter of which are the focus of this article.<sup>4</sup> This section defines the modern high-tech approach to policing, weighs its benefits against its risks, and describes how law enforcement usually obtains predictive policing technology.

### A. *Summary of Predictive Policing Technology*

Predictive analytics rely on “sophisticated computer programs, extensive data sets, and professional data analysts.”<sup>5</sup> Advanced technology helps to “uncover non-obvious relationships” and make predictions that were previously impossible.<sup>6</sup> One of predictive policing’s most distinctive characteristics is its ability to predict future behavior from past data, moving beyond the dataset of past crimes.<sup>7</sup> Consider a town with one hundred convenience stores. Old predictive methods could use previous robberies at three convenience stores to predict future robberies at those stores or stores near them. A new predictive policing algorithm might combine past crimes with other data to inform officers of a greater chance of a robbery occurring at a fourth convenience store across town. The technology provides a more granular prediction, and algorithms al-

---

3. *See id.* at 1 (defining predictive policing). In 2009, the National Institute of Justice defined it as “taking data from disparate sources, analyzing them, and then using the results to anticipate, prevent and respond more effectively to future crime.” Beth Pearsall, *Predictive Policing: The Future of Law Enforcement?*, 266 NIJ J. 16, 16 (2014).

4. *See generally*, WALTER L. PERRY ET AL., PREDICTIVE POLICING 9 (RAND Corporation 2013).

5. Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1510 (2013).

6. PERRY, ET AL., *supra* note 4, at 2; *See also*, Andrew G. Ferguson, *Predictive Policing and Reasonable Suspicion*, 62 EMORY L.J. 259, 281–84 (2012) (discussing how near repeat theory and risk terrain modeling in predictive policing makes predictions otherwise impossible without advanced technology).

7. *See* Shaun B. Spencer, *When Targeting Becomes Secondary: A Framework for Regulating Predictive Surveillance in Antiterrorism Investigations*, 92 DENVER U. L. REV. 493, 499 (2015).

ready exist that can predict crimes based on likely location, offender, victim, and time.<sup>8</sup>

The predictive power of these algorithms hinges on both the analyst's role in designing the model and the data input. Analysts often take an "active role in defining the parameters of the actual data-mining analysis and the creation of clusters, links, and decision trees which are later applied."<sup>9</sup> This allows analysts to influence which variables algorithms take into account<sup>10</sup> and engrain their positive and normative decisions into future predictions. However, algorithm design is only half the story. Without representative data, even a perfectly-designed algorithm is unable to make predictions that accurately reflect the situation on the ground.<sup>11</sup> Expecting the algorithm to perform properly under these conditions would be like expecting Google Maps to find your destination after you typed in the wrong address.

When well-designed and backed by good data, predictive policing algorithms have been a boon to a number of law enforcement goals: overcoming budget restrictions, increasing surveillance, and preventing crimes before they occur. First and foremost, algorithms allow departments to get more out of a smaller force, cutting crime without significantly expanding budgets.<sup>12</sup> For example, the "Blue Crush" program predicts precise times and locations of crimes so that departments will allocate their officers accordingly.<sup>13</sup> In addition, law enforcement has deployed algorithms in the hopes of deterring criminal behaviors and warning community members of high-risk locations and times.<sup>14</sup>

Predictive policing is oriented towards surveillance rather than targeting individuals with particularized suspicion or making arrests.<sup>15</sup> Thus, the algorithms implicate few, if any, Fourth Amendment concerns. The algorithms are employed without targeting or

---

8. PERRY, ET AL., *supra* note 4, at 8; *id.* at 14.

9. Zarsky, *supra* note 5, at 1518.

10. Batya Friedman & Helen Nissenbaum, *Bias in Computer Systems*, in HUMAN VALUES AND THE DESIGN OF COMPUTER TECHNOLOGY 21, 21–39 (Batya Friedman ed., 1997) (describing the influence of design decision in algorithms generally).

11. PERRY, ET AL., *supra* note 4, at 13 (exploring the importance of good data for predictive accuracy).

12. Joel Rubin, *Stopping Crime before It Starts*, L.A. TIMES (Aug. 21, 2010), <http://articles.latimes.com/2010/aug/21/local/la-me-predictcrime-20100427-1> [<https://perma.cc/R2KF-V5FD>] (discussing the benefits of predictive policing in Los Angeles).

13. Spencer, *supra* note 7, at 505.

14. PERRY ET AL., *supra* note 4, at 2.

15. Spencer, *supra* note 7, at 494.

particularized suspicion.<sup>16</sup> Right now, the most common uses of the technology are predicting stranger offenses and organized crime,<sup>17</sup> although predictive policing has also been employed in other tasks, like predicting terrorist activity, city inspection services, and ranking tax returns for audits on the likelihood of tax evasion.<sup>18</sup> The technology predicts locations where crime is most likely to occur without reference to any individual or organization. In turn, this means the predictive technology alone does not form the basis for reasonable suspicion or probable cause. No search or seizure has occurred at the time of the prediction nor is made on account of that prediction. Because the Fourth Amendment is only implicated upon an unreasonable search or seizure, predictive policing technologies circumvent the Fourth Amendment.

High-tech predictive policing has already begun to take root in local law enforcement.<sup>19</sup> Departments are determining patrol routes and distributing officers throughout their jurisdictions on the basis of these predictions.<sup>20</sup> Although the most common types of predictions are location and time-based, Chicago has also used

---

16. Ferguson, *supra* note 6, at 287 (noting that there are “no reported cases on predictive policing in the Fourth Amendment context”). The Fourth Amendment comes into play after police officers have made an arrest or search and resolves around the rationale for that arrest. While predictive policing algorithms may facilitate eventual arrests, they are not yet used to provide the basis for arrests. Therefore, the use of such algorithms are not an aspect of the Fourth Amendment reasonable suspicion or probable cause inquiries.

17. PERRY ET AL., *supra* note 4, at 3 (discussing the use of predictive policing in stranger offenses); Scott Harris, *Product Feature: Predictive Policing Helps Law Enforcement “See around the Corners,”* POLICE CHIEF, <http://www.policechiefmagazine.org/product-feature-predictive-policing-helps-law-enforcement-see-around-the-corners> [<https://perma.cc/3HNP-W6LB>] (arguing that predictive policing best addresses organized criminal activity).

18. V.S. Subrahmanian, *Introducing the Software That Can Predict New Leaders of Terror Groups*, OBSERVER (London) (Sept. 15, 2013), <http://www.theguardian.com/world/2013/sep/15/al-qaida-terrorist-leader> [<https://perma.cc/TZ33-WWFM>] (discussing the use of predictive policing to combat terrorist activity); VIKTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, *BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK, AND THINK* 7-8, 11-12, 15 (2013) (discussing the use of predictive policing for city inspections in New York City); ERIC SIEGEL, *PREDICTIVE ANALYTICS: THE POWER TO PREDICT WHO WILL CLICK, BUY, LIE, OR DIE* 11 (2013) (discussing the use of predictive policing in IRS tax audits).

19. Andrew Papachristos, *Use of Data Can Stop Crime by Helping Potential Victims*, N.Y. TIMES, ROOM FOR DEBATE (Nov. 18, 2015, 10:13 AM), <https://www.nytimes.com/roomfordebate/2015/11/18/can-predictive-policing-be-ethical-and-effective/use-of-data-can-stop-crime-by-helping-potential-victms> [<https://perma.cc/L6D4-7JP2>] (discussing how predictive policing is being used in Chicago, Kansas City, Boston, New Orleans, and New Haven).

20. *See supra* nn. 17–18.

predictive policing to assign individuals a “score” representing their likelihood of becoming involved with a violent crime.<sup>21</sup> In a city of 2.7 million, Chicago police have identified the 1,400 people most at risk and joined forces with social workers for “interventions” in the hope of helping them avoid crime.<sup>22</sup> There is little evidence to-date of widespread adoption of Chicago’s use for predictive technologies.<sup>23</sup>

### B. Benefits of Predictive Policing

By improving the accuracy of estimates of where and when crimes will occur, predictive policing can make departments more effective and efficient.<sup>24</sup> The algorithms are also cheaper than increasing police force sizes, and they cut back on the amount of unnecessary harassment people face in the street, as I explain further below.<sup>25</sup> At least in theory, predictive policing is an alternative to “flawed investigative tools” that leads to safer streets at a lower economic cost and a lower social cost.<sup>26</sup>

Beyond the attempt to reduce unnecessary surveillance actions, predictive policing seeks to reduce the social costs of policing in three other ways. Policing in America has always been a contentious issue that cannot be entirely understood without reference to

---

21. See Jeremy Gorner, *Chicago Police Use ‘Heat List’ as Strategy to Prevent Violence*, CHI. TRIB. (Aug. 21, 2013), [http://articles.chicagotribune.com/2013-08-21/news/ct-met-heatlist20130821\\_1\\_chicago-police-commander-andrew-papachristos-heat-list/](http://articles.chicagotribune.com/2013-08-21/news/ct-met-heatlist20130821_1_chicago-police-commander-andrew-papachristos-heat-list/) [<http://perma.cc/H5MB-JKKF>] (discussing the Chicago “heat list”).

22. Monica Davey, *Chicago Police Try to Predict Who May Shoot or Be Shot*, N.Y. TIMES (May 23, 2016), [https://www.nytimes.com/2016/05/24/us/armed-with-data-chicago-police-try-to-predict-who-may-shoot-or-be-shot.html?\\_r=1](https://www.nytimes.com/2016/05/24/us/armed-with-data-chicago-police-try-to-predict-who-may-shoot-or-be-shot.html?_r=1) [<https://perma.cc/KP5E-T4FC>]. I discuss the concern these “interventions” may be pretext later in this Note.

23. Beth Pearsall, *supra* note 3, at 17 (2014) (specifying predictive policing’s use in “analytic tools and techniques like hot spots, data mining, crime mapping, geospatial prediction and social network analysis . . .”).

24. M. Todd Henderson, Justin Wolfers, & Eric Zitzewitz, *Predicting Crime*, 52 ARIZ. L. REV. 15, 34–38 (2010). The ability of predictive policing to actually reduce crime is itself still uncertain. One recent study found that a predictive policing program was not able to significantly reduce property crime. PRISCILLIA HUNT, JESSICA SAUNDERS, & JOHN S. HOLLYWOOD, *EVALUATION OF THE SHREVEPORT PREDICTIVE POLICING EXPERIMENT* 49 (2014). It did, however, reduce costs by 6–10%, and the negative results were specific to one program, not predictive technologies generally. *Id.* at 47, 49–50.

25. Jane Bambauer, *Other People’s Papers*, 94 TEX. L. REV. 205, 224 (2015).

26. Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 HARV. L. & POL’Y REV. 15, 28–30 (2016) (discussing each alternative to predictive policing in turn).

its racialized history.<sup>27</sup> In the last several years, this issue has come to a head with the widely publicized deaths of Black men and the rise of the Black Lives Matter movement.<sup>28</sup> Predictive policing provides a basis for the distribution of policing resources that curtails police discretion, which has been identified as one of the root causes of biased policing.<sup>29</sup> In general, policing also sweeps in more individuals than it should with unnecessary intrusions and physical invasions. Any time an officer stops, questions, or searches an individual or their possessions there is a cost in time, humiliation, autonomy, and dignity to that individual.<sup>30</sup> Stops, questionings, and searches that prove unfruitful also impose that same harassment cost but without any law enforcement benefit to show for it.<sup>31</sup> Predictive policing theoretically minimizes those costs by avoiding unnecessary encounters. The production of data also facilitates audits and third-party reviews of police departments,<sup>32</sup> lending legitimacy to law enforcement and providing a means to improve existing policies.

No one is arguing that predictive policing is the solution to all challenges in American law enforcement, but supporters contend that predictive algorithms, along with other tools, can make policing more effective, efficient, legitimate, and fair.

---

27. See generally Cassandra Chaney & Ray V. Robertson, *Racism and Police Brutality in America*, 17 J. OF AFR. AM. STUD. 480 (2013) (providing an overview of racism in policing over the last several decades).

28. I acknowledge this has been a publicized issue for decades and also recognize the success of the efforts of Alicia Garza and her co-activists in reinvigorating public discussion on the issue. *About*, BLACK LIVES MATTER (Mar. 11, 2017), <http://blacklivesmatter.com/about> [<https://perma.cc/TX63-QUC3>] (summarizing the national movement's rise after several widely publicized deaths).

29. PHILLIP A. GOFF ET. AL., URBAN INST., *THE SCIENCE OF POLICING EQUITY: MEASURING FAIRNESS IN THE AUSTIN POLICE DEPARTMENT* 14 (2016) (using quantitative analysis to conclude discretion increased racially biased outcomes). Of course, explicit racism also plays a large role in racist outcomes.

30. See Bambauer, *supra* note 25.

31. I recognize some in law enforcement communities would argue that even an unsuccessful law enforcement encounter helps to maintain a feeling of safety and deter crimes. Even if one supports this method of law enforcement, I think all would still acknowledge there is less benefit to an officer action that interferes with innocent activity.

32. Joh, *supra* note 26, at 29.

### C. Hazards of Predictive Policing<sup>33</sup>

Government surveillance has unsettled Americans since the founding of our country,<sup>34</sup> and the modern rise of big data-fueled predictions in the commercial context has further unnerved American consumers.<sup>35</sup> Using big data techniques from the private sector to refine government surveillance risks popular discontent. Any level of surveillance is to some degree inherently intrusive.<sup>36</sup> Surveillance requires increased attention on individuals, prying into our private matters, and, in many cases, stops and/or searches. Popular understanding of predictive policing framed by science fiction,<sup>37</sup> along with legitimate concerns about misappropriation of data and predictions,<sup>38</sup> could harm communities' relationships with and trust in the police.<sup>39</sup>

Another concern with predictive policing is that the algorithms will fail to correct grievances over policing while giving departments and politicians the grounds to declare "mission accomplished."<sup>40</sup> Because predictive algorithms rely on preexisting data, biased data can generate biased predictions.<sup>41</sup> For example, if drug arrests in a

---

33. I start by comparing the costs and benefits of predictive policing to provide a framework for considering transparency needs, but in so doing, do not mean to suggest the question of whether to use predictive policing can be resolved through a simple cost-benefit analysis. Assuming cost-benefit analysis could resolve this issue ignores the fact that many of the hazards of predictive policing I discuss in this section, particularly discriminatory ones, should be treated as hard stops to any new policing techniques.

34. See U.S. CONST. amend. IV.

35. See Kashmir Hill, *How Target Figured Out a Teen Girl Was Pregnant before her Father Did*, FORBES (Feb. 16, 2012), <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#391d1db66686> [<https://perma.cc/TE5C-EKYU>]; Charles Duhigg, *How Companies Learn your Secrets*, N.Y. TIMES (Feb. 16, 2012), <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html> [<https://perma.cc/S9QW-W56F>].

36. Bambauer, *supra* note 25.

37. See, e.g., MINORITY REPORT (DreamWorks 2002).

38. See Evan Perez, *NSA: Some Used Spying Power to Snoop on Lovers*, CNN (Sept. 27, 2013), <http://www.cnn.com/2013/09/27/politics/nsa-snooping> [<https://perma.cc/3UBU-7SEC>].

39. PERRY ET AL., *supra* note 4, at 81–83.

40. Cf. Adam M. Samaha, *Regulation for the Sake of Appearance*, 125 HARV. L. REV. 1563, 1620–33 (2012) (arguing that the appearance of success with broken windows policing allows police to continue with the practice even if the broken windows strategy does not work in actuality).

41. Sarah Brayne, *Stratified Surveillance: Policing in the Age of Big Data* (2015) (unpublished Ph.D. dissertation, Princeton University) (on file with Mudd Manuscript Library, Princeton University). Of particular concern is the role algorithms might play in "amplify[ing] feedback problems" because of the use of prior outputs to shape future inputs. Matt Reynolds, *Biased Policing is Made Worse by Errors*

community have previously skewed towards a particular racial group, then a prediction of future drug crimes based on that data will also skew towards that community. The complexity of algorithms also makes these biases harder to identify and remove because fixing the problem requires some understanding of the algorithm or computer code itself.<sup>42</sup> An algorithm may further conceal the issue by not directly relying on race but instead incorporating race as a factor through correlated variables.<sup>43</sup> The *appearance* of a “neutral mechanism” knocks the wind out of the sails of social justice movements working to redress these wrongs. Police departments can respond to such movements by claiming they have already instituted the desired change, even while that change proves hollow in reality.

Even assuming algorithms are accurate and unbiased, some fear that police will not rely on the algorithms.<sup>44</sup> Beyond wasting the resources invested in these algorithms, activists fear predictive policing will mask policing issues and become an impotent alternative to meaningful proposals.<sup>45</sup>

#### D. *How Law Enforcement Obtains the Technology*

Given the technology’s sophistication, law enforcement typically purchases predictive policing algorithms from outside contrac-

---

*in Pre-Crime Algorithms*, NEWSIDENTIST (Oct. 4, 2017), <https://www.newscientist.com/article/mg23631464-300-biased-policing-is-made-worse-by-errors-in-pre-crime-algorithms> [<https://perma.cc/ZE6R-VNHV>]; *see also* Danielle Ensign et. al, *Runaway Feedback Loops in Predictive Policing*, PROCEEDINGS OF MACHINE LEARNING RESEARCH 81:1–12 at 8–10 (2018), <http://proceedings.mlr.press/v81/ensign18a/ensign18a.pdf> [<https://perma.cc/N2HR-98V6>] (using Predpol to study runaway feedback problems with predictive policing).

42. Jens Ludwig, Director, Uni. of Chi. Crime Lab, Keynote Address at the Policing and Accountability in the Digital Age Conference (Sep. 15, 2016); Michael L. Rich, *Machines as Crime Fighters*, 30-WTR CRIM. JUST. 10, 13 (2016).

43. Anna M. Barry-Jester, Ben Casselman, & Dana Goldstein, *Should Prison Sentences Be Based on Crimes that Haven’t Been Committed Yet?*, FIFTYEIGHT (Aug. 4, 2015), <https://fivethirtyeight.com/features/prison-reform-risk-assessment> [<https://perma.cc/DL95-W9U2>].

44. Henderson, Wolfers, & Zitzewitz, *supra* note 24, at 23 (expressing concern that predictions may “be lost in the noise of the station house”).

45. Kami C. Simmons, *Police Technology Shouldn’t Replace Community Resources*, N.Y. TIMES, ROOM FOR DEBATE (Nov. 18, 2015, 3:22 AM), <http://www.nytimes.com/roomfordebate/2015/11/18/can-predictive-policing-be-ethical-and-effective/police-technology-shouldnt-replace-community-resources> [<https://perma.cc/7VWW-XAB3>] (arguing predictive algorithms fail to account for soft factors that are not in the algorithm and are a poor replacement for community policing); Joh, *supra* note 26, at 32 (arguing predictive algorithms eliminate good discretion that officers may gain from personal and community relationships).

tors rather than developing the algorithms themselves.<sup>46</sup> A number of companies have already profited from selling predictive algorithms to the government with varying purposes and levels of sophistication.<sup>47</sup> Some of these companies provide substantial support so that departments better understand the capabilities and limitations of the technology, as well as best practices for incorporating the predictions into existing department schemes. Other companies focus more exclusively on development and sales, leaving the practical implementation of the algorithms to law enforcement departments.<sup>48</sup> Either way, it is clear that the private sector has been at the forefront of developing this technology.<sup>49</sup>

But the involvement of private actors complicates the hazards of predictive policing. Separating development from implementation is sheltering the algorithm from public review, which amplifies concerns about noisy data, improper biases, and flawed algorithms.<sup>50</sup> The intellectual property rights of private companies add a layer of secrecy to predictive policing. Private developers who do not wish to relinquish information to the public can hide behind proprietary rights, which raises accountability concerns. Whose job is it to ensure models are working correctly? Who do we hold liable when things go wrong? Do we blame the developer or the user? When fingers point in both directions, it creates an accountability gap. The combination of high-tech surveillance and private development necessitates comprehensive transparency.

---

46. See, e.g., PREDPOL, <http://www.predpol.com> (providing an example of a private company's product); see also Cynthia Rudin, *Predictive Policing: Using Machine Learning to Detect Patterns of Crime*, WIRED (Aug. 2013), <https://www.wired.com/insights/2013/08/predictive-policing-using-machine-learning-to-detect-patterns-of-crime> [<https://perma.cc/XN38-FK9K>] (giving an example of collaboration with a private university).

47. See, e.g., *id.*; IBM100 - Predictive Crime Fighting, IBM100, <http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/crimefighting> [<https://perma.cc/LNX7-MBVS>] (describing IBM's products like Blue Crush); HUBSPOT, <https://www.hubspot.com> [<https://perma.cc/EN2X-ZNFY>]; PALANTIR, <http://www.palantir.com/solutions/law-enforcement> [<https://perma.cc/8ZEP-LFQK>].

48. PERRY ET AL., *supra* note 4, at 126–27 (arguing developers should provide more decision support for resource allocation and continuing assistance to departments that purchase their technology).

49. See *supra* Part I.b.

50. PERRY ET AL., *supra* note 4, at 89 (outlining the problem of “noisy and conflicting data”).

## II.

## TRANSPARENCY: THE CURRENT ARGUMENTS

A. *Routes to Information*

Oversight requires some degree of transparency and, in the case of predictive policing, that transparency should be oriented towards the general public. The question of to *whom* should predictive practices be transparent antecedes the question of *how* should that transparency be achieved. If private developers and local law enforcement currently hold the keys to that vault of information, the question becomes whether transparency is better achieved by handing those keys to another or by simply leaving the vault open for any curious passerby. Several institutions could be tasked with the responsibility of overseeing predictive policing: law enforcement, courts, legislatures, or civilian review boards. However, flaws with each of these institutions leaves public oversight as the best option.

Law enforcement has institutions designed for oversight and review.<sup>51</sup> However, self-regulation within law enforcement comes with years of baggage, including the code of silence, conflicts of interest, and a perceived lack of objectivity.<sup>52</sup> Data-collection-based technologies have exacerbated these challenges with added complexity often poorly understood by local law enforcement users.<sup>53</sup>

---

51. See 34 U.S.C. § 12601 (2018) (giving the DOJ Civil Rights Division authority to bring suits against local police departments). See generally N.Y.C. COMM’N TO COMBAT POLICE CORRUPTION, PERFORMANCE STUDY: A FOLLOW-UP REVIEW OF THE IAB’S COMMAND CENTER (Aug. 1999), <https://www1.nyc.gov/assets/ccpc/downloads/pdf/Performance-Study-A-Follow-up-Review-of-the-Internal-Affairs-Bureau-Command-Center-August-1999.pdf> [<https://perma.cc/2UTE-FNKG>] (summarizing the NYPD’s Internal Affairs Bureau).

52. Steven D. Seybold, *Somebody’s Watching Me: Civilian Oversight of Data-Collection Technologies*, 93 TEX. L. REV. 1029 (2015), 1041–42 (summarizing the weaknesses of self-regulation); Dina Mishra, *Undermining Excessive Privacy for Police: Citizen Tape Recording to Check Police Officer’s Power*, 117 YALE L.J. 1549, 1552 (2008) (“Policy corruption can undermine internal monitoring and sanctions even where they do apply.”). Police departments face a challenge in self-regulation. Culture and the desire to stand together makes it difficult for whistleblowers to report colleagues and superiors. This challenge has led many to advocate for civilian review boards and similar external checks on departments.

53. See Jack M. Balkin, *The Constitution in the National Surveillance State*, 93 MINN. L. REV. 1, 4, 7–9 (2008) (tracing the growth of private-public coordination in security technology); Justin Jouvenal, *Police Are Using Software to Predict Crime. Is It the ‘Holy Grail’ or Biased Against Minorities?*, WASH. POST (Nov. 17, 2016), [https://www.washingtonpost.com/local/public-safety/police-are-using-software-to-predict-crime-is-it-a-holy-grail-or-biased-against-minorities/2016/11/17/525a6649-0472-440a-aae1-b283aa8e5de8\\_story.html](https://www.washingtonpost.com/local/public-safety/police-are-using-software-to-predict-crime-is-it-a-holy-grail-or-biased-against-minorities/2016/11/17/525a6649-0472-440a-aae1-b283aa8e5de8_story.html) [<https://perma.cc/E5FB-XQSF>] (discussing

Courts, the traditional oversight mechanism for much of law enforcement's activities,<sup>54</sup> are an especially weak check on predictive policing because they are limited by constitutional law in how they may restrain law enforcement. The Fourth Amendment's protections apply only after there is an invasion of a reasonable expectation of privacy.<sup>55</sup> Because predictive policing is used for surveillance, which does not require reasonable suspicion, police can observe citizens without any court interference.<sup>56</sup> Courts may never consider predictive policing policies if no relevant case is ever brought.

Legislatures, on the other hand, face political pressure. As Professor Donald Dripps explains, legislatures face political pressures to adopt the strongest law enforcement policy rather than the fairest or most reasonable policy.<sup>57</sup> Driven by elections, legislators fear the perception of public safety cut-backs and are hard-pressed to craft rules that limit law enforcement resources.

In addition to procedural and political challenges, both courts and legislatures face another problem in making any information they learn public. DOJ and internal police reviews have worked to conceal technologies from the public, using devices such as non-

---

law enforcement's incomplete understanding of how predictive policing technology works).

54. Stephen D. Mastrofski & James J. Willis, *Police Organization Continuity and Change: Into the Twenty-First Century*, 39 CRIME & JUST. 55, 115 (2010) ("Courts oversee a range of police practices: personnel and administrative matters, as well as practices in the field.").

55. See *Katz v. United States*, 389 U.S. 347 (1967). See generally William J. Stuntz, *The Political Constitution of Criminal Justice*, 119 HARV. L. REV. 780 (2006) (arguing reliance on Constitutional law has hampered attempts to remedy problems in criminal justice).

56. Seybold, *supra* note 52, at 1044–45 (describing the weakness of judicial oversight). Judicial oversight is especially weak when it comes to the surveillance of individuals in public places, even when that oversight relies upon technology; *United States v. Steinhorn*, 739 F. Supp. 268, 272 (D. Md. 1990) (citing *United States v. Batchelder*, 442 U.S. 114, 124 (1979)) (holding law enforcement does not need reasonable suspicion to observe and investigate persons' public conduct).

57. Donald A. Dripps, *Criminal Procedure, Footnote Four, and the Theory of Public Choice; or Why Don't Legislatures Give a Damn about the Rights of the Accused*, 44 SYRACUSE L. REV. 1079, 1080 (1993) (using public choice theory to explain why reliance on legislatures to constrain law enforcement will prove futile). Legislatures are also easier to buy, as evidenced by PredPol's \$2 million dedicated to lobbying the Arizona statehouse. Ali Winston & Ingrid Burrington, *A Pioneer in Predictive Policing is Starting a Troubling New Project*, VERGE (Apr. 26, 2018, 1:36 PM), <https://www.theverge.com/2018/4/26/17285058/predictive-policing-predpol-pentagon-ai-racial-bias> [<https://perma.cc/44ZW-4PPV>].

disclosure agreements to avoid challenges to their practices.<sup>58</sup> Courts often review technologies *in camera*, and, in general, American courts are opaque with little public scrutiny.<sup>59</sup> Thus, these institutions prove fairly weak at distributing information to the public.<sup>60</sup> Opening up information to the public avoids these transparency hurdles, helps to ensure meaningful oversight, and furthers several normative goals that I discuss below.

### B. Arguments for Transparency

Previous advocates for transparency have offered variations on five arguments: (1) democratic legitimacy, (2) the prevention of misconduct, (3) a remedy for injustices, (4) better substantive policies, and (5) community trust.

First, transparency lends democracy legitimacy. Transparency has inherent value: openness in government has been a stalwart companion to democracy, as it creates the well-informed public necessary to sustain democracy.<sup>61</sup> The Freedom of Information Act (FOIA) and its state counterparts recognize the basic right of American citizens to obtain government information,<sup>62</sup> and many argue that the necessity of an informed citizenry for popular sovereignty

---

58. Aff. of FBI Supervisory Special Agent Bradley S. Morrison, Chief, Tracking Technology Unit, Operation Technology Division in Quantico, Virginia, at 2, Apr. 11, 2014, attach. to City's Verified Answer, *Hodai v. City of Tucson*, No. C20141225 (Ariz. Super. Ct. Apr. 14, 2014).

59. Stephanos Bibas, Essay, *Transparency and Participation in Criminal Procedure*, 81 N.Y.U. L. REV. 911, 920–31 (2006) (discussing the exclusion of outsiders from judicial proceedings).

60. There is an argument for enlisting civilian review boards as the means of obtaining and distributing information with predictive policing. I do not focus on it here because many existing constraints and normative arguments that apply to civilian review boards would also apply to transparency with the general public. In addition, the boards have yet to catch on and become popular in the United States. Udi Ofer, *Getting It Right: Building Effective Civilian Review Boards to Oversee Police*, 46 SETON HALL L. REV. 1033, 1053 (2016) (only six of the fifty largest police departments currently have civilian review boards with any serious authority). For an argument in favor of civilian review boards, *see id.*

61. *Scherr v. Universal Match Corp.*, 297 F. Supp. 107, 110 (S.D.N.Y. 1967) (quoting 12 ASCAP Copyright Law Symposium 96, 105 (1961)); Jeremy Bentham, *An Essay on Political Tactics* (1791), *reprinted in* 2 THE WORKS OF JEREMY BENTHAM, 299, 310–12 (John Bowring ed., Russell & Russell 1962) (1843) (“To conceal from the public the conduct of its representatives, is to add inconsistency to prevarication . . .”).

62. 5 Ill. Comp. Stat. 140/1 (2006) (“ . . . all persons are entitled to full and complete information regarding the affairs of government . . . .”); *Sullo & Bobbitt, PLLC v. Abbott*, 2012 WL 2796794 (N.D. Tex. 2012) (working under the assumption there is no absolute right to information held by the government).

trumps personal privacy rights.<sup>63</sup> As the Republic rules only with the consent of the governed, the governed must be informed of the Republic's material activities in order to agree to its rule.<sup>64</sup>

Second, transparency aids accountability, preventing police misconduct.<sup>65</sup> An informed citizenry is able to check against corruption and “protect the public from secret government activity.”<sup>66</sup> Unlike my first argument, this is a monitoring claim rather than a social contracting claim. While evidence of wrongdoing is not a prerequisite to disclosure,<sup>67</sup> transparency helps to prevent and redress police abuses of power. Electoral incentives can lead politicians to make popular but ineffective decisions.<sup>68</sup> Transparency gives the public the ability to check the work of the political branches rather than relying on government officials to self-report. Just as transparency has changed policies in the past, such as restricting the use of the controversial Stingray technology,<sup>69</sup> transparency could revise predictive policing policies, as well.

Third, when misconduct does occur, whether with law enforcement policies or predictive policing, transparency helps provide a remedy to the aggrieved party. Part of the role of adjudicatory bodies is to provide legal and public recognition that a wrong has occurred.<sup>70</sup> That recognition provides some relief to the aggrieved

---

63. See, e.g., Jesse H. Alderman, *Police Privacy in the iPhone Era?: The Need for Safeguards in State Wiretapping Statutes to Preserve the Civilian's Right to Record Public Police Activity*, 9 FIRST. AMEND. L. REV. 487, 523 (2011).

64. Barry Friedman, *Secret Policing*, 2016 U. CHI. LEGAL F. 99, 120 (2016) (“... the members of the public — the electorate are their bosses. And the bosses have a right to know what is going on.”).

65. Adam M. Samaha, *Government Secrecy, Constitutional Law, and Platforms for Judicial Intervention*, 53 UCLA L. REV. 909, 970 (2006) (“Popular accountability depends on information access . . .”).

66. *Lambries v. Saluda Cty. Council*, 409 S.C. 1, 15 (2014); see also *Elkins v. Federal Aviation Admin.*, 99 F. Supp. 3d. 90, 95 (D.D.C. 2015) (“The basic purpose of FOIA is to ensure an informed citizenry . . . needed to check against corruption and to hold the governors accountable to the governed.” (quoting *John Doe Agency v. John Doe Corp.*, 493 U.S. 146, 152 (1989))).

67. E.g., *Chicago All. for Neighborhood Safety v. City of Chicago*, 808 N.E.2d 56, 73 (Ill. App. Ct. 2004) (holding plaintiffs need not show illegal or unethical conduct to prevail in an Open Records suit for information regarding police activities).

68. See Dripps, *supra* note 57, at 1089 (describing the weaknesses of legislative oversight of law enforcement).

69. H. 4522, 121st Leg. (S.C. 2015).

70. The concept of moral signaling to build social cohesion has deep roots in criminal jurisprudence. See generally H.L.A. Hart, *Social Solidarity and the Enforcement of Morality*, 35 U. CHI. L. REV. 1 (1967); EMILE DURKHEIM, *SOCIAL COHESION AND ANOMIE* (1895).

and helps demarcate the boundaries of acceptable behavior.<sup>71</sup> Transparency with predictive policing allows the public to ensure law enforcement recognizes any wrongful uses of predictive policing against aggrieved individuals and communities. It also helps inform what is an acceptable use of predictive policing.

Fourth, transparency produces better policies. Much of this argument boils down to the age-old idea that two heads are better than one. Transparency allows a greater number of parties, with a greater variety of interests, to review predictive policing. The public includes organizations devoted to analysis that can give more attention to issues than law enforcement.<sup>72</sup> As previously discussed, mechanization often carries with it hidden subjectivities and errors creating biases in application.<sup>73</sup> The Breathalyzer, for example, was a black box not fully understood by many in law enforcement. Widespread “public acceptance and cultural prestige” delayed in-depth analysis of its flaws for several decades, and now the certainty of criminal sentences has come under doubt.<sup>74</sup> Similarly, widespread acceptance of predictive policing could risk faults lying undetected for years. Transparency ameliorates that problem by allowing scrutiny early in the process, catching flaws before adoption rather than decades after. The crowdsourcing of predictive policing review is especially important because its complexity requires expertise, time, and attention, which non-governmental organizations can provide.<sup>75</sup>

Fifth, transparency builds community trust. This argument can cut two ways. Revelations of furtive predictive policing could outrage citizens in the same way revelations of the furtive use of Sting-

---

71. Cf. Jenia I. Turner, *Policing International Prosecutors*, 45 N.Y.U. J. INT'L L. & POL. 175, 205–07, 212 (2012) (making this argument in the context of international prosecutions).

72. See ELIZABETH T. BORIS, *NONPROFIT ORGANIZATIONS IN A DEMOCRACY – ROLES AND RESPONSIBILITIES*, 41–46 (2006).

73. See Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 10–16 (2014); Kenneth A. Bamberger, *Technologies of Compliance: Risk and Regulation in a Digital Age*, 88 TEX. L. REV. 669, 676 (2010).

74. Andrea Roth, *Trial by Machine*, 104 GEO. L.J. 1245, 1260–76 (2016); Reginald Fields, *Ohio Commits \$6.4 Million For New Breathalyzers That Face Lawsuits Across The Country*, CLEVELAND PLAIN DEALER (Nov. 17, 2008), [http://blog.cleveland.com/metro/2008/11/ohio\\_commits\\_64\\_million\\_for\\_ne.html](http://blog.cleveland.com/metro/2008/11/ohio_commits_64_million_for_ne.html) [https://perma.cc/LE6P-DLQW]; *Lawyers: Massachusetts Withheld Evidence About Breathalyzers*, INS. J. (Aug. 28 2017), <https://www.insurancejournal.com/news/east/2017/08/28/462537.htm> [https://perma.cc/E9BV-QA3T].

75. BORIS, *supra* note 72, at 41–46.

rays did.<sup>76</sup> What citizens do not know, they cannot get mad about.<sup>77</sup> On the other hand, transparency may increase community trust in the long run. The refusal to answer requests for public inspection diminishes public trust.<sup>78</sup> Transparency gives the public answers and ensures law enforcement provides information upfront. When reporters broke stories about Stingrays, public trust took a serious hit, not just because of the use of the technology, but because law enforcement hid the technology from the public.<sup>79</sup> Assuming an agency's use of predictive policing will eventually come to light, it is better for public trust for law enforcement, rather than journalists, tabloids, and blogs, to break the story.

### III.

#### ADDITIONAL ARGUMENTS FOR TRANSPARENCY SPECIFIC TO PREDICTIVE POLICING

I add three arguments to make the case for transparency with predictive policing: transparency (1) provides a democratic check where democratic oversight is particularly weak, (2) aids the effectiveness of predictive policing through deterring crime, and (3) cuts down on an unfair distribution of costs.

##### A. *Transparency as an Additional Deterrent*

Transparency with predictive policing deters criminal activity by increasing trust in the system of law enforcement and notifying the public of improved law enforcement techniques.

##### 1. Improved Effectiveness through Broader Trust in Law Enforcement

Forty-eight percent of the public lacks confidence in police.<sup>80</sup> Many feel themselves on the wrong side of a gulf separating insiders and outsiders of the system. Those on the outside feel they are unfairly targeted through discriminatory law enforcement practices and have no say in their design.<sup>81</sup> Much of the discontent with po-

---

76. Monte Reel, *Secret Cameras Record Baltimore's Every Move from Above*, BLOOMBERG BUSINESSWEEK (Aug. 23, 2016), <https://www.bloomberg.com/features/2016-baltimore-secret-surveillance> [<https://perma.cc/PY58-ZMFX>].

77. Of course, that Orwellian argument sacrifices democratic legitimacy.

78. Steven D. Zansberg & Pamela Campos, *Sunshine on the Thin Blue Line: Public Access to Police Internal Affairs Files*, 22 COMM. LAW. 34, 34 (2004).

79. See Reel, *supra* note 76.

80. Jeffrey M. Jones, *In U.S., Confidence in Police Lowest in 22 Years*, GALLUP (Jun. 19, 2015), <http://www.gallup.com/poll/183704/confidence-police-lowest-years.aspx> (finding a 52% confidence in police – the lowest since 1993).

81. See BLACK LIVES MATTER, *supra* note 28.

licing today does not come from substantive law but from procedural injustices.<sup>82</sup> An individual who believes he is stopped by an officer because he is speeding has a much more positive view of law enforcement, as well as greater trust in officers, than an individual who believes he is stopped by an officer on account of his race or religion.<sup>83</sup>

Those who believe in the fairness of procedures are more likely to obey police commands and laws even when authorities are not present.<sup>84</sup> Tom R. Tyler, a professor of psychology and law at Yale University, has led much of this research, finding a correlation between those who believe the legal system is fair and those who obey the law.<sup>85</sup> Tyler's research suggests that most people comply with the law not for fear of getting caught, but out of a belief that the law and legal system are legitimate.<sup>86</sup> Personal experiences with the law and media stories shape that sense of legitimacy,<sup>87</sup> so positive experiences and stories help to build a self-regulating society. Fewer laws are broken when people appreciate the system behind those laws.

This is where the argument regarding predictive policing's ability to build community trust in law enforcement comes into play. Transparency builds trust by giving the public answers and avoiding media scandals.<sup>88</sup> Transparency also allows the public to play an active role in shaping policy, promoting buy-in from communities affected by those policies.<sup>89</sup> Combatting disparate attention from law enforcement, predictive policing provides a reasoned, numerical method for distributing law enforcement resources.<sup>90</sup> While concerns about racial bias continue with predictive

---

82. See Tom R. Tyler et al., *Street Stops and Police Legitimacy: Teachable Moments in Young Urban Men's Legal Socialization*, 11 J. EMPIRICAL LEGAL STUD. 751 (2014).

83. Tom R. Tyler, *Procedural Justice, Legitimacy, and the Effective Rule of Law*, 30 CRIME & JUST. 283, 329 (2003) (analyzing through psychology).

84. *Id.* at 284.

85. TOM R. TYLER, *WHY PEOPLE OBEY THE LAW* 4–5 (1990). Professor Tyler first popularized the theory that the legitimacy of the legal system could benefit compliance with the law, and upon additional testing has found a moderately strong correlation between faith in the legal system and obedience to the law. *E.g.*, Tom R. Tyler & Jeffrey Fagan, *Legitimacy and Cooperation: Why Do People Help the Police Fight Crime in Their Communities?* 6 OH. ST. J. CRIM. L. 231, 246 (2008) (finding a correlation over .5 for obligation, trust, confidence in the legal system (indices for legitimacy) and identification with the police).

86. *Id.* at 178.

87. *Id.*

88. Zansberg & Campos, *supra* note 78, at 34.

89. Reel, *supra* note 76.

90. See BLACK LIVES MATTER, *supra* note 28.

algorithms because algorithms incorporate preexisting data that may be biased, ex ante transparency can help to resolve those concerns. Outside actors can review algorithms to identify problems like racial bias and propose effective solutions. At the very least, a willing commitment to transparency and active effort to improve policy should encourage greater trust in and legitimization of law enforcement. This trust and legitimacy will in turn further the ultimate law enforcement goal of increasing compliance with the law.

## 2. Improved Effectiveness through a Deterrence Effect

Throughout this article, I have assumed that predictive policing works at least to some degree. Assume for a moment that it does not work—that police officers could get just as much out of rolling dice. Even then, transparent predictive policing still provides a social benefit—deterrence.

In 2013, the Memphis Police Department found itself relying on luck when its predictive policing proved to be ineffective.<sup>91</sup> Memphis had relied upon results from the Blue Crush system for over three years to allocate resources throughout the city and determine when officers should assume their posts. Three years after implementing the system, crime had decreased thirty percent in the metropolitan area, and the department had repeatedly publically attributed the improvement to Blue Crush.<sup>92</sup>

A system audit, however, soon revealed that “reams of data” had not been input into the system. Without that data, the system could not accurately predict where crimes were going to occur.<sup>93</sup> The police may as well have been rolling dice, but the department’s effectiveness had substantially improved. How could this have happened? The answer could rest in the web of social factors that influences crime rates and continues to elude most scholars on the topic.<sup>94</sup> The better answer is that appearance shapes reality.<sup>95</sup>

---

91. Leslie A. Gordon, *A Byte Out of Crime*, A.B.A.J., Sept. 2013, at 18, 19.

92. See Andrew Ashby, *Operation Blue C.R.U.S.H. Advances at MPD*, MEMPHIS DAILY NEWS (Apr. 7, 2006), <https://www.memphisdailynews.com/editorial/Article.aspx?id=30029> [<https://perma.cc/UQ8Q-YXVC>]; *Blue Crush Statistics*, WMC ACTIONNEWS5, <http://www.wmcactionnews5.com/story/4924493/blue-crush-statistics>; *Memphis Police Department Reduces Crime Rates with IBM Predictive Analytics Software*, IBM NEWS ROOM (Jul. 21, 2010), <https://www-03.ibm.com/press/us/en/pressrelease/32169.wss> [<https://perma.cc/RB4B-BZJD>].

93. Gordon, *supra* note 91.

94. Peter-Jan Engelen, et al., *What Determines Crime Rates? An Empirical Test of Integrated Economic and Sociological Theories of Criminal Behavior*, 53 THE SOC. SCI. J. 247 (2016) (noting the uncertainty of which factors are related to crime rates and adding their own quantitative analysis to the mix).

For example, the appearance of a weak fiscal state at a bank creates a bank run, which in turn brings that weak fiscal state to fruition.<sup>96</sup> Whether or not the bank was in actuality weak to begin with is of little importance. What matters for the end result is the *appearance* of instability, and even a false or illogical belief can impose such results.<sup>97</sup> Applied to predictive policing, the appearance of a predictive algorithm in Memphis may have helped reduce crime. The Memphis community was aware of the algorithm, believed in its ability to predict where crime was likely to occur, and became convinced police would foil criminal plots. Community members chose not to attempt crime in the first place, and the crime rate fell.<sup>98</sup> While many obey the law out of the perceived legitimacy of the law alone, for those weighing the likelihood of getting caught, better policing tools matter.<sup>99</sup>

The influence of public perceptions can augment the decrease in crime resulting from predictive policing. Of course, a properly applied algorithm would have benefitted Memphis more than one with missing inputs. I am not arguing that the effectiveness of the algorithm is irrelevant, but instead that transparency—sharing with the public the fact that a predictive policing algorithm is in use—has substantial benefits for law enforcement. The LAPD has instituted audits of its own Blue Crush system.<sup>100</sup> Transparency with the LAPD system and the audit furthers the deterrence mission of the department.<sup>101</sup> Moreover, transparency decreases crime without putting citizens or police officers in harm's way. This argument does not address what aspect of the predictive algorithm should be released, and the argument that certain information could be used by criminals to avoid law enforcement persists. Regardless, transparency provides an opportunity to avoid the tremendous social

---

95. Samaha, *supra* note 40, at 1582–97 (2012) (discussing the relationship between appearance and reality).

96. *Id.* at 1592–96 (providing the example of a bank run).

97. Robert K. Merton, *The Self-Fulfilling Prophecy*, 8 ANTIOCH REV. 193, 195 (1948) (discussing the superficial importance of truth or logic to beliefs for results).

98. *See supra* note 96.

99. Tyler, *supra* note 83, at 329.

100. Guy Adams, *The Sci-Fi Solution to Real Crime*, INDEPENDENT (Jan. 11, 2012), <https://www.independent.co.uk/news/world/americas/lapds-sci-fi-solution-to-real-crime-6287800.html> [<https://perma.cc/W3DX-EBHQ>] (describing the LAPD's double-blind testing of its technologies to prevent flawed data).

101. An audit is one means of publicizing the fact that predictive policing is in place. This makes the community more aware of the policing tool, which can deter crime. *See supra* nn. 95–99.

costs of intervening in suspected criminal affairs, because officers can rely upon the deterrent effect of predictive policing.

*B. Transparency as an Accountability Gap-Filler*

Predictive policing escalates challenges with accountability because it lacks a good democratic check. Without public transparency, errors are likely to go on undetected. Just like the fear of undetected flaws was at play with the Breathalyzer,<sup>102</sup> flaws in predictive policing have also raised concerns, especially the incorporation of old biases or flawed data.<sup>103</sup> Normally, law enforcement, legislatures, and courts would provide the democratic check ensuring the technology works as promised. As discussed above, these institutions prove especially ineffective in the context of predictive policing, as the technology is designed by an entity different from the one implementing it.<sup>104</sup>

Law enforcement lacks the resources to understand predictive algorithms and hold private developers accountable. Private companies develop predictive algorithms because they have the resources to develop the technology.<sup>105</sup> This means law enforcement never develops the same understanding of the algorithms as the private companies. This disparity in knowledge leaves law enforcement at the whims of analysts who have considerable discretion in algorithm design.<sup>106</sup> A single actor within law enforcement could in theory become an expert on a predictive algorithm, but, given the market fragmentation for predictive policing, expertise on every algorithm in use by law enforcement is likely unobtainable and unaffordable.<sup>107</sup> Law enforcement would either have to expend considerable resources hiring experts or grant one private developer a monopoly over the market. Of the many predictive technology options available, picking just one, or a few, is unlikely to satisfy the needs of departments across widely ranging communities. Either option impairs the expected benefits of predictive policing.

Courts could theoretically provide a platform for expert testimony and transparency, but courts will rarely encounter this de-

---

102. Roth, *supra* note 74, at 1260–72.

103. Brayne, *supra* note 41.

104. See *supra* Part II.a for a discussion of why other institutions would prove ineffective.

105. PREDPOL, *supra* note 46; Rudin, *supra* note 46.

106. Zarsky, *supra* note 5, at 1518.

107. See Miriam A. Cherry & Robert L. Rogers, *Markets for Markets: Origins and Subjects of Information Markets*, 58 RUTGERS L. REV. 339, 342 (2006) (discussing the fragmentation of the market).

bate. Unlike the Breathalyzer, which provides probable cause for arrests, predictive policing algorithms are just a means of distributing surveillance resources throughout a community. They do not provide probable cause or even reasonable suspicion.<sup>108</sup> There is no invasion of reasonable expectations of privacy to provide grounds for a suit in court. The delayed realization of concerns with the Breathalyzer provided an imperfect solution. Tiptoeing around the jurisdiction of the court, the situation with predictive policing delayed further.

The resulting lack of democratic oversight is more than just a problem of erroneous results.<sup>109</sup> We give a unique power to law enforcement to exert considerable force against us and engage in activities that would often be unlawful if carried out by a private citizen.<sup>110</sup> In the context of predictive policing, the government is delegating that power to algorithm developers. The information disparity between law enforcement and developers results in an *unchecked* delegation of democratically granted authority to a private party.

Public transparency resolves this issue by providing a check on developers. Transparency allows private parties who have greater expertise and resources to better inform the public by analyzing algorithms, their implementation, and their effects on communities.<sup>111</sup> Building on previous notions of the “private attorney general,”<sup>112</sup> these private parties can fill the information gap—educating both the public and public officials.

---

108. Seybold, *supra* note 52, at 1044–45. *But see* Ferguson, *supra* note 6 at 262 (2012) (suggesting predictive policing has Fourth Amendment ramifications). Courts will see FOIA suits, but the debate in those suits is over whether the suing party may gain access to predictive policing information. Before that information has been released to the public, there is no one with expertise in the area to evaluate the technology. The judge may review the information in camera, but a judge’s in camera evaluation of predictive policing practically cannot be as thorough or scientific as external review by a testifying expert.

109. BARRY FRIEDMAN, UNWARRANTED: POLICING WITHOUT PERMISSION 5–6 (2017).

110. *Id.*

111. BORIS, *supra* note 72, at 41–46.

112. *See generally* Myriam E. Gilles, *Reinventing Structural Reform Litigation: Deputizing Private Citizens in the Enforcement of Civil Rights*, 100 COLUM. L. REV. 1384 (2000); Carl Cheng, *Important Rights and the Private Attorney General Doctrine*, 73 CAL. L. REV. 1929 (1985).

C. *Transparency as Guarantee of Fairness*

Predictive policing algorithms determine the fairness of the distribution of the burden of policing; they help determine who and where to target for additional surveillance.<sup>113</sup> In effect, they determine the quantity of law enforcement surveillance that individuals, businesses, and organizations will undergo within their community. Unguarded by the Fourth Amendment, the sweetheart of courts and scholars alike, surveillance has not received as much academic attention.<sup>114</sup> But, those concerned with surveillance note that surveillance is inherently discretionary<sup>115</sup> and costly to those the government chooses to watch.<sup>116</sup> Surveillance is inherently intrusive — it will encroach on subjective preferences for privacy, inhibit free expression, stigmatize, and increase the number of unwarranted police interventions.<sup>117</sup> The cost of this intrusion raises the question of how law enforcement determines who will bear that cost.

Predictive policing provides one answer, but predictive policing with public transparency is the better answer. When algorithms work correctly, the costs of surveillance are borne by those the government has a basis to believe are more likely engage in criminal activity or those who have the greatest need for government protection. However, as shown by the MPD, algorithms are vulnerable to clerical mistakes or flawed methodologies that result in arbitrary predictions.<sup>118</sup> This imposes the costs of surveillance arbitrarily across a community and removes the rational basis for those selected for greater surveillance.

---

113. PERRY ET AL., *supra* note 4, at 2.

114. Joh, *supra* note 26, at 18–19 (claiming surveillance discretion has received little scholarship).

115. *Id.* at 15.

116. See H.R. SUBCOMM. ON CIVIL AND CONSTITUTIONAL RIGHTS, H.R. COMM. ON THE JUDICIARY, 98TH CONG., REP. ON FBI UNDERCOVER OPERATIONS 2–3 (Comm. Print 1984) (discussing the stigma of investigations); Jeremy Gomer, *supra* note 21, (arguing investigation can lead to anxiety); Bambauer, *supra* note 25 (discussing the relevance of investigation for expression).

117. *Id.* But see Ana Viseu, Andrew Clement, & Jane Aspinall, *Situating Privacy Online*, 7 INFO., COMM., & SOC'Y 92 (arguing the classic “nothing to hide argument” that surveillance programs do not infringe on privacy unless the target is engaged in criminal activities).

118. Cf. E. Scott Reckard, *Data Compilers' Secret Scores Have Consumers Pegged—Fairly or Not*, L.A. TIMES (Apr. 8, 2014), [www.latimes.com/business/la-fi-secret-consumer-scores-20140409,0,6240971.story](https://www.latimes.com/business/la-fi-secret-consumer-scores-20140409,0,6240971.story) [https://perma.cc/322D-NHNY] (describing clerical mistakes in the commercial context).

Additionally, algorithms that are fed biased data will perpetuate that bias.<sup>119</sup> Currently, certain communities are over-policed.<sup>120</sup> In some forms, skewed data are a result of individual officer determinations, while in other forms, department-wide determinations skew results.<sup>121</sup> The result is biased data reflecting disproportionate enforcement against lower-income communities and racial minorities.<sup>122</sup> Predictive policing that uses these data for future predictions incorporates its old biases, perpetuating them.

As discussed above, even a flawed predictive policing algorithm benefits jurisdictions through deterrence. Every member of the jurisdiction benefits from these improvements, but the flawed algorithm imposes the cost of enforcement arbitrarily among only a few members of society. In effect, a few are selected without any basis for a tax in the form of increased surveillance that benefits the entire group. Even worse, biased data can lead low-income individuals and racial minorities to pay the cost for that community benefit.<sup>123</sup>

One could respond that predictive policing could help to further entrench biased policing. A legitimating theory—quantitative analysis should replace policing discretion—deflects public scrutiny of law enforcement. The broken windows theory helped police justify cracking down on public nuisance crimes, and the ensuing attribution of decreases in crime to that theory further engrained it.<sup>124</sup> The same problem could arise where police use predictive algorithms to justify decisions if the public is unaware of its actual utility. Allowing successes to be wrongly attributed to predictive policing legitimates and builds the popularity of predictive policing models, making it harder to remove or change in the future. Biased policing in algorithms could outlive biased policing in other methods already recognized as flawed.

---

119. Joh, *supra* note 26, at 30.

120. See generally Eric J. Miller, *Role-Based Policing: Restraining Police Conduct “Outside the Legitimate Investigative Sphere”*, 94 CAL. L. REV. 617 (2006) (arguing some communities are over-policed).

121. See *Cost-Benefit Analysis in Criminal Justice*, POLICEONE (May 2, 2013), <https://www.policeone.com/grants/articles/6218674-Cost-benefit-analysis-in-criminal-justice> [<https://perma.cc/63HV-5QMW>].

122. Miller, *supra* note 120. *But see* Jeffrey Brantingham et al., *Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial*, 5 J. STATS. & PUB. POL. 1, 5 (2017) (failing to reject their null hypothesis that the predictive policing technology increased racial bias in Los Angeles).

123. *Id.*

124. Samaha, *supra* note 40, at 1620–32.

Transparency will allow external institutions to conduct audits and the public to pressure law enforcement to correct flawed algorithms. Combining this with my previous arguments, transparency doubles down on deterrence where algorithms are successful and fair. Where algorithms are not successful and fair, transparency provides a much-needed democratic check to ensure a fair law enforcement scheme. Despite the need for transparency with predictive policing, the current statutory scheme provides little opportunity to access the information.

#### IV. RESTRICTIONS ON PUBLIC ACCESS TO PREDICTIVE POLICING INFORMATION

The process of predictive policing creates information about the algorithm or code—the dataset input into those algorithms, the incorporation of predictive policing into existing law enforcement strategies, and the predictions themselves. In this section, I address the avenues and obstacles to obtaining access to each of those forms of information.

##### A. *Open Records Laws*

Although it may appear that the existence of open records laws would ensure transparency around predictive policing, the broad exceptions to these laws frustrate that purpose.<sup>125</sup> Law enforcement has not traditionally shared vast amounts of internal data with the public.<sup>126</sup> The incorporation of technology developed by private companies, including predictive policing algorithms, has further

---

125. There is, of course, always the possibility that the private company elects to voluntarily release information. This degree of disclosure may not be as complete, nor is it universal across the industry, but there are companies that have done it. Letter from Michael R. Gehrman, Assistant Corp. Counsel, to J. Ader (May 18, 2018) (on file with Muckrock, <https://www.muckrock.com/foi/elgin-7770/foia-elgin-police-dept-predpol-documents-51858/#file-190433> [https://perma.cc/EE9J-PNLR]) (releasing information after a court battle began); *CivicScape*, GITHUB (June 5, 2018), <https://github.com/CivicScape/CivicScape> [https://perma.cc/K8DC-VAF4] (CivicScape released information about its technology); Josh Kaplan, *Predictive Policing and the Long Road to Transparency*, S. SIDE WEEKLY (July 12, 2017), <https://southsideweekly.com/predictive-policing-long-road-transparency> [https://perma.cc/7BNP-9VZB] (Chicago released information regarding predictive policing results, but not the algorithm, after seven years of refusing to provide the information.).

126. See Friedman, *supra* note 64.

hindered transparency.<sup>127</sup> When the public, whether non-government organizations or individuals, have wanted to check in on the actions of their government, they have traditionally turned to their statutory rights under open records laws.<sup>128</sup>

FOIA, the federal open records law, makes government information available to private citizens upon request.<sup>129</sup> The theoretical underpinning of the law is that access to “government information is a basic right of American citizens.”<sup>130</sup> Transparency provides the public the ability to scrutinize government actions and, thus, hold their government accountable.<sup>131</sup> Given the rights-based notion of FOIA, records are made available regardless of purpose or motivation such that “even a person with only ‘idle curiosity’ may request disclosure of information.”<sup>132</sup>

While FOIA favors open access to government activity, it still balances the public’s right against the interest of the government in keeping some information secret.<sup>133</sup> The act has a number exceptions for information pertaining to national defense, foreign policy, law enforcement, commercial trade secrets, and other categories of information that could cause harm to the government or other parties if shared with the general public.<sup>134</sup> A substantial number of statutory guidelines and judicial decisions delineate the boundaries of each of these categories.<sup>135</sup>

---

127. See *Chrysler Corp. v. Brown*, 441 U.S. 281, 293 (1979) (recognizing a private sector interest in secrecy in the decision to not disclose information).

128. Several legislatures have expressly recognized the purpose of open records laws is to create transparency in government. 37A AM. JUR. 2D, Freedom of Information Acts § 2.; Cal. Gov. Code § 6250; 5 ILL. COMP. STAT. 140/1 to 140/2 (2010).

129. See, e.g., *Renegotiation Bd. v. Bannerkraft Clothing Co.*, 415 U.S. 1, 18 (1974) (recognizing the process by which FOIA achieves government transparency).

130. Monique C.M. Leahy, *Proof Supporting Disclosure under State Freedom of Information Acts*, 132 AM. JUR. PROOF OF FACTS 3d 1 (2013) (arguing access to government information is a basic right).

131. See 37A AM. JUR. 2D, Freedom of Information Acts § 1; *ACLU v. U.S. Dept. of Justice*, 750 F.3d 927 (D.C. Cir. 2014).

132. *San Lorenzo Valley Community Advocates for Responsible Educ. v. San Lorenzo Valley Unified Sch. Dist.*, 139 Cal. App. 4th 1256 (6th Dist. 2006).; 37A AM. JUR. 2D, Freedom of Information Acts § 408. (stating a requestor need not state a reason).

133. *McKinley v. F.D.I.C.*, 744 F. Supp. 2d 128 (D.D.C. 2010), judgment aff’d, 647 F.3d 331 (D.C. Cir. 2011).

134. 37A AM. JUR. 2D, Freedom of Information Acts § 1.

135. See 5 U.S.C.A. § 552(b)(1) (West 2006) (providing the categories of FOIA exemption).

The fact that predictive policing algorithms are privately-developed does not alone render FOIA inapplicable. Even government files originally created by private citizens often become subject to FOIA once in the possession of the government.<sup>136</sup> For example, in *TJS of New York, Inc. v. New York States Dept. of Taxation and Finance*, software sold to the government by a private contractor became a record of an agency. Upon a FOIA request, the court required the software shared with another private citizen over the objections of both the government agency and the government contractor.<sup>137</sup>

Although FOIA does not apply to data shared only with state and local governments,<sup>138</sup> each of the fifty states and many cities have their own open records law. State open records laws are modeled on the federal law<sup>139</sup> and share the same transparency goals as FOIA.<sup>140</sup> Given these similarities in language and purpose, courts draw on federal jurisprudence to interpret their state analogues.<sup>141</sup> Case law on FOIA is thus especially persuasive precedent in state decisions, and courts often draw comparisons across jurisdictions. Therefore, I incorporate both federal and state decisions to illuminate the likely outcome of open records requests for predictive policing information. Despite the need for transparency, I expect that business information and law enforcement exceptions to FOIA disclosure will impair the potential for openness with predictive policing.

---

136. See *U.S. Dept. of Justice v. Tax Analysts*, 492 U.S. 136, 144 (1989) (articulating a two-part for determining what constitutes an agency record).

137. *TJS of N.Y., Inc. v. N.Y. State Dep't of Taxation & Fin.*, 89 A.D.3d 239, 241 (2011) (ruling software acquired by the government subject to New York open records law).

138. Elizabeth O. Tomlinson, *Litigation Under Freedom of Information Act*, 110 AM. JUR. TRIALS 367 (2017) (“FOIA is a federal act and as such, it does not apply to state or local governments.”).

139. See, e.g., ALA. CODE § 36-12-40 (1975). But see David L. Ganz, *Open Public Records Act Litigation*, 128 AM. JURIS. TRIALS 495 (2017) (FOIA “imposes no affirmative obligation to unilaterally disclose information,” whereas state laws “permit or require disclosure by public authorities to the population as a whole.”).

140. See, e.g., *Thomas v. Hall*, 2012 Ark. 66, 399 S.W.3d 387 (2012); *ACLU Found. of S. Cal. v. Superior Court of L.A. Cty.*, 186 Cal. Rptr. 3d 746 (Cal. Ct. App. 2d Dist. 2015); *Cent. Fla. Reg'l Transp. Auth. v. Post-Newsweek Stations, Orlando, Inc.*, 157 So.3d 401 (Fla. 5th DCA 2015); *Office of the Governor v. Davis*, 122 A.3d 1185, 1191 (Pa. Commw. Ct. 2015); *Predisik v. Spokane Sch. Dist. No. 81*, 346 P.3d 737, 739 (Wash. 2015).

141. Leahy, *supra* note 130 (“Because the state acts usually have been modeled on the Federal FOIA, courts draw on the federal counterpart for judicial construction and legislative history.”).

B. *The Business Information Exception*

1. The Structure and Applications of the Exception

The business information exception to FOIA disclosure prevents required disclosures of certain business-related information.<sup>142</sup> Commercial law protects company profits by allowing companies to keep trade secrets, a doctrine which also applies to trade secrets shared with the government.<sup>143</sup> For example, the trade secrecy doctrine protects the design of privately developed voting machines.<sup>144</sup>

A number of non-physical assets fall under the category of trade secrets. The law protects business-related information if (i) there is economic value from the information not being generally known to other persons who could obtain economic value from its disclosure and (ii) there have been reasonable efforts to maintain its secrecy.<sup>145</sup> The law protects company formulas, compilations, programs, devices, methods, and processes.<sup>146</sup>

Applying the law to the voting machine example, there exist company designs for the physical machine and information about how to keep votes safe from cyber-attacks. There also exists a system for tabulating votes entered on the machines. A voting machine company profits from protecting that information from competitors who might, for example, use it to undercut their government contracts or develop improved machinery. To avoid these situations, the company may use passwords on its computers, release info only on a “need to know” basis, require employees or purchasers to sign confidentiality agreements, or issue verbal instructions to those they work with not to reveal the information.<sup>147</sup> As long as the voting machine company makes reasonable efforts to protect that

---

142. 5 U.S.C.A. § 552(b)(4) (West 2016).

143. RESTATEMENT (FIRST) OF TORTS § 757 cmt. B (AM. LAW INST. 1939) (describing trade secrets doctrine generally); see David S. Levine, *Secrecy and Unaccountability: Trade Secrets in Our Public Infrastructure*, 59 FLA. L. REV. 135, 147 (2007) (describing the protection of trade secrets of the government under the same commercial doctrine applied to trade secrets held by private actors). The Freedom Of Information Act exception also protects matters that are “commercial or financial information obtained from a person and privileged or confidential.” 5 U.S.C.A. § 552(b)(4) (West 2016).

144. *About*, DIEBOLD ELECTION SYSTEMS, <https://www.dieboldnixdorf.com/en-us/about-us> (last visited Sept. 8, 2018).

145. Uniform Trade Secrets Act § 1(4).

146. *Id.*

147. See *Liberty Am. Ins. Grp., Inc. v. Westpoint Underwriters, L.L.C.*, 199 F. Supp. 2d 1271 (M.D. Fla. 2001); *RKI, Inc. v. Grimes*, 177 F. Supp. 2d 859 (N.D. Ill. 2001); *Equifax Servs., Inc. v. Examination Mgmt. Servs., Inc.*, 453 S.E.2d 488 (Ga.

information, the trade secrets doctrine protects the company's information from falling into the hands of potential competitors.<sup>148</sup>

Given its importance to private industry,<sup>149</sup> the trade secrets doctrine has been incorporated into Exemption Four of FOIA, which protects "trade secrets and commercial or financial information obtained from a person [that is] privileged or confidential."<sup>150</sup> The exception technically covers two categories of information. The former provides protections under a more technical definition of trade secrets, while the latter builds upon the underlying policy of trade secrets doctrine to cover privileged and confidential commercial information generally. The "commercial or financial information" category is much broader and applies to "anything pertaining or relating to or dealing with commerce."<sup>151</sup>

State law reflects FOIA and the federal courts' interpretation of the exemption, including the protection of source codes as trade secrets.<sup>152</sup> For example, Illinois includes trade secrets and commercial information in the same two distinct categories.<sup>153</sup> On the whole, both federal and state open records laws allow the government, or a private contractor, to claim that an information request should be denied on IP grounds.

---

Ct. App. 1994); *Fred's Stores of Miss., Inc. v. M & H Drugs, Inc.*, 725 So. 2d 902 (Miss. 1998).

148. *See, e.g., Cisco Sys. Inc. v. Lynn*, No. 05-CV-03043 (N.D. Cal. Jul. 27, 2005).

149. *See Wesley M. Cohen, Richard R. Nelson, & John P. Walsh, Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)* 13 (Nat'l Bureau of Econ. Research, Working Paper No. 7552, 2000), <http://www.nber.org/papers/w7552.pdf> [<https://perma.cc/6HLZ-2LJP>] (finding among 1,478 manufacturing firms, secrecy ranked first or second in importance for product innovations).

150. 5 U.S.C. § 552(b)(4) (West 2016); *see also Critical Mass Energy Project v. Nuclear Regulatory Comm'n*, 975 F.2d 871, 879 (D.C. Cir. 1992) (holding voluntary submission to the government did not waive trade secrecy rights and trade secrets must only satisfy any one prong of the FOIA exception). Was this holding that any one of the prongs is enough or was it one specific prong?

151. *American Airlines, Inc. v. National Mediation Bd.*, 588 F.2d 863, 870 (S.D.N.Y. 2003).

152. *See generally Leahy, supra* note 130 (describing structure of state open records laws). Source codes are simply the building blocks of any computer program. Maybe add a little bit of background on source codes or an example of when a source code could be treated as a trade secret.

153. *See* 5 ILL. COMP. STAT. 140/7(1)(g)(ii) (2018). Illinois strays from the structure of the federal law by leaving agencies discretion over whether to disclose some trade secrets and confidential information, whereas FOIA mandates secrecy. *Id.*

Business information exceptions extend to types of information similar to predictive policing algorithms. These exceptions have protected source codes,<sup>154</sup> as well as computer software and proprietary technology.<sup>155</sup> Also, they have protected the structuring and informational basis for these technologies, including predictive technologies. Proprietary technology provided Illinois law enforcement with the rationale to deny disclosure of the ten variables used in a predictive policing algorithm.<sup>156</sup> There exists a substantial foundation in both federal and state court decisions for denying open records laws requests for information related to the IP of private government contractors. This foundation lays the groundwork for excluding predictive policing information from the public eye as well, stymying critical public oversight efforts.

## 2. The Rationale for Extending the Business Information Exception to Predictive Policing

Courts could extend the business information exception to predictive policing, but such extension would be detrimental to the public's ability to hold law enforcement accountable. In the next two sections, I explain how courts could arrive at the conclusion that predictive policing information should be kept secret before showing why that conclusion would be a mistake. Preventing the disclosure of information serves the interests of companies engaged in or likely to engage in contracts with the government.<sup>157</sup> Because private companies currently sell predictive policing technology to the government, the same interests apply. Given the government's emphasis on protections for small businesses, those commercial protections are arguably even more important for predictive policing companies, which are typically small.<sup>158</sup> Because the disclosure of information may place American companies at a competitive dis-

---

154. See, e.g., *GlobeRanger Corp. v. Software AG United States of Am.*, 836 F.3d 477 (5th Cir. 2016).

155. See 48 C.F.R. § 52.227-14(a) (defining for civilian acquisitions "data" as including computer software).

156. Davey, *supra* note 22.

157. See Nathan T. Nieman, Note, *Reforming the Illinois Freedom of Information Act: An Opportunity to Repair the Leaky Boat*, 58 DEPAUL L. REV. 529, 556 (2009) (arguing the Illinois FOIA ought to expressly notify parties whether their trade secrets will be disclosed to increase its protections for economic growth).

158. The Small Business Act in 1953 recognized the importance of small businesses to the United States economy. 15 U.S.C.A. § 631 (West 2010). Many predictive policing companies, like PredPol, are small businesses. PredPol, Inc., Notice of Exempt Offering of Securities (Form D) (June 16, 2014) (reporting a revenue between \$1 million and \$5 million). *But see* IBM Reports 2016 Fourth Quarter and Full-Year Results (Jan. 19, 2017), <https://www.ibm.com/investor/att/pdf/IBM->

advantage,<sup>159</sup> there is a purely economic justification for the trade secrets exception. Predictive policing companies are a part of that economy, so the exception applies.

Protecting the IP of predictive policing companies helps to ensure law enforcement can continue to obtain the technology by incentivizing market participation. Recall that predictive policing can provide substantial benefits to police departments, but it can only provide those benefits if private companies continue to develop and sell the technologies. Removing the ability of predictive policing companies to turn a profit could result in smaller companies losing the support of venture capital investors or larger companies substituting more profitable products for their current predictive policing projects.<sup>160</sup> Limitations on remedies further disincentivize investment.<sup>161</sup> More likely than exiting the market, private companies will raise their prices to account for the increased risk of government contracting.<sup>162</sup> This will result in larger outlays for the technology from law enforcement, greater opportunity costs, and fewer cities being able to afford the technology.

There is a final ramification of decreased expected profits that relates specifically to the market fragmentation of predictive policing.<sup>163</sup> The available capital of small companies is more likely to limit the development of predictive algorithms than for large companies.<sup>164</sup> Small companies can only design so sophisticated a model because they can only invest so much before making a sale. Decreasing IP protections for predictive policing companies en-

---

4Q16-Earnings-Press-Release.pdf [<https://perma.cc/ZL9U-WWZH>] (reporting annual revenue in 2016 of \$32.8 billion).

159. See Cohen, Nelson, & Walsh, *supra* note 149 (finding IP important for company innovation).

160. Nader Mousavi & Matthew J. Kleiman, *When the Public Does Not Have a Right to Know: How the California Public Records Act is Deterring Bioscience Research and Development*, 2005 DUKE L. & TECH. REV. 23, ¶ 29.

161. One such limitation on recovery is the inability to receive an injunction against the government for data disclosure. See 48 C.F.R. § 52.233-1. Another limitation is the doctrine of “waiver,” which requires the government to reveal information it would otherwise keep secret because it previously disclosed the information. See *Watkins v. U.S. Bureau of Customs & Border Prot.*, 643 F.3d 1189, 1198 (9th Cir. 2011).

162. See Mousavi & Kleiman, *supra* note 160 (suggesting companies may raise prices to account for unpredictable IP rights).

163. Cherry & Rogers, *supra* note 107.

164. The economic principles I discuss in this paragraph do not extend beyond the fundamentals and can be found in any introductory economics text. See e.g., N. GREGORY MANKIW, *PRINCIPLES OF ECONOMICS* 272–73 (2012) (discussing how economies of scale for larger companies can increase efficiencies and free up capital).

hances the' risk that they will start producing even less sophisticated models. The reduced protections incentivize smaller investments with faster returns. Less sophisticated models may not be inherently bad, but earlier I noted that many of the weaknesses of predictive policing relate to potential flaws in the models. A favorable economic environment, together with a demand for fairer models, could help start to resolve those concerns. Moreover, better models could promise greater effectiveness and efficiency in law enforcement, doubling down on the benefits of predictive policing.

In sum, limited IP protections weaken the market for the technology and potentially the quality of the technology itself. Because investments in predictive policing rely almost entirely on intellectual assets rather than physical assets, IP rights assume a determinative role in business decisions. With government contracting in predictive policing technology, the industry turns on the guarantee of these rights.

Even if there is a solid justification for strong IP concerns here, why do companies require secrecy to protect their IP? The trade secrets doctrine is not the only means of protection. Private companies could also rely on copyright or patent law.<sup>165</sup> However, each of these measures is unlikely to satisfy many predictive policing investors, especially if the investment is a more sophisticated predictive algorithm.<sup>166</sup>

IP rights present for private developers two additional hurdles. First, the rights attach automatically upon creation of the work, but protect "only the physical embodiment" not the "concepts or ideas underlying the work."<sup>167</sup> Much of what differentiates one predictive model from another are inventive variables and combinations of data. Even if a company does not copy another's predictive model verbatim, copying the underlying ideas behind the model could undercut a company's competitive advantage. Trade secrets, unlike copyrighted materials, do not require one to make "an exact or nearly exact copy of an algorithm or complication of information to

---

165. Nancy O. Dix et al., *Fear and Loathing of Federal Contracting: Are Commercial Companies Really Afraid to Do Business with the Federal Government? Should They Be?*, 33 PUB. CONT. L.J. 5, 11 (2003) ("An example of a product-specific issue is software, which in certain cases may be entitled to both patent and copyright protections.").

166. With both protections, private actors can only pursue damages in a suit against the government, not any injunctive relief. *See* Levine, *supra* note 143, at 178.

167. 17 U.S.C.A. § 102 (1976).

be liable . . . .”<sup>168</sup> Second, there are limits on copyright rights in government contracting. Government permission is often required both to copyright works first produced under a government contract and to include copyrighted work in deliverables to the government.<sup>169</sup>

Like copyrights, patents also fail to protect the underlying ideas of the predictive algorithms. Patents have been filed for predictive policing and arguably provide more appropriate protections than copyrighting.<sup>170</sup> Yet, patentees of predictive technology may still fear that their now public algorithms will be ripped off by competitors given the importance of discrete decisions and ideas within the models.<sup>171</sup> Developers of predictive technology also face a difficult choice *ex ante*. If they obtain a public patent for their technology, they must indefinitely rely upon the patent for their protection because they are no longer able to argue that the technology is a trade secret.<sup>172</sup> Trade secrets may be converted to patented material. Uncertainty of their future situation could drive predictive policing companies into the arms of trade secrecy law instead of patent law because doing so lends them the flexibility to change their mind and obtain a patent at a later date.

I am not arguing that these financial determinations are an adequate justification for keeping predictive policing secret. I am simply showing that the policy-focused FOIA determinations of courts could fit the algorithms under this exception of an open records law. Given the current situation of the predictive policing industry, private actors have the incentive to push for greater IP protections, and courts have a platform to recognize those interests. The current law provides both a committed advocate for secrecy and a legal ground from which to argue for that secrecy.

### C. *The Law Enforcement Exception*

#### 1. The Exception in General

Courts could also reduce transparency by fitting algorithms under law enforcement exceptions to open records laws. These exceptions protect “records compiled for law enforcement pur-

---

168. Michelle L. Evans, *Establishing Liability for Misappropriation of Trade Secrets*, 91 AM. JURIS. PROOF OF FACTS 3D 95 (2017) (referencing RESTATEMENT (FIRST) OF TORTS § 757 (AM. LAW INST. 1939)).

169. FAR 52.227-14

170. *See, e.g.*, U.S. Patent No. 8,949,164 (issued Feb. 3, 2015).

171. Dix et al., *supra* note 165, at 13–14.

172. *Id.*

poses.”<sup>173</sup> The exceptions define law enforcement broadly to include everything from surveillance tactics to immigration enforcement techniques and strategies to uncover tax fraud.<sup>174</sup> The underlying theme that connects these categories is an intent to enforce a law or regulation, and courts have denied the exception for any other purpose, pretextual rationale, or improper use.<sup>175</sup> The law remains grounded in the advancement of law enforcement, and so most circuits apply less exacting standards when considering exemptions for police departments than when considering the same exemptions for other agencies.<sup>176</sup>

A number of judicial and statutory limitations help to define the scope of law enforcement exceptions and capture their purpose. Statutes contain specific limitations on what constitutes a law enforcement purpose, and courts balance law enforcement justifications against public interests in disclosure.<sup>177</sup> Although the numbering and language of statutes vary, most hew closely to the limits provided in FOIA.<sup>178</sup> Sections 7(C) and 7(E) of the federal statute allow exemption under the law enforcement rule when the production of “. . . law enforcement records or information . . .

(C) could reasonably be expected to constitute an unwarranted invasion of personal privacy, . . .

(E) would disclose techniques and procedures for law enforcement investigations or prosecutions, or would disclose guidelines for law enforcement investigations or prosecutions if such disclosure could reasonably be expected to risk circumvention of the law . . .”<sup>179</sup>

7C protects personal information that law enforcement gathers and uses for legitimate purposes. 7E addresses information about the law enforcement strategies themselves. Although state codes may change the formatting, they generally reflect the same protec-

173. 5 U.S.C. § 552(b)(7)(A); *Jordan v. U.S. Dep’t of Justice*, 668 F.3d 1188 (10th Cir. 2011).

174. U.S. DEP’T. OF JUST. MANUAL U.S. ATT’YS MANUAL § 3-17.100 (describing potential application of law enforcement exceptions).

175. *Id.*

176. *See, e.g.*, *Puerto Rico v. U.S. Dep’t. of Justice*, 823 F.2d 574, 583–84 (D.C. Cir. 1987) (applying a rational nexus test).

177. This occurs especially in state court decisions. *See, e.g.*, *Chicago All. for Neighborhood Safety v. City of Chicago*, 808 N.E.2d 56, 73 (Ill. App. Ct. 2004); *ACLU of Oregon, Inc. v. City of Eugene*, 380 P.3d 281, 284 (Or. 2016).

178. *See Tomlinson, supra* note 138, at 367.

179. 5 U.S.C. A. § 552(b)(7) (2000). These are only two sections of several other possible law enforcement sections. I only mention these two because they are the exceptions most directly applicable to predictive policing.

tive goals. The limits to the FOIA law enforcement exception demonstrate the statute's commitment to ensuring an enforcement purpose. Courts further respect that purpose by limiting the exception temporally — the exception only applies so long as there is an active law enforcement proceeding or investigation that demands secrecy.<sup>180</sup>

## 2. The Exception's Applicability to Predictive Policing

Both 7C and 7E could apply to predictive policing. Exception 7C may apply to predictive policing because the algorithms generate data specific to individuals. Protection of personal information includes not just big data information like addresses, credit scores, incomes, and such, but also an individual's status as a target for additional surveillance. In *Union Leader Corp. v. U.S. Department of Homeland Security*, the Court prevented disclosure of a list of names because doing so would reveal the characteristics that put an individual on the list and the surveillance consequences for those individuals.<sup>181</sup> The court determined that the traits that merit police suspicion also merited privacy. To the extent that predictive policing can generate lists of individuals likely to be involved with crime, *Union Leader Corp.* would justify secrecy with the outputs of individual-based algorithms. Even if the determinative variables are harder to isolate with predictive policing than normal strategies, the stigmatizing impact from mere association with any law enforcement investigation constitutes a ground for applying Exception 7C.<sup>182</sup>

Although less apparent than individual-based predictions, an argument also exists for concealing information related to location-based predictive policing. Revealing location-based predictions may in effect disclose individuals or businesses residing within that location. The emphasis on stigmatization from *Union Leader Corp.* furthers this argument. For example, a business located within an area that an algorithm predicts will see more violent crime will understandably receive less business after that prediction is made public through a FOIA request. The argument that location-based predic-

---

180. *Citizens for Responsibility & Ethics in Washington v. U.S. Dep't of Justice*, 746 F.3d 1082, 1097 (D.C. Cir. 2014) ("Exception 7(A) is temporal in nature.").

181. *See Union Leader Corp. v. U.S. Dep't of Homeland Sec.*, 749 F.3d 45, 51–53 (1st Cir. 2014).

182. *Shapiro v. U.S. Dep't of Justice*, 34 F. Supp. 3d. 89, 95–97 (D.D.C. 2014) ("Exemption 7(C) recognizes that the stigma of being associated with any law enforcement investigation affords broad privacy rights to those who are connected in any way . . ."). Also note, individual-based arguments could also apply to targeted organizations, businesses, or advocacy groups.

tions can still embarrass, stigmatize, and expose to public scrutiny members of the community has particular application to precise predictive technologies that can make predictions as narrow as a 400 x 400 foot square.<sup>183</sup> Personal information concerns help to protect the inputs and outputs of algorithms from the disclosure under open records laws.

Exception 7E could serve to protect the algorithms themselves. Strategy exceptions can justify not disclosing information that may reach investigatory targets, but those justifications evaporate once the prediction-based surveillance ends and a new prediction takes its place. Because the same algorithm is used in current and future predictions, transparency before the surveillance or investigation ends provides criminals the means to figure out what the models will predict and circumvent law enforcement. Therefore, the predictions themselves have a limited lifespan of protection under 7E, while the algorithms and computational methods have an indefinite lifespan under 7E.<sup>184</sup>

The burden of proof placed on the government is easily met, which simplifies the task of keeping algorithms secret. The government need not prove disclosure *will* allow circumvention of the law, only that it *risks* circumvention.<sup>185</sup> In addition, the exception's broad language encompasses many law enforcement records. Guidelines include any "means by which agencies allocate resources for law enforcement investigations." Techniques and procedures include "the means by which agencies conduct investigations."<sup>186</sup> Between these two categories, protections extend as far as law enforcement manuals, policy guidance documents, settlement guidelines, monographs, and emergency plans.<sup>187</sup>

Precedent benefits agencies attempting to fit predictive policing algorithms under exceptions to protect law enforcement goals. The algorithms help determine the distribution of officers and constitute a procedure through which agencies conduct investigations. Moreover, precedent suggests investigations within the meaning of the FOIA exception need not be suspicion-based, which means pre-

---

183. PERRY ET AL., *supra* note 4, at 65 (explaining how precise algorithms predictions can be).

184. 7E protects techniques and procedures for law enforcement, which would include predictive policing techniques and the strategies implementing it. 7E also protects guidelines where transparency would risk circumvention of the law. Guidelines on how to determine resource allocation under predictive policing would fit under this second clause.

185. 5 U.S.C.A. § 552(b) (7)(E) (West 2016).

186. Tomlinson, *supra* note 138.

187. *Id.*

dictive algorithms are protected even though they are generally used at the surveillance stage of law enforcement action.<sup>188</sup> In *American Civil Liberties Union of New Jersey v. FBI*, an FBI racial mapping initiative was exempted because revealing the distribution of future surveillance efforts would interfere with later enforcement proceedings and reveal the targets of those efforts. Similarly, in *Showing Animals Respect and Kindness v. U.S. Department of the Interior*, documents on surveillance equipment used to locate and time potential poachers was exempted because it risked circumvention. These decisions lend credence to the idea that efforts to determine where to focus future surveillance can receive protection through exemption. The broad scope of this provision provides ample room for placing new law enforcement devices under its wing and preventing public oversight.

Additional grounding for the inclusion of predictive policing under Exception 7E comes from prior decisions including similar technologies. Previous cases have recognized the ability to exempt information on drones and drone use in surveillance measures from open records actions.<sup>189</sup> Proponents of secrecy can argue that revealing information about advanced technologies defeats law enforcement's ability to stay one step ahead of criminals and criminal organizations.<sup>190</sup> Such arguments have successfully exempted information systems and computer systems used for tracking movements of criminal activity, as well.<sup>191</sup> Unfortunately for transparency advocates, a sophisticated predictive algorithm would also likely fall within these exemptions.

---

188. *ACLU of New Jersey v. FBI*, 733 F.3d 526 (3d Cir. 2013); *Asian Am. Legal Def. & Educ. Fund v. New York City Police Dep't*, 964 N.Y.S.2d 888 (2013).

189. *Electr. Privacy Info. Ctr. v. Customs & Border Prot.*, 160 F. Supp. 3d 354, 261 (D.D.C. 2016) (granting in part and denying in part a FOIA request because the drone information could fit within Exemption 7(E) but the agency failed to provide sufficient justification for why the Exemption should apply in that particular case); *Citizens for Responsibility & Ethics in Washington v. U.S. Dep't of Justice*, 746 F.3d 1082, 1092 (D.C. Cir. 2014) (finding for the plaintiff on similar grounds).

190. Aff. of FBI Supervisory Special Agent Bradley S. Morrison, Chief, Tracking Technology Unit, Operation Technology Division in Quantico, Virginia, at 2, Apr. 11, 2014, attach. to City's Verified Answer, *Hodai v. City of Tucson*, No. C20141225 (Ariz. Super. Ct. Apr. 14, 2014).

191. *Skinner v. U.S. Dep't. of Justice*, 893 F. Supp. 2d 109, 112–13 (D.D.C. 2012) (properly withholding computer access codes for information systems); *Bishop v. U.S. Dep't. of Homeland Sec.*, 45 F. Supp. 3d 380, 389 (S.D.N.Y. 2014) (protection system for tracking individuals).

### 3. Reasonably Segregable Exception

Even where courts rule some information about predictive policing may be disclosed without risking circumvention of the law or invasions of personal privacy, the “reasonably segregable” requirement protects the rest of the information. Written into FOIA and state open records laws, the reasonably segregable requirement requires agencies to disclose portions of the record that do not fall within an exception when they can be separated from the rest of that record.<sup>192</sup> This means the public can gain access to some information that might otherwise be exempted, but also that courts avoid tough decisions for full disclosure by allowing partial disclosures. With as complicated a tool as predictive policing, partial disclosures will likely fail to provide sufficient insight to really understand how the models are operating within law enforcement.<sup>193</sup> Likewise, courts can require agencies to disclose techniques in only general terms without any specifics,<sup>194</sup> again making it difficult to complete a rigorous analysis of the predictive technology and its use. Partial disclosures have made external audits and accountability difficult in the context of other sophisticated law enforcement tools and would also allow only a shallow analysis of predictive policing.<sup>195</sup>

Over the last several decades, police departments have developed a culture of silence, preferring to keep information secret rather than disclose it to the public.<sup>196</sup> In the context of FOIA, even law enforcement agencies that favor transparency with predictive policing have incentives to deny open records requests and fight court orders. First, “voluntary disclosure in one situation can preclude later claims that records are exempt from release to someone

---

192. Tomlinson, *supra* note 138.

193. Chris J. Hoofnagle, *Big Brother's Little Helpers: How Choicepoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement*, 29 N.C. J. INT'L L. & COM. REG. 595, 598 (arguing multiple documents are necessary to provide context to understand FOIA responses); Stephanie K. Pell & Christopher Soghoian, *Your Secret Stingray's No Secret Anymore: The Vanishing Government Monopoly over Cell Phone Surveillance and its Impact on National Security and Consumer Privacy*, 28 HARV. J. L. & TECH. 1, 39 (2014) (arguing that heavily redacted documents defeat the utility of many FOIA requests).

194. Tomlinson, *supra* note 138.

195. *See supra* pp. 23–25.

196. *See, e.g.*, Jerome H. Skolnick, *Corruption and the Blue Code of Silence*, 3 POLICE PRAC. & RES. 7 (2002); Jeremy R. Lacks, Note, *The Lone American Dictatorship: How Court Doctrine and Police Culture Limit Judicial Oversight of the Police Use of Deadly Force*, 64 N.Y.U. ANN. SURV. AM. L. 391 (2008); Gabriel J. Chin & Scott C. Wells, *The “Blue Wall of Silence” as Evidence of Bias and Motive to Lie: A New Approach to Police Perjury*, 59 U. PITT. L. REV. 233 (1998).

else.”<sup>197</sup> A law enforcement agency may feel completely comfortable revealing predictive policing information to the ACLU with the understanding that an external review may improve legitimacy and effectiveness. However, the agency will not have the same positive effect for suspected members of criminal organizations, but still find itself bound to disclosure because of its previous cooperation with the ACLU’s request. Second, FOIA decisions create precedent for future disclosure or non-disclosure. With courts comparing technologies in FOIA decisions, police departments that put up little fight to the disclosure of predictive policing force themselves into a much greater conflict over the disclosure of the next big thing in policing, whatever it may be. Departments fear negative precedent for requests for predictive technology requests, as well as requests for whatever the next big thing in policing becomes.

Concerns over evasion of the law lends law enforcement a strong argument for fitting predictive policing information under the current exception. Predictive technologies are designed to help police departments determine where to send officers. If the public knows how those predictions are made, then criminal organizations can avoid those locations. Ensuring law enforcement can continue to function effectively is part of the rationale for balancing tests and the litany of complicated exceptions. Pure transparency is unlikely in this context, as experience with previous technologies demonstrates below.

#### *D. The Unlikelihood of Accessing Predictive Policing Information*

The public’s difficulty getting information with Stingrays exemplifies the quintessential problems the public will face getting information about predictive policing.<sup>198</sup> Stingray technology allows police to track individuals by tracking the movement of their cellphones. The cellphones ping off of cell towers installed with the stingray technology. Illinois law enforcement used Stingrays for years without the public being fully aware of that use. The Stingray story did eventually break, and attorneys applied for information under the state open records law. They waited for a response from law enforcement for months. Ultimately law enforcement delivered the information, but the documents were so heavily redacted much of the attorneys’ inquiries could not be answered. Meanwhile, police had begun to use the Stingrays to track protestors of the po-

---

197. *Lieber v. Board of Trustees of S. Ill. Univ.*, 680 N.E.2d 374 (Ill. 1997) (citing *Cooper v. U.S. Dep’t of the Navy*, 594 F.2d 484, 485–86 (5th Cir. 1979)).

198. Pell & Soghoian, *supra* note 193, at 34–39. My information about the stingray technology and this story come from the article by Pell & Soghoian.

lice.<sup>199</sup> Transparency was slow and half-hearted, and because of that fact, accountability was negligible.

Receiving information about predictive policing will likely be more difficult. Unlike Stingrays, where the concern is the implementation of the technology (for example, tracking police protestors), predictive policing raises concerns about the design of the technology. Predictive policing also presents a much more complex question. Information exists about inputs, algorithms, outputs, and implementation. Each of those steps involves a point where the public could demand public accountability and oversight. The number of steps also presents more opportunities for information to fall into an exception to open records laws. The complexity of predictive policing also makes it more difficult to know what to request from law enforcement. Asking: “Are you using Stingrays and for what purpose?” opens and closes the discussion. That same question is just a starting point with predictive policing, for the strategy also includes questions about algorithm design, inputs, and beyond.

In sum, the current legal framework provides little opportunity for substantial transparency with predictive policing.<sup>200</sup> Executive agencies, the judiciary, and the legislature have each proven themselves mediocre conduits of law enforcement information to the public. More importantly, relying on any government institution perpetuates the problems inherent in government-citizen informational gaps. Organizations, like civilian review boards, may better convey that information to the public and alleviate legitimacy concerns about creating an inner-circle of politicians and bureaucrats in the know. However, civilian review boards have yet to catch on, so the public’s primary opportunity for predictive policing information remains open records laws. Without transparency, there is little

---

199. *Id.* at 34–39.

200. There has been *very* little litigation on this subject matter. What litigation that has occurred has produced mixed results. In Los Angeles, a non-profit’s request for predictive policing information was denied. Brenda Gazzar, *Activists File Lawsuit over LAPD’s Predictive Policing Program*, GOV’T TECH. (Feb. 14, 2018), <http://www.govtech.com/public-safety/Activists-File-Lawsuit-Over-LAPDs-Predictive-Policing-Program.html>. Others, such as the Brennan Center in New York, have made more success. Rachel Levinson-Waldman & Erica Posey, *Court: Public Deserves to Know How NYPD Uses Predictive Policing Software*, BRENNAN CTR. FOR JUSTICE (Jan. 26, 2018), <https://www.brennancenter.org/blog/court-rejects-nypd-attempts-shield-predictive-policing-disclosure> [<https://perma.cc/4RAB-5JC9>] (gaining access through the courts to email correspondence, some historical output data, notes from the developer, a summary of results of trials on the products, and policies regarding the technology, but not receiving the algorithm itself).

opportunity for oversight over this new, rapidly evolving technology.

### CONCLUSION

Transparency has long been extolled as a requisite and virtue of democracy. Predictive policing's unique characteristics emphasize this need for transparency. Despite this heightened need, the momentum behind open records laws may sweep predictive policing under confidential business information and law enforcement exceptions. In this gap between normative ideals and positive expectations, exists an opportunity for practitioners to distinguish the unique need for transparency with predictive policing from other law enforcement methods. Establishing that precedent now lays a foundation for similar arguments in the continuously developing field of predictive analytics. While practitioners work to distinguish predictive policing to help craft precedent favoring transparency, lawmakers should turn their focus to how the language of open records laws fails to differentiate among technologies and consider creating exceptions for novel police technologies like predictive policing.

While total transparency may not be feasible with predictive policing, drawing the line with too blunt a tool threatens to leave the public in the dark about the realities of this new tactic. Given the growing coordination between the public and private sectors in areas like law enforcement, the increasing reliance on sophisticated technology, and disparities in information about those technologies, the challenge to transparency has grown more complex and the need for transparency more acute. Predictive policing demonstrates why attorneys cannot concede that difficult battle.



# END-TO-END AUTHENTICATION: A FIRST AMENDMENT HOOK TO THE ENCRYPTION DEBATE

LEONID GRINBERG\*

Introduction .....	174
I. A Technical Primer .....	180
A. Encryption .....	180
B. User Authentication .....	185
C. The Assisted MITM Attack .....	187
D. A Brief Digression on San Bernardino .....	190
E. In Summary .....	192
II. Authority for Issuing Such an Order .....	192
A. The All Writs Act .....	193
B. The Stored Communications Act .....	195
C. The Burr-Feinstein Legislation .....	197
III. The Rights of the Service Providers .....	198
A. The Exchange of Cryptographic Keys is Speech Under the First Amendment .....	199
B. Compelled Speech Triggers Strict Scrutiny .....	205
C. The Routine Use of an Order Compelling Attestations About Cryptographic Keys Cannot Survive Strict Scrutiny .....	208
D. Banning the Services Altogether .....	211
IV. The Rights of the Users .....	212
A. Service Providers Cannot Assert Users' Rights on a Fourth Amendment Theory .....	213
B. The First Amendment Does Not Protect Users ...	217
Conclusion .....	222

---

\* N.Y.U. School of Law, J.D., 2018; Massachusetts Institute of Technology, S.B., 2014; Editor-in-Chief, *Annual Survey of American Law*, 2017–18. The views expressed in this Note are solely mine and not my employer's. This Note was submitted in partial completion of the requirements of the Scholarship for Service, which has very generously supported my enrollment in law school. I would like to thank Zach Goldman and Randy Mitch for their leadership in the program during my time at NYU, and Zach Goldman in particular for his mentorship throughout this project. The *Annual Survey* has been my lifeline at NYU, and I would especially like to acknowledge the Notes Writing Program, Harry Black, and Brian Gottlieb for helping me develop this Note. I would also like to thank Ben Weissmann for insightful discussions in the early stages of this project. Finally, I would like to thank Beth Findley—for everything, really.

## INTRODUCTION

The past few years have seen a proliferation of messaging services offering “end-to-end encryption.”<sup>1</sup> The popularity of and demand for these services increased after Edward Snowden’s 2013 leaks, which revealed that communication service providers were cooperating with state surveillance efforts.<sup>2</sup> A conversation that is encrypted “end to end” is one in which only the participants have access to the messages in a readable form. When messages are in transit, no one, including the service provider itself, has access to the unencrypted content. By encrypting messages from end to end, a messaging service provider can assure its users that cooperation with the government would be impossible because the company does not have the ability to decrypt the messages. Only when a message reaches its intended recipient can it be read intelligibly.<sup>3</sup>

One of the most prominent examples of end-to-end encrypted messaging services is Apple’s iMessage,<sup>4</sup> which, as of 2016, processed 200,000 messages per second (which amounts to 63 quadrillion messages per year),<sup>5</sup> and has featured end-to-end encryption as a default setting since 2011.<sup>6</sup> There are plenty of other

---

1. Ken Kantzer, *Yet Another End-To-End Encrypted App*, PKC SECURITY (Dec. 16, 2016), <https://www.pkcsecurity.com/yet-another.html> [<https://perma.cc/NGS4-7LAN>] (“It seems that every week, yet another end-to-end encrypted app is unleashed on the world.”).

2. See generally Mark Mazzetti & Michael S. Schmidt, *Ex-Worker at C.I.A. Says He Leaked Data on Surveillance*, N.Y. TIMES (June 9, 2013), <http://www.nytimes.com/2013/06/10/us/former-cia-worker-says-he-leaked-surveillance-data.html>. Media outlets across the world have posted the original leaked confidential, starting with Snowden’s original trove. Much of this has been collected by the Electronic Frontier Foundation. See *NSA Primary Sources*, ELECTRONIC FRONTIER FOUND. <https://www.eff.org/nsa-spying/nsadocs> [<https://perma.cc/J7PF-CXXJ>].

3. Andy Greenberg, *Hacker Lexicon: What is End-to-End Encryption?*, WIRED (Nov. 25, 2014, 9:00 AM), <https://www.wired.com/2014/11/hacker-lexicon-end-to-end-encryption/> [<https://perma.cc/PV9Z-QEVB>].

4. As of iOS 10, the iMessage app was renamed as “Messages,” and “iMessage” became the name of a framework allowing third-party developers to add extension to the main app. See *iMessage Apps*, APPLE, <https://developer.apple.com/imessage/> [<https://perma.cc/LF37-NLWZ>]. For simplicity and the distinctive name, this Note uses the name “iMessage” to describe the messaging service.

5. Kif Leswing, *Apple Says People Send as Many as 200,000 iMessages per Second*, BUS. INSIDER (Feb. 12, 2016, 2:08 PM), <http://www.businessinsider.com/eddy-cue-200k-imessages-per-second-2016-2> [<https://perma.cc/YA7P-QE93>].

6. Press Release, Apple, *New Version of iOS Includes Notification Center, iMessage, Newsstand, Twitter Integration Among 200 New Features* (June 6, 2011), <http://www.apple.com/pr/library/2011/06/06New-Version-of-iOS-Includes-Notification-Center-iMessage-Newsstand-Twitter-Integration-Among-200-New-Features.html> [<https://perma.cc/ZC8A-FSL6>].

end-to-end encrypted messaging services available on the market, such as Facebook's WhatsApp, which boasts over one billion users of its own.<sup>7</sup> Facebook Messenger, which is built into the social networking website and is also available as a separate mobile application, also features end-to-end encryption, although it must be specially enabled by the user.<sup>8</sup> Consequently, a significant fraction of the world's population is communicating using end-to-end encryption, perhaps without even knowing it.<sup>9</sup>

Former FBI Director James Comey has referred to the proliferation in the use of end-to-end encrypted messaging services as the "going dark problem" and has argued that such services should be reined in.<sup>10</sup> Some commentators have breathed new life into a 1990s-era proposal called "key escrow," wherein the keys needed to decrypt a user's content would be preemptively given to the government but held "in escrow" until a valid court order grants access to

---

7. *One Billion*, WHATSAPP BLOG (Feb. 1, 2016), <https://blog.whatsapp.com/616/One-billion> [<https://perma.cc/38N9-E98L>].

8. See Gail Kent, *Hard Questions: Why Does Facebook Enable End-to-End Encryption*, FACEBOOK NEWSROOM (May 7, 2008), <https://newsroom.fb.com/news/2018/05/end-to-end-encryption/> [<https://perma.cc/WF6M-BPK2>].

9. There is no single reason why the demand for end-to-end encryption is so high, even among the law-abiding. Some consumers distrust the government, either on principle or after news of some abuse. Reports about increased surveillance on mosques, for example, likely provided incentives for some Muslims to encrypt their communications to avoid harassment by the government. Cf. Lee Mathews, *Encrypted Email Signups Skyrocketed After Trump Victory*, FORBES (Nov. 14, 2016, 4:39 PM), <https://www.forbes.com/sites/leemathews/2016/11/14/encrypted-email-signups-skyrocketed-after-trump-victory> [<https://perma.cc/7PKE-BPYK>]. Others may be worrying not about the government but about hackers who could steal their personal or embarrassing communications from their service provider. In 2014, for example, a hacker accessed iCloud and Gmail accounts belonging to celebrities and posted nude photographs of them online. Tom Sykes, *Celebrity Nude Photo Hack: Images of Miley Cyrus, Kristen Stewart, Tiger Woods and More Leak Online*, THE DAILY BEAST (Aug. 22, 2017, 4:34 AM), <https://www.thedailybeast.com/celebrity-nude-photo-hack-images-of-miley-cyrus-kristen-stewart-tiger-woods-and-more-leak-online> [<https://perma.cc/4FH9-TRY5>]. Anyone who had stored only encrypted copies of the images would not have been affected by the hack.

10. See, e.g., Adam Mazmanian, *Comey Renews Encryption Plea on Capitol Hill*, FCW (July 8, 2015), <https://fcw.com/articles/2015/07/08/comey-encryption-hearing.aspx> [<https://perma.cc/6RX9-L5HN>] ("The problem, as Comey sees it, is that criminals, terrorists and other malefactors are 'going dark,' by using end-to-end encryption built into mobile device operating systems offered by Apple and Google, and available in some communications software, like WhatsApp."); see also *Going Dark*, FBI, <https://www.fbi.gov/services/operational-technology/going-dark> [<https://perma.cc/4BX6-JDAQ>]; David E. Sanger & Sheera Frenkel, "Five Eyes" Nations Quietly Demand Government Access to Encrypted Data, N.Y. TIMES (Sept. 4, 2018), <https://www.nytimes.com/2018/09/04/us/politics/government-access-encrypted-data.html> [<https://perma.cc/AGQ3-4MJT>].

the key on a case-by-case basis.<sup>11</sup> Many civil rights and technical groups have responded negatively to this proposal<sup>12</sup> (just as they did two decades ago<sup>13</sup>). One of the most common critiques is that the government would be unable to adequately protect the billions of decryption keys it was holding “in escrow” against theft by criminals or rival state governments.<sup>14</sup>

The reality, however, is that key escrow would probably not be an effective method of solving the “going dark problem.” Any two actors communicating on an uninterrupted, public connection can establish an encrypted channel of communication using simple, well-known algorithms, even in the presence of eavesdroppers.<sup>15</sup> Of course, the government could ban the creation of such software, but the simplicity and ubiquity of the code in computer science textbooks would undermine the ban.<sup>16</sup> Furthermore, history has shown that similar efforts have been unsuccessful. The Digital Millennium Copyright Act, for example, has done little to lessen the availability of file-sharing applications and websites, notwithstanding their illicit status.<sup>17</sup>

Therefore, the government will probably look for other ways to gain access to encrypted messages. In light of that, I propose focusing the discussion on *user authentication*—that is, technology that

11. See, e.g., Ellen Nakashima & Barton Gellman, *As Encryption Spreads, U.S. Grapples with Clash Between Privacy, Security*, WASH. POST (Apr. 10, 2015), [https://www.washingtonpost.com/world/national-security/as-encryption-spreads-us-worries-about-access-to-data-for-investigations/2015/04/10/7c1c7518-d401-11e4-a62f-ee745911a4ff\\_story.html](https://www.washingtonpost.com/world/national-security/as-encryption-spreads-us-worries-about-access-to-data-for-investigations/2015/04/10/7c1c7518-d401-11e4-a62f-ee745911a4ff_story.html) [<https://perma.cc/4L98-PNRH>].

12. E.g., HAROLD ABELSON ET AL., KEYS UNDER DOORMATS: MANDATING INSECURITY BY REQUIRING GOVERNMENT ACCESS TO ALL DATA AND COMMUNICATIONS, (2015), <https://dspace.mit.edu/handle/1721.1/97690> [<https://perma.cc/86QA-ATA4>].

13. E.g., HAL ABELSON ET AL., THE RISKS OF KEY RECOVERY, KEY ESCROW, AND TRUSTED THIRD-PARTY ENCRYPTION (1997), <https://doi.org/10.7916/D8GM8F2W> [<https://perma.cc/Z2DU-2QMP>].

14. See, e.g., ABELSON ET AL., *supra* note 12, at 15 (“Providing access over any period of time to thousands of law enforcement agencies will necessarily increase the risk that intruders will hijack the exception access mechanisms. . . . [T]he challenge of guaranteeing access to multiple law enforcement agencies in multiple countries is enormously complex. It is likely to be prohibitively expensive and also an intractable foreign affairs problem.”).

15. See Greenberg, *supra* note 3. See generally *infra* Part I.

16. See, e.g., Aaronson, *infra* note 100.

17. See, e.g., Koren Helbig, *11 Numbers That Show How Prolific Illegal Downloading Is Right Now*, PRI (Apr. 20, 2014, 12:43 PM), <https://www.pri.org/stories/2014-04-20/11-numbers-show-how-prolific-illegal-downloading-right-now> [<https://perma.cc/DPU9-93SY>] (reporting that Americans illegally downloaded more than 96.8 million songs in 2012, more than any other country).

verifies for each user that the messages she reads come from the counterparty in the conversation—instead of focusing it on keys. As noted above and explained in detail below, any two individuals speaking on a reliable connection can establish an encrypted channel. Authentication is the technology that provides that reliability<sup>18</sup> through the use of keys that are known to be associated with a particular user. Thus, authentication is sufficient to establish an end-to-end encrypted conversation. But, though sufficient, it is also necessary. After all, for any conversation to remain confidential, each party must know *with whom* she is speaking. It does no good for the sender to encrypt a message for the recipient if the key used for the encryption actually belongs to a third party. End-to-end encryption could be effectively circumvented by any eavesdropper who could get his or her own key inserted into the users' encryption algorithms (a technique called a "man-in-the-middle attack"). If there were no reliable way for distributing keys and matching them to particular identities, this would be a very effective way for the government to listen in on encrypted conversations. The government could simply supply its own keys to a service provider and convince (or compel) it to claim that the keys are associated with a target user.

I am not aware of any case in which the government has actually replaced one user's keys with its own, but the possibility has been raised many times in the past, in both popular and technical literature.<sup>19</sup> In the case of iMessage, for example, even though Ap-

---

18. See generally *infra* Part I.B. By "reliability" I mean the abstract notion that messages are not tampered with or forged by others. The physical integrity of the channel—the audio quality of a phone conversation, for example—is assumed.

19. See, e.g., Manisha Ganguly, *WhatsApp Design Feature Means Some Encrypted Messages Could Be Read by Third Party*, THE GUARDIAN (Jan. 13, 2017, 6:00 AM), <https://www.theguardian.com/technology/2017/jan/13/whatsapp-design-feature-encrypted-messages> [<https://perma.cc/3H6K-HVDF>] (describing the potential for the vulnerability in WhatsApp); *Apple's iMessage Defense Against Spying Has One Flaw*, WIRED (Sept. 8, 2015, 1:10 PM), <https://www.wired.com/2015/09/apple-fighting-privacy-imessage-still-problems/> [<https://perma.cc/L43K-KGZN>]; Greenberg, *supra* note 3 ("Even end-to-end encryption isn't necessarily impervious from snooping. Rather than try to actually break the encryption, for instance, an eavesdropper may try to impersonate a message recipient so that messages are encrypted to their public key instead of the one the sender intended. After decrypting the message, the snoop can then encrypt it to the recipient's actual public key and send it on again to avoid detection."); Cyril Cattiaux et al., *iMessage Privacy*, QUARKSLAB (Oct. 17, 2013), <http://blog.quarkslab.com/imessage-privacy.html> [<https://perma.cc/4W63-U3X5>]; Matthew Green, *Can Apple read your iMessages?*, CRYPTOGRAPHY ENGINEERING (June 26, 2013), <https://blog.cryptographyengineering.com/2013/06/26/can-apple-read-your-imessages/> [<https://perma.cc/2PL3-BJJY>].

ple never has access to unencrypted messages, nor the keys needed to decrypt encrypted messages, there is nothing stopping Apple from configuring its system to misreport the party with whom a user is speaking and encrypting a message so as to be readable by a third party. That third party could be the government.

It is not an accident that so much ink has been spilled on what seems like a hypothetical problem. The concern is not just about encryption. After all, authentication is useful for many applications, many having nothing to do with end-to-end encryption. Proving that one is who he or she claims to be is central to the Internet's functionality. Users would not be comfortable sending intimate emails or texts, or for that matter logging into a bank's website, if they were not confident that they were communicating with their lover or their bank. The government would be tampering with the technology that makes that confidence possible, which I am broadly referring to as "authentication" (and in the case of consumer messaging services, "user authentication").<sup>20</sup> And yet, as mentioned earlier, a reliable channel, with authenticated participants in the conversation, is sufficient on its own to create a confidential channel. In other words, unlike key escrow, which is widely resisted and is unlikely to work, this method is both very effective and has far-reaching consequences.

I refer to this hypothetical government technique as an "assisted man-in-the-middle (MITM) attack" and the corresponding hypothetical court order compelling assistance by the service provider as an "assisted MITM order." As mentioned above and explained further in Part I, a "man-in-the-middle attack" is a hacking technique in which an adversary intercepts communications and tampers with them in some way. In this context, the "adversary" (from the perspective of those wishing to keep their messages private) is the government and its "attack" is "assisted" by the service provider. As Part III emphasizes, the nuances of the roles the service provider and the government play are central to the discussion: the assistance from the service provider is not simply technical, but also communicative—it facilitates the process by placing its imprimatur on the communications provided by the government, making the communications seem as if they are coming from the service provider itself.

As cited earlier, there is plenty of prior technical work acknowledging the concerns about an assisted MITM attack. This paper explores the issue through a legal lens. Though very effective, an

---

20. See generally *infra* Part I.B.

assisted MITM attack is likely to have constitutional ramifications. An assisted MITM order implicates the compelled speech doctrine because it would require service providers to display and send particular messages chosen by the government, thus potentially violating the service providers' First Amendment rights. And because key distribution affects not just encryption but also authentication, there is a broader public policy concern—one that has less to do with conventional notions of privacy and more with identity. In an era in which an increasingly large percentage of communications take place over a network, anyone texting, emailing, or posting on social media has a higher stake in being able to digitally prove who they are to their audiences. The law has not yet evolved to recognize this need. Indeed, as this paper argues, assuming the Fourth Amendment is satisfied, *users'* constitutional rights are unlikely to be violated if the government acted in this way—even in circumstances where the users seemingly have strong free speech interests. This cutting-edge context raises novel questions as to whether the First Amendment *should* be extended to recognize such rights—a “right to integrity of identity,” if you will—which heretofore have mostly been protected only, if at all, by common law and statutory personality rights.<sup>21</sup>

This Note proceeds in four Parts. Part I is a technical primer on encryption and authentication technologies. It explains in basic terms how the technologies work, describes iMessage (as an operative example) in some detail, and explains how the hypothetical surveillance technique would operate. It also briefly explains the relationship between the technologies and issues discussed in this Note and those in the controversy between Apple and the FBI following the San Bernardino terrorist attacks in 2015. Part II argues that there is existing statutory authority for the government to issue an assisted MITM order under the All Writs Act and the Stored Communications Act. It also highlights some legislation introduced after the Apple-FBI controversy. Part III argues that irrespective of the source of authority, the order would violate service providers' First Amendment rights. Part IV shifts the focus to the users and analyzes their interests. It concludes that absent a claim of injury by the service provider, such techniques would likely be constitutional and without remedy to the users. It suggests, however, that it may be worth considering an expansion of privacy rights in some in-

---

21. *See, e.g.*, *Toney v. L'Oreal USA, Inc.*, 406 F.3d 905, 908–09 (7th Cir. 2005) (contrasting federal copyright in a given image with a state statute providing for a right of publicity, defined as “the very identity or persona of the plaintiff as a human being”).

stances into a constitutional realm—motivated by the First, rather than the Fourth Amendment.

## I. A TECHNICAL PRIMER<sup>22</sup>

The issues discussed in this Note do not require any technical knowledge beyond familiarity with some common cryptographic concepts and terminology. This primer supplies that background.

### A. *Encryption*

A fundamental building block of most cryptographic operations is the “*key*.” A key is just a number, typically a very big one,<sup>23</sup> that is used as an input into some cryptographic process. In the case of “*encryption*,” the key is used to “scramble” the message. “*Decryption*,” the reverse operation, uses a key to “unscramble” the message. There are many different ways to implement an encryption scheme, some of which use the same key for both the encryption and decryption steps (this is called “*symmetric encryption*”) and some of which use different but mathematically related keys (“*asymmetric encryption*”).

Although end-to-end encryption in messaging services requires asymmetric encryption (the reason for that will be evident in a moment), basic symmetric encryption is a useful starting place for understanding encryption technology more generally. Thus, as a first, concrete example, consider a simple cipher in which the sender shifts the letters in her message by a predetermined amount. For

---

22. Because cryptography is such a vast subject whose technical details would not be of interest to most readers, I have chosen not to individually cite many of the general claims made in this primer. Instead, I direct interested readers to an excellent textbook by Professor Michael Sipser. See generally MICHAEL SIPSER, INTRODUCTION TO THE THEORY OF COMPUTATION ch. 10.6 (3d ed. 2013).

23. For example, the popular Advanced Encryption Standard, issued in 2001 by the National Institute of Standards and Technology, uses keys of at least 128 bits. NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, ANNOUNCING THE ADVANCED ENCRYPTION STANDARD (AES) (2001), <https://csrc.nist.gov/csrc/media/publications/fips/197/final/documents/fips-197.pdf> [https://perma.cc/FD3Z-FEAT]. A key length of 128 bits translates to approximately 39 digits. One of the encryption keys used in iMessage (there are several layers of encryption involved) is 1280 bits, APPLE, iOS SECURITY 60 (2018), [https://www.apple.com/business/docs/iOS\\_Security\\_Guide.pdf](https://www.apple.com/business/docs/iOS_Security_Guide.pdf) [https://perma.cc/JQ39-9GS7], which translates to approximately 887 digits. For comparison, the number of atoms in the observable universe has approximately 80 digits. John Carl Villanueva, *How Many Atoms Are There in the Universe?*, UNIVERSE TODAY (July 30, 2009), <https://www.universetoday.com/36302/atoms-in-the-universe/> [https://perma.cc/RUL5-J7YF].

example, suppose the message is “HELLO” and the letters are to be shifted by a distance of 5. The resulting message is “MJQQT” (because “H” is the eighth letter in the alphabet and “M” is the thirteenth, “E” is the fifth letter and “J” is the tenth, and so on). The original “HELLO” message is called the “*plaintext*” whereas the incomprehensible “MJQQT” message is called the “*ciphertext*.” The key used in this operation is “5”—the amount by which the letters were shifted. The shifting operation is the “encryption” step. After the recipient receives the ciphertext, he can shift the letters back down (the “decryption” step) by the key to recover the plaintext. The complete sequence of steps—the sender and recipient agree on a key, the sender shifts up every letter in the plaintext by the key to create the ciphertext, the sender transmits the ciphertext, the recipient shifts down every letter in the ciphertext by the same key to recover the plaintext—is called an “*encryption scheme*.”

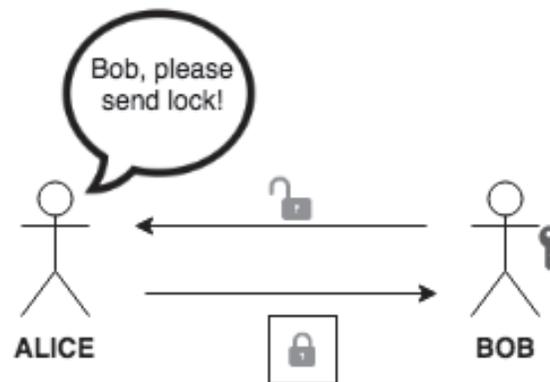
Apart from being easy to circumvent by just guessing every number between 0 and 25 (a weakness that can be mitigated by variations in how the key is used<sup>24</sup>), symmetric encryption schemes suffer from a fundamental limitation—the sender and receiver need to have agreed on a key in secret prior to sending the message. This is not always possible. For example, it would be impractical to encrypt a connection between a web browser and a website (say, Amazon.com) in this way because there is no way for Amazon and every one of its customers to agree on a secret key prior to making the connection. Without encrypting the connection, however, it would be impossible for customers to securely shop online with their credit cards. Anyone intercepting the connection would have access to the sensitive information.

Such encrypted connections can be established through asymmetric encryption schemes. These schemes are also called “*public key*” encryption schemes because one of the keys—the one used for encryption—is made public to the world. The decryption key is mathematically related to but separate from the encryption key and is kept private. The public and private keys together are called the “*key pair*.” A sender uses the recipient’s publicly available encryption key to encrypt her message, and the recipient decrypts it with his decryption key, which he has kept to himself.

---

24. See Ron Rivest, *6.857 Lecture 2*, MIT 1–3 (Sept. 9, 1997), <http://web.mit.edu/6.857/OldStuff/Fall97/lectures/lecture2.pdf> [<https://perma.cc/EB9T-YSG6>] (discussing “one-time pad,” a generalized version of this technique that is a theoretically “unbreakable” form of encryption, at least when unlimited computing resources are available).

Although it is hard to provide a mathematically simple version of such a scheme by way of example, a physical analogy can help. Imagine that Alice wants to send a package to Bob<sup>25</sup> but cannot meet with him ahead of time to exchange keys to a shared lock. Instead, she calls Bob and asks him to send her a padlock to which only Bob has the key, but with the padlock in an unlocked state. When Alice receives the lock, she puts her package in a box and closes the padlock. She can't unlock the padlock at this point because she doesn't have the key. When she sends it, anyone who intercepts the package is also unable to open it. Only Bob, once he receives the package, can open the padlock and retrieve the package. Figure 1, below, illustrates the mechanics of this scheme.



**Figure 1:** Illustrating how Alice can send Bob a secure package without prearranging a shared key.

Notice that in this scheme, the package remains secure even if someone is opening and inspecting (but not tampering with) Alice and Bob's mail. All that an interceptor sees initially is an unlocked padlock and then, later, a box locked with that padlock. At no point

25. For some reason, cryptography literature ubiquitously labels the sender as "Alice" and the recipient as "Bob." I adopt the convention here and take advantage of the disambiguation in the gendered pronouns. According to one source, the first use of the name was in a paper by Ron Rivest, Adi Shamir, and Leonard Adleman, the inventors of the classic "RSA cryptosystem." See Quinn Dupont & Alana Cattapan, *Alice & Bob, A History of the World's Most Famous Cryptographic Couple*, <http://cryptocouple.com/Alice%20and%20Bob%20-%20DuPont%20and%20Cattapan%202017.pdf> [<https://perma.cc/V22P-L5VG>]. Despite its age, RSA is still core to the iMessage encryption system (among many others). See APPLE, *supra* note 23, at 60.

does either the key or the unlocked box go through the mail.<sup>26</sup> Although the devil will always be in the technical details, this is the basic reason that two parties speaking to each other can establish a secure conversation even in the presence of an eavesdropper.

A typical messaging service offering “end-to-end encryption” looks a lot like this, but it has an additional feature: the service offering the transportation of messages is the same one that effectuates the encryption. To match the physical analogy to that arrangement, imagine that instead of Alice calling Bob on some separate channel (like a phone call) to request a padlock, FedEx were to preemptively give each of its customers a key to a unique padlock that it kept in storage and promised not to keep copies of the keys. FedEx would keep the padlocks in an unlocked state. When Alice wants to send Bob a package, she merely asks FedEx to put on “Bob’s” padlock. When Bob receives his package, he can unlock the padlock with the key that only he has and give the unlocked padlock back to FedEx. As long as FedEx keeps its word about attaching the right padlock and not retaining copies of the keys, Alice and Bob can be sure that neither FedEx, nor anyone intercepting FedEx packages, will see the inside of the package.

This revised, fanciful arrangement is essentially how modern messaging services that rely on end-to-end encryption work. The service provider facilitates not just the transportation between user accounts (which, in many cases, is an account system that the service provider itself controls), but also manages the keys for its users. The private decryption keys are kept safe on the users’ devices without the service provider retaining or having access to them. When one user wants to send an encrypted message to another, she asks the service provider for the recipient’s public encryption key. Having encrypted the message, she sends the ciphertext over to the user, who uses his private decryption key to recover the plaintext.<sup>27</sup>

---

26. Of course, in the physical world, it may be possible to construct a key just from careful inspection of a padlock, but that is not the case with the mathematical system after which this analogy is modeled. Similarly, the assumption is that it is impossible (or at least intractably difficult) to simply “guess” the mathematical key in a way where one could pick a physical lock. *But see infra*, note 27 (explaining that this assumption is not proven).

27. Asymmetric encryption schemes come with a massive disclaimer, and this footnote is as good a place for it as any: they are not proven to be secure. Many different asymmetric encryption algorithms have been proposed over the years, some of which have shown themselves to be convincing enough to researchers that trillion-dollar industries have been built on the assumption that they cannot be defeated in any practical amount of time. And yet, to date, no one has been able to prove that no clever algorithm exists that can quickly recover a plaintext from a

Different implementations of end-to-end encryption schemes work differently in practice, but this illustration roughly describes them all. One important special case, particularly relevant to this Note, is Apple's iMessage. There, a user has a different key pair associated with each device she owns.<sup>28</sup> Thus, a user's iPhone, iPad, and MacBook all have their own decryption keys stored privately on each respective device, with the corresponding public encryption keys stored by Apple. When Alice wishes to send a text message over iMessage to Bob, Apple sends *all* of Bob's encryption keys to Alice. Alice's device sends multiple copies of the message—one for each device's encryption key—and Apple forwards them to Bob's corresponding devices for decryption.<sup>29</sup> That is how it is possible for Bob to see incoming messages on all of his devices, but why a newly purchased device will not show old messages (unless it was restored from a backup of an old device). Figure 2, below, illustrates how end-to-end encryption works on Apple's iMessage in the case where Bob owns just one device.

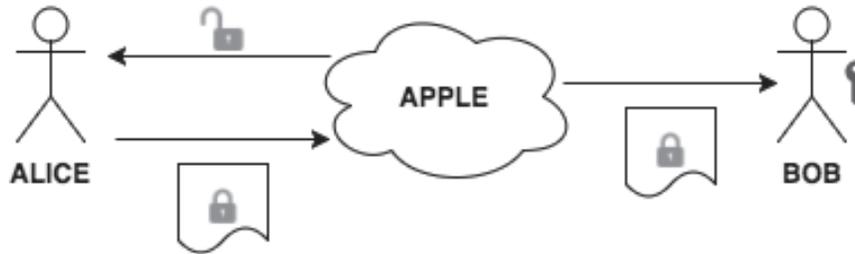
---

ciphertext encrypted with an encryption key in an asymmetric encryption scheme. Furthermore, it *has* been shown that quantum computers, if they can ever be reliably built at scale, can efficiently break a subset of commonly used asymmetric encryption schemes, including, notably, RSA (*see supra* note 25). *See generally* Stephanie Blanda, *Shor's Algorithm—Breaking RSA Encryption*, AMS BLOGS (Apr. 30, 2014), <https://blogs.ams.org/mathgradblog/2014/04/30/shors-algorithm-breaking-rsa-encryption/> [<https://perma.cc/FDB2-XED7>]. There is a million-dollar bounty for a mathematical solution to a highly generalized version of this open problem—a figure that is laughably small compared to the economic ramifications of a proof that asymmetric encryption schemes do not, in fact, work. *See generally P vs NP Problem*, CLAY MATHEMATICS INST., <http://www.claymath.org/millennium-problems/p-vs-np-problem> [<https://perma.cc/5C3H-TG3L>]. In fact, the so-called “P-vs-NP problem” has ramifications that extend far beyond the realm of cryptography, including into logistics, *id.*, and even far-flung fields like the drawing of voting districts, Michah Altman & Michael McDonald, *The Promise and Perils of Computers in Redistricting*, 5 DUKE J. CONST. L. & PUB. POL'Y 69, 81 (2010).

In any case, this Note proceeds, as the rest of the world has, on the assumption that at least some version of asymmetric encryption schemes in use today does work and is not, crudely put, “easily breakable,” at least if correctly implemented. Incidentally, it is generally understood that correct implementations are a big “if,” and that in almost all cases, bugs in the implementation are what tend to cause security vulnerabilities rather than some fundamental mathematical flaw in the underlying encryption scheme. *See, e.g.*, Bruce Schneier, *Security Pitfalls in Cryptography*, SCHNEIER ON SECURITY (1998), [https://www.schneier.com/essays/archives/1998/01/security\\_pitfalls\\_in.html](https://www.schneier.com/essays/archives/1998/01/security_pitfalls_in.html) [<https://perma.cc/D97M-DPDL>] (describing a variety of attack vectors having nothing to do with encryption).

28. *See* APPLE, *supra* note 23, at 60.

29. *Id.*



**Figure 2:** Illustrating how Alice can send Bob an encrypted message so Apple can't read it. The lock represents Bob's public encryption key and the key represents his corresponding private decryption key. The locked message indicates that the message is encrypted. Bob can decrypt the message with his private key when he receives it.

The next Section describes authentication, a sort of counterpart to encryption.

### B. User Authentication

Encryption and decryption are not the only cryptographic operations. There are also a large number of operations that broadly fall into the category of "authentication." Unlike encryption, which is designed to ensure *secrecy*, authentication is designed to ensure *integrity*. Authentication technologies, among other things, enable senders to digitally "sign" their messages and allow recipients to "verify" those signatures. These signatures can provide various assurances, including that the message was not modified after it was sent and that the message was sent by a particular sender. This Note is primarily concerned with the latter functionality, which I call "*user authentication*."<sup>30</sup>

User authentication can be effectuated through asymmetric keys, just like encryption. In that context, the key that remains private is used for signing, while the key that is publicly disclosed is

30. The concept of signing and verifying messages is not limited to end-to-end communication services. Websites, for example, use the same technology to assure web browsers that they are operated by a particular real-life entity. See generally *Certificate Authorities & Trust Hierarchies*, GLOBALSIGN, <https://www.globalsign.com/en/ssl-information-center/what-are-certification-authorities-trust-hierarchies/> [<https://perma.cc/SS57-LTME>]. Furthermore, the exact same technology is used by Apple to protect against unauthorized updates to the iPhone operating system, which was the technology at issue in the San Bernardino controversy. See *infra* Part I.D. I chose the term "user authentication" to emphasize that this technology is what enables users to trust their counterparties' reported identities, and to distinguish from other similar authentication-related operations. It is not a term of art used in the field generally.

used for verifying. Although the standard terms, “private key” and “public key,” are also used in the context of authentication schemes, I will use the non-standard terms, “signing key” and “verification key,” for clarity.

As noted in the introduction, end-to-end encryption services would be of little use if the users did not know with whom they were speaking.<sup>31</sup> For this reason, user authentication is built into them as well. When Alice wishes to send a message to Bob, she not only encrypts her message with Bob’s publicly available encryption key, but also signs it with her private signing key. When Bob receives the message, he first verifies the signature with Alice’s publicly available verification key and then decrypts it with his private decryption key. As long as the users are confident that the publicly available keys in fact belong to the users they purport to belong to (or, more precisely, that the corresponding private keys were solely in the possession of those users), both the confidentiality and integrity of the system are assured.

To illustrate, suppose Alice and Bob are using iMessage to communicate. Apple maintains both users’ keys in a large directory. When Alice requests Bob’s key(s), she gets back a message that, in simplified form, looks something like Figure 3:

---

31. For this reason, I say that user authentication is necessary for end-to-end encryption to work. As the previous Section demonstrates, it is also “sufficient” because asymmetric encryption schemes enable users to create a private communication channel while speaking publicly. Thus, given a public authenticated channel (that is, where everything said is public but attributable to the speakers), asymmetric encryption can be used to enable a private conversation. In fact, as it turns out, researchers have proposed ways to maintain confidential conversations in public without using encryption (or keys) at all. One famous proposal is called “Chaffing and Winnowing.” See Ronald L. Rivest, *Chaffing and Winnowing: Confidentiality without Encryption*, MIT LAB FOR COMPUT. SCIENCE (Mar. 18, 1998), <http://people.csail.mit.edu/rivest/chaffing-980701.txt> [https://perma.cc/ZPB9-894N]. Professor Ron Rivest proposed this scheme in the context of the 1990s key escrow debate. See *supra* notes 11–14 and accompanying text. However, the Chaffing and Winnowing scheme still assumes an authenticated channel because, as noted earlier, a conversation cannot be truly confidential without the parties being sure about each others’ identities.

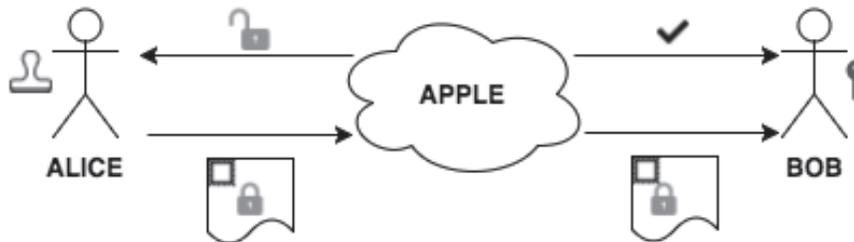
```

{
  user: "Bob",
  pubkeys: [
    "12345",
    "67890"
  ]
}

```

**Figure 3:** Illustrating a simplified example of a key transmission for user “Bob.” In this example, Bob has two public keys, “12345” and “67890.” These public keys are associated with two private keys, each of which is stored on a device (e.g., Bob’s iPhone and iPad) and not shared with anyone else, including Apple.

The security of iMessage depends on the integrity of that directory—that is, on the assurance that the public keys listed for a particular Apple account are in fact associated with the owner of that account. The full scheme is illustrated in Figure 4, below.



**Figure 4:** Illustrating the added authentication layer on top of the encryption layer. The stamp represents Alice’s private signing key and the checkmark represents her corresponding public verification key. The stamped and locked message indicates that the message is encrypted with Bob’s public encryption key and signed with Alice’s private signing key. When Bob receives the message, he can verify it with the public verification key and decrypt it with his private decryption key.

### C. The Assisted MITM Attack

As the discussions in the first two Sections demonstrate, the schemes used in many messaging services that offer end-to-end encryption actually require a great deal of trust in the service provider. First, the service provider must keep its word that the private keys are only stored on its users’ devices and are not secretly copied to the service provider’s servers. Apple’s iMessage is proprietary software, so there is no practical way to independently verify compliance with that requirement. There are, however, some open source end-to-end messaging programs, such as Signal, where ex-

perts are able to verify that the private keys remain secure.<sup>32</sup> But second, and more fundamentally, encryption and authentication schemes like the ones described here require that the public keys, which the service provider sends to a user, in fact match the identity of the counterparty that they supposedly belong to. That fundamental issue of matching keys to users is broadly called “*key distribution*.” An inherent weakness in any end-to-end encryption scheme that relies on a centralized directory of users for key distribution is that the service that operates the centralized directory must be trusted to always correctly identify which key belongs to which user.

Thus, because the correspondence between users and their public keys is a core requirement for all end-to-end encrypted systems, it is a source of potential vulnerabilities. If a third party were able to replace a user’s public key with its own, it would undermine the system’s security. Specifically, if it replaced Bob’s encryption key, it would be able to decrypt messages sent to him; if it replaced Alice’s verification key, it would be able to forge messages to seem as if they were written by her. The general form of this technique—intercepting communications and tampering with them before sending them on—is called a “man-in-the-middle (MITM) attack.” Because this Note focuses on the government acting as the third-party interceptor, it assumes that this key replacement is accomplished through some form of legal coercion, such as a subpoena or court order, rather than by technical malfeasance. Nonetheless, to adopt the common parlance of security literature, I refer to this technique as an “attack,” meaning simply an action that is intended to undermine some security guarantee. I refer to the government’s coercing the service provider to assist the government as an “assisted MITM attack” and to the legal process used in the coercion (be it subpoena, court order, or otherwise) as an “assisted MITM order.”

In the case of iMessage, the assisted MITM attack would work as follows. The government would generate its own encryption key pair and then send Apple an order that compels it to transmit to a sender the government’s public encryption key, instead of (or along with) one belonging to the intended recipient. For example, if Bob was the target and the government wanted to read all messages sent to him, it would compel Apple to replace Bob’s encryption key with the government’s encryption key. Ordinarily when Alice sends Bob a message, her device would normally en-

---

32. See *Signal*, SIGNAL, <https://signal.org/>; *Signal*, GITHUB, <https://github.com/signalapp> [<https://perma.cc/AKD5-7E3U>] (containing all source code for Signal software).

crypt it with Bob’s encryption key. However, because the government compelled Apple to send the *government’s* encryption key and claim it belonged to Bob, the government would be able to intercept and decrypt the message with its corresponding decryption key. To complete the attack, the government could generate a separate authentication key pair and compel Apple to claim that the government’s verification key was Alice’s. The government could then resend the message to Bob by signing with its own signing key. When Bob received the message, his device would ask Apple for Alice’s verification key, and Apple would be compelled to give the key that was supplied by the government. Thus, it would appear as if the message came from Alice. Neither Alice nor Bob would be able to detect what happened. Figure 5, below, illustrates the complete attack.

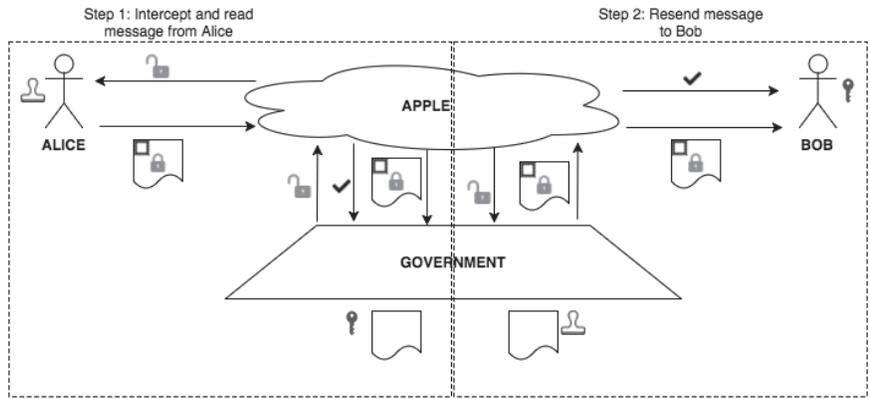


Figure 5: Illustrating the completed assisted MITM attack. The dashed rectangles represent the two steps of the process. In the first step, the government stands in the shoes of Bob. Alice receives an encryption key that she believes belongs to Bob but in actuality belongs to the government. In the second step, the government stands in the shoes of Alice. Bob receives a verification key that he believes belongs to Alice but in actuality belongs to the government. By conscripting Apple to forward the keys and represent that they belong to each party, the government is able to decrypt messages and forward them along to the other party without being detected.

As mentioned earlier, iMessage is designed so that each user can have multiple public keys (for both encryption and authentication) associated with her account: each key corresponds to a separate Apple device (on which is stored a corresponding private key). Thus, Apple has already built up the infrastructure for an even simpler version of this attack. The government’s key pair would just be treated as a new “device” associated with each user account, and its

public key would appear as one of the “pubkeys” in Figure 3, above. When Alice sent a message to Bob, the government would silently get a copy. To be fair, when a new device is associated with an account, the user typically receives an alert. Thus, if the government simply added a new key to Bob’s account, he may receive a warning.<sup>33</sup> Apple has never commented on whether it could disable that warning without alerting the user.

#### D. *A Brief Digression on San Bernardino*

Although not directly related to the topic of this Note, the extensive coverage of the 2015 San Bernardino terrorist attacks and of the subsequent controversy between Apple and the FBI make it worth taking a brief detour to explain how the issues in that case relate to the ones addressed here. Despite the presence of some of the same actors (Apple and the government) and the same themes (encryption on Apple devices, law enforcement needs, and the First Amendment), the issues are in fact very different.

The facts, briefly, are as follows. After the attack,<sup>34</sup> the FBI recovered a locked Apple iPhone 5C owned by one of the perpetrators. The iPhone contained encrypted content, but before the FBI could access that encrypted content, it had to unlock the iPhone by entering the suspect’s passcode. The iPhone was configured to automatically delete all the contents stored on the device after 10 incorrect unlock attempts (this is an optional feature present on newer iPhone devices).<sup>35</sup> The FBI requested Apple’s assistance in either disabling the feature or recovering the unencrypted content. Apple declined to help, and the FBI filed for and obtained a court order, issued under the All Writs Act,<sup>36</sup> compelling Apple’s assistance. Apple appealed the order, but the litigation became moot after an unidentified entity helped the FBI recover the contents.<sup>37</sup>

---

33. APPLE, *supra* note 23, at 60.

34. See generally Michael S. Schmidt & Richard Pérez-Peña, *F.B.I. Treating San Bernardino Attack as Terrorism Case*, N.Y. TIMES (Dec. 4, 2015), <https://www.nytimes.com/2015/12/05/us/tashfeen-malik-islamic-state.html>.

35. *Apple-FBI Battle Over San Bernardino Terror Attack Investigation: All The Details*, L.A. TIMES (Feb. 19, 2016, 6:08 PM), <http://www.latimes.com/business/technology/la-fi-tn-apple-fbi-20160219-htmstory.html> [<https://perma.cc/C7C9-HP5Z>].

36. See *infra* Part II.A for a detailed discussion of the All Writs Act.

37. Jack Date et al., *Justice Department Withdraws Request in Apple iPhone Encryption Case After FBI Accesses San Bernardino Shooter’s Phone*, ABC NEWS (Mar. 28, 2016, 10:05 PM), <http://abcnews.go.com/Technology/justice-department-withdraws-request-apple-iphone-encryption-case/story?id=37986428> [<https://perma.cc/2F54-G3NN>].

The validity of the AWA order has never been ultimately determined.<sup>38</sup>

The primary distinction between the order received by Apple in the aftermath of San Bernardino and an assisted MITM order as discussed in this Note lies in the differences between the role of encryption—and, consequently, the necessary technical assistance—in the two cases. Apple explained that just as with its iMessage service, it does not have the technical capability to directly decrypt the contents of a locked iPhone because it does not have a copy of the key used to encrypt it. It could, however, send an update to the operating system that would disable the security features, allowing the FBI to simply try all possible combinations without the device automatically erasing all of its contents after 10 incorrect attempts. One of the reasons the FBI needed Apple’s assistance is that the device hardware separately required all code updates to be digitally signed (in the same sense of signing as described in Part I.B) with the verification key hardcoded into the phone itself. Only Apple has the signing key necessary to sign updates; this is a security feature that prevents malicious attackers from writing false operating system updates to take over the device.<sup>39</sup>

Thus, the FBI’s request for assistance in the aftermath of San Bernardino is properly characterized as not only asking Apple to write code for it, but also for signing that code as, in some sense, “authentic.” Part of Apple’s argument against the validity of the San Bernardino order rested on a First Amendment theory that being forced to write computer code violates the First Amendment. That argument should not be confused with the discussion in Part III, which does not have to do with writing code, but rather with transmitting a particular message. The analogy to my argument would be a complaint that the order required Apple to designate the code as “authentic.” Apple never made that argument, and for good reason: by signing the code it was asked to write, it would simply be attesting that the code was “genuine Apple code,” which would be true. The signature would only mean that the code was written by Apple, not that Apple had a subjective desire to write it.

---

38. In fact, in an unrelated case, Magistrate Judge James Orenstein declined to issue a substantially identical order, finding that it was *not* appropriate under the All Writs Act. *In re Apple, Inc.*, 149 F. Supp. 3d. 341, 353–75 (E.D.N.Y. 2016). Thus, the issue is in the nascent stages of a potential circuit split.

39. L.A. TIMES, *supra* note 35.

*E. In Summary*

Needless to say, cryptography is a complex subject. For the purposes of this Note, however, all that is necessary to understand is the meaning of an assisted MITM order, which is simply a court order compelling the exchange of keys as illustrated in Figure 5, above. With that, this Note proceeds to a purely legal discussion.

## II.

## AUTHORITY FOR ISSUING SUCH AN ORDER

To my knowledge, there is no known incident in which the government actually issued anything like an assisted MITM order.<sup>40</sup> However, the possibility has long been discussed in technical and public literature<sup>41</sup> and, as explained in the previous Part, it would not be technically difficult to implement. This Part argues that there are at least two sources of legal authority that the government could rely on when issuing an assisted MITM order.

---

40. A recent report from Reuters describes an attempt by the Department of Justice to compel Facebook to “break the encryption in its popular Messenger app,” which allows users to encrypt some conversations end-to-end, including when making one-on-one calls through the program. Dan Levine & Joseph Menn, *Exclusive: U.S. Government Seeks Facebook Help to Wiretap Messenger—Sources*, REUTERS (Aug. 17, 2018 4:34 PM), <https://www.reuters.com/article/us-facebook-encryption-exclusive/u-s-government-seeks-facebook-help-to-wiretap-messenger-sources-idUSKBN1L226D> [<https://perma.cc/RM6Z-4J9L>]. The case remains under seal, so it is unknown whether the government is requesting an assisted MITM order or trying some other technique. However, Facebook is apparently arguing that “it can only comply with the government’s request if it rewrites the code relied upon by all its users to remove encryption or else hacks the government’s current target.” *Id.* An assisted MITM attack should not require alterations to any code, although it is hard to infer too much from unattributed sources in a non-technical news article.

An arguably related fact pattern comes from two cases in Washington, apparently stemming from the same arrest, in which a police officer used a seized iPhone to impersonate the arrestee in order to arrange some drug deals in a sting operation. *State v. Hinton*, 319 P.3d 9 (Wash. 2014); *State v. Roden*, 321 P.3d 1183 (Wash. 2014). The police officer did not seek court authorization prior to sending the text messages. *Hinton*, 319 P.3d at 11; *Roden*, 321 P.3d at 1185. The Supreme Court of Washington vacated both convictions on state constitutional and statutory grounds. In both cases, the court focused on the privacy expectations of the individuals who were tricked into communicating with the police officer. *Hinton*, 319 P.3d at 17 (basing ruling off of the state constitution); *Roden*, 321 P.3d at 1189–90 (state statute). Neither of the cases addressed either Apple’s interests or the interests of the owner of the iPhone who was being impersonated by the police officer. See also Sacharoff, *infra* note 126 (discussing listeners’ interests in the context of the First Amendment).

41. See *supra* note 19.

A. *The All Writs Act*

The All Writs Act (AWA) grants federal courts authority to “issue all writs necessary or appropriate in aid of their respective jurisdictions and agreeable to the usages and principles of law.”<sup>42</sup> The AWA acts as a “gap-filler” that gives courts the authority to issue orders that are not expressly provided for by other laws but which are necessary to effectuate some existing authority.<sup>43</sup> The preexisting authority may be some inherent authority vested in the courts or derived from some statutory or constitutional authority.<sup>44</sup>

A watershed decision connecting the AWA to advanced electronic investigative techniques is *United States v. New York Telephone Company*.<sup>45</sup> There, a court in the Southern District of New York found probable cause to issue an order that authorized FBI agents to install a “pen register”<sup>46</sup> to monitor several phone lines operated by the New York Telephone Company. The order, issued under the AWA and Rule 41 of the Federal Rules of Criminal Procedure, directed the company to assist the agents in installing the pen registers and the FBI to compensate the company for its efforts.<sup>47</sup> The company declined to comply with the order, arguing that such an order could only be issued through the special procedures established by the Wiretap Act.<sup>48</sup> The Second Circuit determined that pen registers do not fall under the scope of the Wiretap Act and that the district court had the authority to issue such an order under Rule 41 and the AWA.<sup>49</sup> The Second Circuit concluded, however, that the district court abused its discretion in ordering the company to offer technical assistance.

---

42. 28 U.S.C. § 1651(a) (2012).

43. See Brian M. Hoffstadt, *Common-Law Writs and Federal Common Lawmaking on Collateral Review*, 96 Nw. U. L. REV. 1413, 1460–61 (2002).

44. See, e.g., *Harris v. Nelson*, 394 U.S. 286, 299 (1969) (holding that the AWA confers a power to hold evidentiary hearings in connection with habeas corpus proceedings based on “[the presence] of habeas corpus jurisdiction and the duty to exercise it”).

45. 434 U.S. 159 (1977).

46. *Id.* at 161. “A pen register is a mechanical device that records the numbers dialed on a telephone . . . [without] . . . overhear[ing] oral communications. [It] does not indicate whether calls are actually completed.” *Id.* at 161 n.1.

47. *Id.* at 161.

48. *Id.* at 163. The Wiretap Act refers to Title III of the Omnibus Crime Control and Safe Streets Act of 1968. See *In re Google Inc. Cookie Placement Consumer Privacy Litig.*, 806 F.3d 125, 135 (3d Cir. 2015) (“The federal Wiretap Act is codified at 18 U.S.C. § 2510 *et seq.*”); Pub. L. 90-351, 82 Stat. 212.

49. 434 U.S. at 164.

The Supreme Court upheld the order in full. The Court agreed that the Wiretap Act does not cover pen registers<sup>50</sup> and found that to the extent that the Federal Rules of Criminal Procedure do not explicitly confer the ability, the AWA fills the gap.<sup>51</sup> Furthermore, the Court found that the district court did not abuse its discretion because “the [c]ompany was a third party [not] so far removed from the underlying controversy that its assistance could not be permissibly compelled.”<sup>52</sup> The Court found it significant that a district court found probable cause of a “criminal enterprise” which was potentially using “the [c]ompany’s facilities . . . on a continuing basis.”<sup>53</sup> It also emphasized that the company would be compensated for its assistance and that “compliance with [the order] required minimal effort on the part of the [c]ompany and no disruption to its operations.”<sup>54</sup>

Although the analog telephonic technology at issue in *New York Telephone* is very different from digitally-encrypted messages sent over the Internet, the legal analysis of the AWA remains the same. Apple is in an analogous position to the New York Telephone Company and the effort needed to effectuate an assisted MITM order would be minimal. The Supreme Court in *New York Telephone* pointed out that the company “regularly employs such devices without court order”<sup>55</sup> for various reasons. Apple similarly already has infrastructure for associating new keys with user accounts (such as when users purchase new devices), and there is no technical reason why Apple could not simply include a different key proposed by the government.<sup>56</sup> In other words, although the technical issues are different, the question of the powers of a court as conferred by the AWA are essentially identical.<sup>57</sup> The AWA would confer the necessary authority for a court to issue an assisted MITM order.

---

50. *Id.* at 165–68.

51. *Id.* at 168–70.

52. *Id.* at 174.

53. *Id.*

54. *Id.* at 175.

55. 434 U.S. at 174.

56. *See supra* Part I.C.

57. One distinction between an assisted MITM attack and the pen register in *New York Telephone* is that, depending on how the MITM attack was carried out, ongoing assistance from Apple may be required. The Ninth Circuit held that this distinction is not sufficient to defeat the authority conferred by the AWA. *See In re Application of the United States for an Order Authorizing an In-Progress Trace of Wire Commc’ns over Tel. Facilities*, 616 F.2d 1122, 1130 (9th Cir. 1980).

B. *The Stored Communications Act*

The Stored Communications Act (SCA)<sup>58</sup> addresses requirements for third-party service providers who store electronic data that belongs to their subscribers.<sup>59</sup> It proscribes voluntary disclosure of user data under some circumstances<sup>60</sup> and provides various levels of protection for compelled disclosure, depending on the service and the type of data requested.<sup>61</sup> Furthermore, the SCA allows a court to issue a “gag order” that prevents the service provider from notifying its subscriber that the provider was issued, or responded to, a data request.<sup>62</sup>

The SCA makes a distinction between “electronic communication services” and “remote computing services.” The former refers to “any service which provides to users thereof the ability to send or receive wire or electronic communications,”<sup>63</sup> whereas the latter refers to “the provision to the public of computer storage or processing services by means of an electronic communications system.”<sup>64</sup> An “electronic communications system,” in turn, is defined as “any wire, radio, electromagnetic, photooptical or photoelectronic facilities for the transmission of wire or electronic communications, and any computer facilities or related electronic equipment for the electronic storage of such communications.”<sup>65</sup> The term “electronic storage” is specifically defined to include “any temporary, intermediate storage of a wire or electronic communication incidental to the electronic transmission thereof.”<sup>66</sup> And finally, the term “electronic communication” is defined to include “any transfer of signs, signals, writing, images, sounds, data, or intelligence of any nature transmitted in whole or in part by a wire, radio, electromagnetic, photoelectronic or photooptical system,” but not analog audio

---

58. 18 U.S.C. §§ 2701–2712 (2012).

59. The SCA is actually Title II of the Electronic Communications Privacy Act, which in turn was an amendment to the Wiretap Act (Title III of the Omnibus Crime Control and Safe Streets Act of 1968). *Electronic Communications Privacy Act of 1986 (ECPA)*, 18 U.S.C. § 2510–22, JUSTICE INFO. SHARING, <https://it.ojp.gov/PrivacyLiberty/authorities/statutes/1285> [<https://perma.cc/B9HB-GJYX>]. Recall that the New York Telephone Company argued that the order compelling its help could not be issued under the AWA because it was foreclosed by the Wiretap Act.

60. *See* § 2702(a).

61. *See* § 2703.

62. *See* § 2705(b).

63. § 2510(15).

64. § 2711(2).

65. § 2510(14).

66. § 2510(17)(A).

communications, communications from tone-only paging devices or mobile tracking devices, or electronic fund transfers.<sup>67</sup>

It is difficult to say with certainty how these 1980s-era concepts should apply to twenty-first century technologies. Most modern Internet-enabled communication services could probably be characterized as either electronic communications services or remote computing services, depending on which application is at issue.<sup>68</sup> Regardless, the requirements of section 2703 would govern compelled disclosure of user data. A message stored for 180 days or less—which would include messages sent via iMessage, since Apple only keeps encrypted messages for one month<sup>69</sup>—can only be disclosed “pursuant to a warrant issued using the procedures described in the Federal Rules of Criminal Procedure (or, in the case of a State court, issued using State warrant procedures) by a court of competent jurisdiction.”<sup>70</sup> Messages stored longer than that can also be requested by court order or subpoena.<sup>71</sup>

Thus, to the extent the SCA modifies the landscape since *New York Telephone*, it provides the necessary substrate to which the AWA can be applied. An order issued under section 2703 for “the contents of a wire or electronic communication, that is in electronic storage in an electronic communications system for one hundred and eighty days or less” clearly describes a message in transit, given that storage incidental to a transmission is included in the definition of “electronic storage.”<sup>72</sup> Additionally, section 2706, which allows for reimbursement for services complying with governmental orders, specifically includes reimbursement for “such costs as are reasonably necessary and which have been directly incurred in searching for, assembling, reproducing, or otherwise providing such information.”<sup>73</sup> The broad language comfortably includes decryption, since decryption would be a required step in “assembling, reproducing, or otherwise providing” the messages. The SCA therefore most likely also provides courts with the authority necessary to issue an assisted MITM order, at least in the context of decryption.

---

67. § 2510(12).

68. *See, e.g.*, *Crispin v. Christian Audigier, Inc.*, 717 F. Supp. 2d 965, 980–91 (C.D. Cal. 2010) (concluding that social networking sites could be treated as either electronic communication services or remote computing services, depending on how they were being used).

69. APPLE, *supra* note 23, at 62.

70. § 2703(a).

71. § 2703(b).

72. *See supra* note 66 and accompanying text.

73. § 2706(a).

### C. *The Burr-Feinstein Legislation*

In the wake of the 2015 San Bernardino attacks and the resulting controversy between FBI and Apple,<sup>74</sup> Senators Richard Burr and Dianne Feinstein introduced legislation that would have implicitly granted authority for an assisted MITM order (as well as the order sought by the FBI to compel Apple to assist the FBI in accessing the suspect's locked iPhone following the San Bernardino Attacks).<sup>75</sup> The bill, whose short title was "Compliance with Court Orders Act of 2016,"<sup>76</sup> would have required that "a covered entity that receives a court order from a government for information or data shall provide such technical assistance as is necessary to obtain such information or data in an intelligible format or to achieve the purpose of the court order."<sup>77</sup> The term "covered entity" was defined to mean "a device manufacturer, a software manufacturer, an electronic communication service, a remote computing service, a provider of wire or electronic communication service, a provider of a remote computing service, or any person who provides a product or method to facilitate a communication or the processing or storage of data."<sup>78</sup> That expansive definition would easily include Apple as well as any other provider of messaging software.

Although the Burr-Feinstein legislation was specifically written to address Apple's non-compliance with the FBI's request in the wake of the San Bernardino attacks,<sup>79</sup> which is different from the vulnerability discussed here, the legislation was not limited to that context. First, the bill repeatedly refers to "communications services" as one of its primary targets.<sup>80</sup> Even more telling, the bill defines "technical assistance"—which the covered entities, including providers of communication services, must render—to include *either* delivering data in an intelligible format "concurrently with its transmission" *or* "expeditiously, if stored by a covered entity or on a

---

74. *See supra* Part I.D.

75. S. \_\_\_, 114th Cong. (2016) (discussion draft) [hereinafter Burr-Feinstein], <https://www.burr.senate.gov/imo/media/doc/BAG16460.pdf> [<https://perma.cc/AQR8-P3BK>].

76. *Id.* § 1.

77. *Id.* § 3(a)(1).

78. *Id.* § 4(4).

79. *See* Dustin Volz & Mark Hosenball, *Senate Proposal on Encryption Gives Judges Broad Powers*, REUTERS (Mar. 21, 2016, 6:22 PM), <http://www.reuters.com/article/us-apple-encryption-legislation/senate-proposal-on-encryption-gives-judges-broad-powers-idUSKCN0WN2B1> [<https://perma.cc/4YQ9-LKN5>] (discussing the legislation in the context of the San Bernardino attacks).

80. *See* Burr-Feinstein §§ 2(4), 3(e), 4(4), 4(6).

device.”<sup>81</sup> In other words, the legislation foresaw *both* a case analogous to San Bernardino—where the information was already stored on the device—and a situation where the information is being transmitted in real time. Although the bill does not explicitly authorize a prophylactic approach like an assisted MITM order, any “assistance” to obtain decrypted information of data “concurrently with its transmission” would presumably require some work on the part of the communications service provider prior to the transmission of the data. Thus, the legislation would have almost certainly authorized an assisted MITM order.

One could argue that the fact that the Burr-Feinstein legislation was proposed and ultimately rejected<sup>82</sup> suggests that Congress has not authorized assisted MITM orders under existing statutes, including the AWA or SCA. For the reasons already described in the preceding sections of this Part, I reject that view. Additionally, Burr-Feinstein was proposed soon after Magistrate Judge James Orenstein concluded, in a separate but factually indistinguishable case, that the AWA does not supply courts with the authority to grant the order the FBI was requesting after San Bernardino.<sup>83</sup> The case was widely publicized due to its conflict with the order authorized by the federal court in California.<sup>84</sup> Thus, it is much more likely that the Burr-Feinstein bill was introduced simply to foreclose any doubt about the AWA in that very novel context. The assisted MITM order setting, although incidentally affected, is in some sense much more similar to other existing wiretap techniques.

### III. THE RIGHTS OF THE SERVICE PROVIDERS

An assisted MITM order conscripts service providers (who are essentially third parties to government surveillance) and, as I elaborate below, forces them to lie. This Part argues that this aspect of the technique is constitutionally fatal because it violates the service providers’ First Amendment rights.

---

81. *Id.* § 4(12).

82. See Dustin Volz et al., *Push for Encryption Law Falters Despite Apple Case Spotlight*, REUTERS (May 27, 2016, 7:46 AM), <http://www.reuters.com/article/usa-encryption-legislation/push-for-encryption-law-falters-despite-apple-case-spotlight-id/USL2N18O0BM> [<https://perma.cc/BF7K-ETBF>].

83. See *supra* note 38.

84. See, e.g., Katie Benner & Joseph Goldstein, *Apple Wins Ruling in New York iPhone Hacking Order*, N.Y. TIMES (Feb. 29, 2016), <https://www.nytimes.com/2016/03/01/technology/apple-wins-ruling-in-new-york-iphone-hacking-order.html> [<https://perma.cc/QP6R-U37E>].

A. *The Exchange of Cryptographic Keys is Speech Under the First Amendment*

A threshold question is whether the mandated exchange of cryptographic keys counts as “speech” for First Amendment purposes. As explained in Part I, cryptographic keys are simply very large numbers, inherently meaningless and in some interfaces, including iMessage, not even displayed to the user. One could argue, therefore, that cryptographic keys are not speech because they are not expressions that are readily comprehensible to the communicating users. This argument is unavailing for several reasons.

First, the communication is not simply that of a cryptographic key, but rather a statement *about* the key—namely, that it belongs to a particular user. When Alice sends a message to Bob, Apple supplies her with a public encryption key that it promises (for example, through its marketing materials<sup>85</sup>) will encrypt her message in such a way that only Bob can read it. As explained in Part I, the technology that enables this requires an associated private decryption key that only Bob has access to, and as illustrated in Figure 3 of Part I.B, the communication explicitly ties the key to the user. Thus, because Bob has exclusive control of the private key, the public key associated with it is necessarily tied to Bob as well. Although the public key is not particularly meaningful to Alice as a number, the attestation about its association with Bob is.

Second, the Supreme Court has adopted a very expansive view of what counts as speech. For example, in *Expressions Hair Design v. Schneiderman*, the Court unanimously ruled that a New York provision that prohibited merchants from imposing a surcharge on customers paying with credit cards—but allowed the economically equivalent practice of describing the price differential as a discount to customers paying by cash—was speech for the purposes of the First Amendment.<sup>86</sup> The communication in *Expressions Hair Design* were simply price stickers—numbers solely communicating a price. Similar to price stickers, which are used to communicate a price, cryptographic keys are used to communicate the identity of parties exchanging messages. Thus, cryptographic keys fit within this expansive interpretation of the right to freedom of speech because the exchanging of keys communicates information regarding party identity.

---

85. As an example, consider Apple’s publicly available document that describes the security features of its iOS devices. See APPLE, *supra* note 23.

86. 137 S. Ct. 1144, 1147 (2017).

It is hard to distinguish numbers that communicate a price with numbers that communicate a cryptographic key, but the biggest differentiating factor, and the one alluded to earlier, is that users do not “understand” the cryptographic key, while they do “understand” a price. This line of reasoning—that whether a communication should count as “speech” turns on its understandability to humans—has been examined in a different line of cases concerning computer code. Although computer code is not the same as cryptographic keys—and I do not here argue that writing code always constitutes speech—cryptographic keys, like some computer code, are speech because they communicate an idea. For example, in *Universal City Studios, Inc. v. Corley*, the Second Circuit reviewed an order, issued under the Digital Millennium Copyright Act, enjoining web site owners from posting computer software that allowed users to play movies on DVDs on devices that were not specially licensed by manufacturers.<sup>87</sup> In comparing computer programs to musical notes, the court held that computer code is speech covered by the First Amendment because it can be used as a means to communicate information between specialists.

Similarly, in *Junger v. Daley*, the Sixth Circuit determined that the fact that computer code can communicate information entitles it to First Amendment protection.<sup>88</sup> There, a law school professor appealed a determination that certain software he wanted to publish was subject to restrictions under the Export Administration Regulations.<sup>89</sup> In reversing the district court’s grant of summary judgment against the professor, the court held that software is protected by the First Amendment because it can be used to communicate ideas.<sup>90</sup> “[A]ll ideas having even the slightest redeeming social importance,’ including those concerning ‘the advancement of truth, science, morality, and arts’ have the full protection of the First Amendment,” the court pronounced.<sup>91</sup>

Both the *Corley* and *Junger* courts noted the fact that computer code comes in two forms: “object code” and “source code.” Object code is code that a physical computer can directly process. A computer ultimately processes electrical signals, and object code is just a representation of those signals. Source code, meanwhile, is a set of instructions written in higher-level concepts intended to be under-

---

87. 273 F.3d 429 (2d Cir. 2001).

88. 209 F.3d 481 (6th Cir. 2000).

89. *Id.* at 483–84. Coincidentally, the software at issue happened to be encryption software. *Id.* at 483.

90. *Id.* at 485.

91. *Id.* at 484 (quoting *Roth v. United States*, 354 U.S. 476, 484 (1957)).

stood by human programmers. Source code is not directly interpretable by a computer and must instead be translated into object code by specialized software.<sup>92</sup> The exact relationship between source code and the resulting object code, as well as the translation process between the two, varies across programming languages and operating systems. The reason that the distinction matters, however, is that object code is usually generated by, and processed by, computers and not humans. Similarly, cryptographic keys are usually processed by software and not directly used by humans. Nonetheless, although both *Corley* and *Junger* based their First Amendment holdings on the fact that computer code can be used to communicate ideas between programmers, neither confined its holdings to source code. *Corley*, for example, explicitly refused to distinguish code based on whether or not it is executable (i.e., object code). “[T]he fact that a program has the capacity to direct the functioning of a computer does not mean that it lacks the additional capacity to convey information, and it is the conveying of information that renders instructions ‘speech’ for purposes of the First Amendment.”<sup>93</sup> While acknowledging that “it would be inconvenient, inefficient and, for most people, probably impossible to” program in object code,<sup>94</sup> courts have nonetheless accepted that the code still “communicates” those instructions and thus should not lose protection, just as a poem translated into Sanskrit should not lose protection just because few understand the language.<sup>95</sup> Similarly, in a copyright case, the Ninth Circuit has held that object code can “be copyrighted as expression” and that object code “also contains ideas.”<sup>96</sup> Likewise, cryptographic keys, which, like object code, are usually meaningless to humans, are nevertheless “expression[s]” that “contain ideas.”

The fundamental problem of drawing a constitutional line between source code and object code is that the latter “is merely one additional translation of speech into a new, and different, language.”<sup>97</sup> This encapsulates the reason that the exchange of crypto-

---

92. *See id.* at 483. *See generally* Daniel Lin et al., *Source Code Versus Object Code: Patent Implications for the Open Source Community*, 18 SANTA CLARA COMPUTER & HIGH TECH. L.J. 235, 238–39 (2002).

93. 273 F.3d at 447.

94. *Id.* at 439.

95. *Id.* at 446.

96. *Sony Comput. Entm’t, Inc. v. Connectix Corp.*, 203 F.3d 596, 602 (9th Cir. 2000).

97. *United States v. Elcom Ltd.*, 203 F. Supp. 2d 1111, 1126 (N.D. Cal. 2002). Elsewhere, I have argued a similar point, suggesting that the line between the two is elusive and is based less on whether the code is “executable” and more on how

graphic keys must be considered “speech” as well. Consider the simple shift-cipher discussed in Part I.A. If a recipient sees the message “MJQQT,” the information that the letters were shifted up by 5 letters is incredibly meaningful. Without a communication to inform the recipient of the shift-cipher, the recipient has no way of comprehending the ideas expressed in the message—i.e., that the message is intended to read “HELLO.” It is, to borrow the analogy, akin to teaching someone enough Sanskrit to understand a previously incomprehensible poem. “The letters were shifted up by 5 letters”—or, equivalently, “the key is 5”—is clearly speech for First Amendment purposes because it communicates information. To be sure, cryptographic keys are generally far larger, are used as inputs into far more complex algorithms, and are not intended to be referenced by humans. But neither is object code. The point is that both do carry a particular meaning. For cryptographic keys, that meaning is quite literally the information needed to understand a message already sent.

In a footnote in *Corley*, the Second Circuit cautioned that “in the rare case where a human’s mental faculties do not intercede in executing the instructions, we have withheld protection.”<sup>98</sup> Courts that have followed this doctrine have relied on this dichotomy that is subtly different from the object-vs-source code dichotomy. If the *recipient* of the code is only the computer—that is, if the communicators are computers—no First Amendment protection should be extended. This dichotomy covers cryptographic keys more cleanly. Because the users are not personally plugging the keys into the cryptographic algorithm, the keys should not count as communications covered by the First Amendment. This would presumably also distinguish the shift-cipher scenario.

The problem with this reasoning is that cryptographic keys *are* used by humans, or at least could be in some common scenarios. For example, although Apple does not show its users the public keys associated with the people they are talking to, many competing products, including Facebook’s WhatsApp, do.<sup>99</sup> The purpose of

---

expressive the language is. Leonid Grinberg, *The Generativity of Programming Languages: Why “Open Source” Is About Expressive Power*, THE FUTURE OF THE INTERNET BLOG (Aug. 12, 2013), <http://blogs.harvard.edu/futureoftheinternet/2013/08/12/the-generativity-of-programming-languages-why-open-source-about-expressive-power/>.

98. *Corley*, 273 F.3d at 448 n.20.

99. See *End-To-End Encryption*, WHATSAPP, <https://faq.whatsapp.com/en/android/28030015/> [<https://perma.cc/FB2F-3XBE>] (describing the “Verify Security Code” feature, which shows a number derived from the security keys). Similarly, many modern browsers display a graphic, often a green lock, to indicate that the

displaying the keys to users is to allow them to visually compare the keys across time and ensure they do not change. For example, if Alice and Bob want to use WhatsApp to secretly communicate, they could meet up in person and visually compare the keys displayed by WhatsApp. Afterwards, when they are no longer next to each other, either one can look at the displayed key and ensure that they are still speaking with the same person. Thus, although neither Alice nor Bob is manually plugging the other's key into the encryption/decryption algorithm, each is still very much relying on the key as a communication about the person she or he is speaking with.

Furthermore, whether keys are being manually plugged into an algorithm is a very brittle, and ultimately unworkable, line on which to base a constitutional judgment. For one thing, what is mathematically difficult for one person is easy for another. It would be time-consuming to decrypt a message by hand when it has been encrypted with an industry-strength encryption scheme, but it would certainly be possible. Only one decryption operation needs to be performed for a given message, so while it would take a fair amount of time, it would be significantly more practical to decrypt a message by hand—even one encrypted with a large key—than for a user to try to manually read or write object code, which might involve hundreds of millions, if not billions, of operations. Moreover, if a user *herself* writes the software to perform the decryption—after all, the specifications are publicly available and relatively straightforward to implement, even for students<sup>100</sup>—would that destroy her constitutional credibility as a bona fide “participant” in the cryptographic exchange?

Moreover, as the *Corley* court noted, “even dry information, devoid of advocacy, political relevance, or artistic expression, has been accorded First Amendment protection.”<sup>101</sup> Several years later, in *Sorrell v. IMS Health Inc.*, the Supreme Court held that information about physicians' preferences in issuing prescription was speech covered by the First Amendment.<sup>102</sup> “Facts,” explained the Court,

---

connection to the website is secured. See, e.g., *How Do I Tell If My Connection to a Website Is Secure?*, MOZILLA, <https://support.mozilla.org/en-US/kb/how-do-i-tell-if-my-connection-is-secure> [<https://perma.cc/MVU9-2DX3>].

100. See, e.g., Scott Aaronson, *6.080/6.089 Lecture 17*, MIT OPENCOURSEWARE at 2–6 (Apr. 15, 2008), [https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-045j-automata-computability-and-complexity-spring-2011/lecture-notes/MIT6\\_045JS11\\_lec14.pdf](https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-045j-automata-computability-and-complexity-spring-2011/lecture-notes/MIT6_045JS11_lec14.pdf) [<https://perma.cc/NY65-KGDD>] (providing and explaining two well-known public-key encryption algorithms in a computer science undergraduate class).

101. 273 F.3d at 446.

102. 564 U.S. 552, 570 (2011).

“are the beginning point for much of the speech that is most essential to advance human knowledge and to conduct human affairs.”<sup>103</sup> Since *IMS Health*, the Court has continued embracing an aggressive view of First Amendment protections, including for purely factual information. For example, it recently invalidated a town’s “sign code”—which had differing provisions for signs directing the public to non-profit group meetings as opposed to signs covering other messages—under a strict scrutiny analysis.<sup>104</sup>

To reiterate, the communications containing cryptographic keys are not just numbers. They are statements of fact sent by the service provider indicating that a particular message was sent by a particular user. The key that Apple sends to the sender’s device—a key that it says belongs to a particular user, but, in reality, belongs to the government—is sent with the imprimatur of Apple, Inc. It is, in the parlance of security engineers but also quite literally, an *attestation* about the message and its security guarantees. This attestation is sent every single time Apple delivers a message from Alice to Bob. Each time Bob sees a little blue bubble<sup>105</sup> with Alice’s name at the top of the screen, Apple has communicated to Bob that Alice sent that message.

There is, in other words, a reputational element at stake. Apple’s communication to Bob—that the message Bob received is from Alice—acts like a signature or a seal. Any system of verification is only as good as the verifier’s word. And the evils of a compelled lie extend beyond even financial harms. If the government forced a prestigious university to print a fake diploma and transcript for an undercover agent, that may impose financial costs on the university, but it would have a more fundamentally damaging effect as well. The diploma from the university would simply *mean less*. The right not to be compelled to cheapen that meaning is protected by the First Amendment.

---

103. *Id.*

104. 135 S. Ct. 2218, 2227 (2015).

105. In the iMessage interface, messages displayed on a blue background indicate that they were sent over the iMessage service, complete with the security guarantees of the service. In contrast, messages displayed on a green background were sent on the “short message service” (SMS), which is the standard “texting” protocol that features no security guarantees. See generally *About iMessage and SMS/MMS*, APPLE, <https://support.apple.com/en-us/HT207006> [https://perma.cc/RX39-XQAU]; Adam Fendelman, *Explaining SMS Messaging and Its Limitations*, LIFEWIRE (May 21, 2018), <https://www.lifewire.com/definition-of-sms-text-messaging-578676> [https://perma.cc/7HG9-ZC2H].

B. *Compelled Speech Triggers Strict Scrutiny*

Having established that the exchange of cryptographic keys constitutes speech, I now argue that compelling a service provider to make false attestations about key ownership triggers strict scrutiny under the compelled speech doctrine.

Although many of the most grandiose utterances about the First Amendment’s guarantees securing freedom of speech concern specifically limitations on the government’s ability to *restrict* speech,<sup>106</sup> the First Amendment just as surely limits the government’s ability to *compel* it. This principle was emphasized in *West Virginia Board of Education v. Barnette*,<sup>107</sup> which, in overruling an opinion issued just three years prior, held that the First Amendment bars a public school from compelling students to salute the American flag.<sup>108</sup> The Court’s pronouncement that the First Amendment prevents the government from compelling students to salute the flag—and thus express a particular viewpoint,<sup>109</sup> even if it may be unpopular—remains one of the most enduring passages about the First Amendment: “If there is any fixed star in our constitutional constellation, it is that no official, high or petty, can prescribe what shall be orthodox in politics, nationalism, religion, or other matters of opinion *or force citizens to confess by word or act their faith therein.*”<sup>110</sup>

The compelled communication that would be at issue if the government issued an assisted MITM order is the service provider’s attestation regarding key ownership. Private keys are by definition private to a particular account, which is to say a particular *identity*; sending a public key and promising that it corresponds to a particular private key amounts to promising that the public key belongs to the identity of the private key holder. While these communications

---

106. *E.g.*, *Citizens United v. FEC*, 558 U.S. 310, 356 (2010) (“The First Amendment confirms the freedom to think for ourselves.”); *Texas v. Johnson*, 491 U.S. 397, 414 (1989) (“If there is a bedrock principle underlying the First Amendment, it is that the government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable.”); *Bose Corp. v. Consumers Union, Inc.*, 466 U.S. 485, 503–04 (1984) (“The First Amendment presupposes that the freedom to speak one’s mind is not only an aspect of individual liberty—and thus a good unto itself—but also is essential to the common quest for truth and the vitality of society as a whole.”).

107. 319 U.S. 624 (1943).

108. *Id.* at 642.

109. No one disputed that the flag salute was an expressive act that implicated the First Amendment. *Id.* at 632 (“There is no doubt that, in connection with the pledges, the flag salute is a form of utterance.”).

110. *Id.* at 642 (emphasis added).

may not be of the same character as a forced confession of loyalty to a nation or its flag, they are real, factual statements about the identity of the parties in communication. These communications are therefore analogous to the relatively “technical speech” that, although having no relation to politics, religion, or other personal viewpoints, is protected by the First Amendment. The compelled utterance of this drier, technical speech is as strictly scrutinized as political speech.<sup>111</sup>

In *Hurley v. Irish-American Gay, Lesbian and Bisexual Group*,<sup>112</sup> the Court explained that the right to be free from compelled speech “applies not only to expressions of value, opinion, or endorsement, but equally to statements of fact the speaker would rather avoid.”<sup>113</sup> By way of example, one of the cases that the Court cited was *Riley v. National Federation of the Blind of North Carolina, Inc.*,<sup>114</sup> which struck down a statute that required fundraisers to “disclose to potential donors, before an appeal for funds, the percentage of charitable contributions collected during the previous 12 months that were actually turned over to charity.”<sup>115</sup> The Court struck down the statute as an unconstitutional regulation of free speech, explaining that “[m]andating speech that a speaker would not otherwise make necessarily alters the content of the speech,”<sup>116</sup> and that although there is “some difference between compelled speech and compelled silence” that “difference is without constitutional significance, for the First Amendment guarantees ‘freedom of speech,’ a term necessarily comprising the decision of both what to say and what *not* to say.”<sup>117</sup> The Court has continued to embrace this jurisprudence: for example, in *National Institute of Family & Life Advocates v. Becerra*, the Court partially relied on *Riley* in applying strict scrutiny to examine a California law that imposed certain disclosure requirements on so-called “crisis pregnancy centers.”<sup>118</sup>

---

111. See, e.g., *Pac. Gas & Elec. Co. v. Pub. Utils. Comm’n*, 475 U.S. 1, 16 (1986) (finding that a public commission may not require an electric company to include newsletters from a consumer advocacy organization in envelopes sent to customers).

112. *Hurley v. Irish-American Gay, Lesbian and Bisexual Grp.*, 515 U.S. 557 (1995).

113. *Id.* at 573.

114. 487 U.S. 781 (1988).

115. *Id.* at 795.

116. *Id.*

117. *Id.* at 796–97.

118. See *Nat’l Inst. Family & Life Advocates v. Becerra*, 138 S. Ct. 2361, 2371 (2018).

These precedents make clear that the applicable level of scrutiny does not depend on whether speech is restricted or compelled, nor on how political or opinionated it is. Instead, it depends on the purpose of the regulation and the type of speech at issue (for example, commercial speech typically enjoys less scrutiny than non-commercial speech<sup>119</sup>). And although the speech in the assisted MITM attack may be dry and technical, First Amendment protections for purely factual speech remain quite high. Thus, if the exchange of cryptographic keys is speech for the purposes of the First Amendment—as I argue it is in the previous Section—then a compelled assertion like the one in an assisted MITM attack is unquestionably a content-based compulsion of speech. After all, it requires the provider to alter the substance of its communication to the recipient of a message by compelling the service provider to falsely attest that a particular user sent the message, when in fact the government sent the message. Content-based restrictions on speech almost always call for strict scrutiny.<sup>120</sup> Furthermore, none of the various exceptions that occasionally reduce judicial scrutiny of content-based restrictions or regulations apply.<sup>121</sup> A key exchange is not commercial speech.<sup>122</sup> Nor is an assisted MITM order primarily a non-speech-related order that happens to incidentally affect speech, since the actual substance of the order is to generate

---

119. See generally *Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n*, 447 U.S. 557, 562–66 (1980).

120. See, e.g., *id.* at 2226.

121. Of course, a comprehensive analysis of every possible exception to boilerplate strict scrutiny First Amendment doctrine would fill a treatise, but one helpful summary was recently published by the Congressional Research Service. KATHLEEN ANN RUANE, CONG. RESEARCH SERV., FREEDOM OF SPEECH AND PRESS: EXCEPTIONS TO THE FIRST AMENDMENT (2014), <https://fas.org/sgp/crs/misc/95-815.pdf> [<https://perma.cc/67Z9-CT35>].

122. See, e.g., *Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n*, 447 U.S. 557, 561 (1980) (defining “commercial speech” as “expression related solely to the economic interests of the speaker and its audience”); see also Victor Brudney, *The First Amendment and Commercial Speech*, 53 B.C. L. REV. 1153, 1154–61 (2012) (proposing various definitions of commercial speech). Professor Brudney separates commercial speech into a “narrow” category that merely describes a proposed transaction, and an “enriched” category that contains additional expression that would otherwise be “ordinary” speech (e.g., a description of how great some product or lifestyle is) but is attached to a proposed transaction. Apple’s whitepaper about its products’ security, see APPLE *supra* note 23, might fall into this latter category. However, the actual communication of the keys would not be commercial speech under either definition, since it is entirely divorced from any kind of transaction; indeed, the one relevant commercial transaction—the purchase of an Apple product—must have already occurred for the conversation to be taking place over iMessage at all.

speech (in the form of falsely attributing a key to a user). It is, quite simply, speech that is being compelled as part of a particular regulatory framework. Regardless of whether or not the government is allowed to compel such speech in a particular situation, compelling a service provider to make a particular attestation invariably triggers strict scrutiny.

*C. The Routine Use of an Order Compelling Attestations About Cryptographic Keys Cannot Survive Strict Scrutiny*

Of course, First Amendment protection does not render all restrictions on a communication impermissible. Under strict scrutiny analysis, restrictions on speech are allowed if they are imposed for compelling government purposes and are narrowly tailored towards fulfilling those purposes.<sup>123</sup> It would be impossible to say here that *no* application of an assisted MITM order in any circumstance would be impermissible. I argue merely that it is impermissible as a matter of routine police practice.<sup>124</sup>

There is no dispute that the government's interest in law enforcement is a compelling interest. The real question is whether this type of regulation is sufficiently narrowly tailored. Although it comes in a very different context, a rich body of case law involving compelled speech against the backdrop of compelling government interests appears in the context of mandatory disclosure requirements. The most important differentiating factor is that in the con-

---

123. *Reed*, 135 S. Ct. at 2231.

124. For brevity's sake, I focus here on criminal investigations rather than on national security missions. Of course, action based on national security are not immune to First Amendment restrictions, and courts have occasionally stricken down such actions on constitutional grounds. *See, e.g.*, *New York Times Co. v. United States*, 403 U.S. 713 (1971) (per curiam) (upholding newspapers' rights to publish the Pentagon Papers notwithstanding national security concerns). That said, I concede that the analysis in this Section would likely change in a national security context where the government interest is higher, and courts tend to accord the executive much greater deference. *See, e.g.*, *In re National Security Letter*, 863 F.3d 1110 (9th Cir. 2017) (finding that a non-disclosure requirement issued with a National Security Letter withstood strict scrutiny). That said, it is also worth noting that, as a practical matter, techniques used for national security are rarely used for ordinary criminal investigations. OFFICE OF THE INSPECTOR GENERAL, U.S. DEPARTMENT OF JUSTICE, A SPECIAL INQUIRY REGARDING THE ACCURACY OF FBI STATEMENTS CONCERNING ITS CAPABILITIES TO EXPLOIT AN IPHONE SEIZED DURING THE SAN BERNARDINO TERROR ATTACK INVESTIGATION 4 n.3 (March 2018), <https://oig.justice.gov/reports/2018/o1803.pdf> [<https://perma.cc/E7V7-2YH8>] (explaining that internal Department of Justice policies require stringent procedures for using national security techniques for criminal cases, and noting that a relatively high-level FBI official was only aware of two instances between 2002 and 2015 when the procedures had to be invoked).

text of an assisted MITM order, the compelled speech (that the key being sent belongs to the identified user) is unequivocally false. And although I am not aware of any case that has explicitly said that false compelled speech is more suspect than truthful compelled speech, that inference can be drawn from one of the most basic justifications for free speech: that free speech facilitates the finding of truth.<sup>125</sup> Alternatively, Professor Laurent Sacharoff has argued that the doctrine against compelled speech is best understood as protecting the *listener's* interests rather than the speaker's.<sup>126</sup> Considered through the lens of the listener, traditional justifications for free speech, including the search for truth and the promotion of the marketplace of ideas, become even clearer.<sup>127</sup> Compelling service providers to falsely attribute keys to its users directly conflicts with this justification for free speech because the "listener," the recipient of the message, is being tricked into believing that he is communicating with a particular person, when in fact he is communicating with the government.

Furthermore, cases involving mandatory disclosure requirements illustrate the importance of protecting the communication of truthful information. For example, in *Zauderer v. Office of the Disciplinary Counsel of the Supreme Court of Ohio*,<sup>128</sup> the Court partially upheld a disciplinary sanction imposed on an attorney for his advertisements. In particular, although the Court found that certain rules proscribing the use of illustrations in attorney advertise-

---

125. See generally Kent Greenawalt, *Free Speech Justifications*, 89 COLUM. L. REV. 119, 130–41 (1989). Given this justification, and the recognition that national security letter (NSL) non-disclosure requirements can withstand strict scrutiny, see *supra* note 124, an interesting question arises as to the legal efficacy of "warrant canaries." A "warrant canary" is a message periodically issued by a technology company that foresees receiving an NSL at some point in the future. The message says something to the effect of "this company has not yet received an NSL [or a similar warrant or court order]." The idea is that once a company receives an NSL, it will quietly remove the message, "killing the canary in the coal mine" and silently signaling that an NSL was received without *technically* violating the nondisclosure provision. A recent paper suggested that an injunction prohibiting the removal of a canary would be more suspect than the provision itself because it would be compelling false speech. See Naomi Gilens, Note, *The NSA Has Not Been Here: Warrant Canaries as Tools for Transparency in the Wake of the Snowden Disclosures*, 28 HARV. J. L. & TECH. 525, 540 (2015). For what it's worth, however, that article was written prior to the Ninth Circuit's decision finding that an NSL letter withstood strict scrutiny. See *supra* note 124.

126. Laurent Sacharoff, *Listener Interests in Compelled Speech Cases*, 44 CAL. W. L. REV. 329 (2008).

127. *Id.* at 374–77.

128. 471 U.S. 626 (1985).

ments were unconstitutional,<sup>129</sup> the Court explained that the government may require mandatory disclosures of some factual information. Emphasizing that the state's "prescription has taken the form of a requirement that appellant include in his advertising *purely factual and uncontroversial* information,"<sup>130</sup> the Court upheld the disclosure requirement, explaining that "because [such] requirements trench much more narrowly on an advertiser's interests than do flat prohibitions on speech, warnings or disclaimers might be appropriately required in order to dissipate the possibility of *consumer confusion or deception*."<sup>131</sup> The logic of this opinion would not hold if the information at issue was false.

Perhaps even more telling are cases concerning mandated abortion disclosures. Courts have been particularly deferential to legislatures in terms of what disclosures they may mandate in the abortion context, notwithstanding significant doubt and resistance by the medical community.<sup>132</sup> And yet, even highly deferential abortion-related cases make clear that demonstrable *false* disclosure cannot be compelled. For example, in *Planned Parenthood v. Rounds*, the Eighth Circuit, sitting *en banc*, reversed an earlier panel's decision striking down a disclosure requirement that it found factually unsubstantiated.<sup>133</sup> But in doing so, it made clear that Planned Parenthood could have succeeded if it established that the mandated disclosure was actually "*untruthful*."<sup>134</sup> The Eighth Circuit simply disagreed with the panel's conclusion that the disclosure was untruthful or misleading.

Here, the compelled communication of the cryptographic keys would be unquestionably untruthful because the keys would not belong to the purported user. That, of course, is the whole point of compelling the sending of those keys. As in *Planned Parenthood v. Rounds*, where the court implied that an untruthful mandatory disclosure would have resulted in Planned Parenthood prevailing

---

129. *Id.* at 649.

130. *Id.* at 651 (emphasis added).

131. *Id.* (emphasis added) (cleaned up).

132. Professor Rebecca Dresser has noted this phenomenon in cataloging some examples of mandated disclosure requirements that "conflict with accepted medical knowledge." Rebecca Dresser, *From Double Standard to Double Bind: Informed Choice in Abortion Law*, 76 GEO. WASH. L. REV. 1599, 1609 (2008). She explains that this is in conflict with traditional notions of informed consent, which generally do not require physicians to warn of health risks that are not recognized by the medical community as having a causal link to a proposed procedure or medicine. *Id.* at 1618–19.

133. 686 F.3d 889, 892 (8th Cir. 2012) (*en banc*).

134. *Id.* at 893 (emphasis added) (cleaned up).

under strict scrutiny, an order requiring service providers to make untruthful statements is unlikely to be the least restrictive, most narrowly-tailored approach to conducting law enforcement operations.

Finally, although users are the ones ultimately targeted, the service providers are the ones whose speech is being burdened and whose reputation is at stake. Since communication service providers succeed by achieving network effects, there are naturally relatively few of them, and fewer still that feature end-to-end encryption. Thus, it is a small group of repeat players that would be routinely and repeatedly subjected to compelled speech. To the extent that strict scrutiny analysis takes into account the burden imposed on the speaker,<sup>135</sup> this additionally weighs against the constitutionality of an assisted MITM order because the same companies would be routinely asked to make false attestations in myriads of cases.<sup>136</sup>

#### D. *Banning the Services Altogether*

The above analysis, while forceful, is narrow. I argue only that, in the context of *existing* systems, the government cannot undermine the key distribution system by compelling service providers to make false attestations to their users.

An entirely separate question is whether such services may be banned altogether. As discussed in Part II.C, the Burr-Feinstein legislation introduced in the wake of the 2015 terrorist attacks in San Bernardino would have required service providers to ensure that they are able to decrypt data when writing encryption software. The

---

135. See, e.g., *Riley v. National Federation of the Blind of North Carolina, Inc.*, 487 U.S. 781, 798 (1988) (considering whether a compelled disclosure is “unduly burdensome” as part of the narrow tailoring factor in a strict scrutiny analysis).

136. Of course, not all compelled assistance provisions are unconstitutional. The Store Communication Act contemplates compelled assistance via subpoena, for example, see *supra* Part II.B, and one must assume that companies have been compelled to write code in response to orders issued under such subpoenas. For example, the Structured Query Language (SQL) is a programming language used to interface with many database systems, the contents of which may be the target of a subpoena. See generally ALAN BEAULIEU, *LEARNING SQL* (2d ed. 2009). What distinguishes the assisted MITM order is that the compulsion is not to write code but to *communicate* a statement (key ownership). To take the example of Apple in the San Bernardino case, the compelled order was to enable the FBI to access the unencrypted contents of the device—a device that it already had in its possession. That assistance would have required Apple to write code—which it argued was speech—but it was not the code that the government was seeking. Analogously, a subpoena for archived tax records served on a company would presumably require communications among the employees to produce the responsive records, but those incidentally necessary communications are not the real subject of the subpoena. The tax records are.

Burr-Feinstein legislation would not implicate the same constitutional issues as above, since it would not result in the government forcing a service provider to make a particular communication—it would simply make services like iMessage, WhatsApp, and Signal illegal.

I take no position on whether a law entirely banning software that features end-to-end encryption would be constitutional. As noted in Section III.A, courts have routinely found computer code to be speech for the purposes of the First Amendment. Thus, software vendors might have a colorable argument that such a ban is a content-based prior restraint on speech. On the other hand, restrictions on software (including, specifically, encryption software) have historically been implemented in terms of export controls rather than restraints on *writing* the software,<sup>137</sup> and one could imagine carefully written legislation that would make it essentially impossible to sell software featuring end-to-end encryption without running afoul of export regulations, even though technically *producing* such software would be legal. And there have been other proposed ways that would inhibit end-to-end encryption, including key escrow.<sup>138</sup> The constitutionality of these proposals is outside the scope of this Note.

I argue only that there is a sustained, and perhaps increasing, call for the government to be able to access messages sent by users over encrypted communication systems. There is at least one way for the government to do so that would require minimal technical work. But in many cases, I believe that way would not be a constitutional one.

#### IV. THE RIGHTS OF THE USERS

The previous Part analyzed an assisted MITM order—a scenario in which the government compels a service provider to falsely

---

137. *Bernstein v. United States* was a prominent series of cases concerning export controls on encryption software in which Daniel Bernstein, a student and later professor studying cryptography, sought to publish certain papers and software code that had originally been subject to export controls. After several trips between the district court in the Northern District of California and the Ninth Circuit, the government ended up loosening the regulations and the case was dismissed as unripe. See, D. J. Bernstein, *Bernstein v. United States*, <https://cr.yo.to/export.html> [<https://perma.cc/K96T-NWJA>] (summarizing the controversy); 65 Fed. Reg. 2,492 (the relaxed rules). The rules currently in effect are quite broad and allow for current end-to-end encrypted products to be sold. See, e.g., 15 C.F.R. § 740.17 (2018).

138. See *supra* notes 11–14 and accompanying text.

attribute government keys to the provider’s users—and concluded that such an order would violate the First Amendment rights of the service provider. This Part examines the rights of the users.

Of course, while the rights of the users and service providers are separate and distinct, they are not completely isolated from one another. If a service provider is served with an assisted MITM order along with a gag order (such as under the Stored Communication Act), the service provider may be the only one in a position to assert its users’ rights. A threshold question, therefore, is whether the users have any privacy rights that the service provider could assert. This Part first reviews some recent case law to show that the third-party standing doctrine forecloses that strategy. It then argues that no separate First Amendment remedy appears to be available for users and concludes by examining some of the public policy ramifications of this status quo.

*A. Service Providers Cannot Assert Users’ Rights  
on a Fourth Amendment Theory*

As a general matter, federal courts strongly disfavor one party asserting another party’s legal claims or rights. This is related to the rationale requiring a party to have standing and is thus rooted in Article III of the Constitution.<sup>139</sup> There are certain exceptions, such as “organizational standing,” which allows organizations to assert certain rights on behalf of their members,<sup>140</sup> as well as certain “close” relationships, such as that of doctor and patient,<sup>141</sup> but these exceptions are few and far between. Additionally, specifically in the context of the Fourth Amendment, a separate doctrine that is also called “third-party standing” precludes the application of the exclusionary rule<sup>142</sup> when the “Fourth Amendment injury” was in-

---

139. See generally, e.g., *Ashwander v. Tenn. Valley Auth.*, 297 U.S. 288, 347 (1936) (Brandeis, J., concurring) (“The Court will not pass upon the validity of a statute upon complaint of one who fails to show that he is injured by its operation.”).

140. *Warth v. Seldin*, 422 U.S. 490, 511 (1975).

141. See, e.g., *Griswold v. Connecticut*, 381 U.S. 479 (1965) (finding that doctors have standing to assert their patients’ privacy interests in obtaining contraception); see also *Kowalski v. Tesmer*, 543 U.S. 125, 130 (2004) (outlining requirements for this exception).

142. In the context of criminal law, the “exclusionary rule” deems inadmissible evidence that was gathered in violation of constitutional rights. *Weeks v. United States*, 232 U.S. 383 (1914), was a seminal case that announced the rule in the context of the Fourth Amendment. *Mapp v. Ohio*, 367 U.S. 643 (1961), extended the rule to state court prosecutions as well. See generally Richard M. Re, Note, *The Due Process Exclusionary Rule*, 127 HARV. L. REV. 1885, 1893–907 (2014) (describing commonly stated rationales for the exclusionary rule).

flicted upon someone other than the defendant.<sup>143</sup> For example, an individual storing his contraband at a friend's house could not avail himself of the exclusionary rule if the friend's house was searched without a warrant.<sup>144</sup> This is based on the view that the Fourth Amendment is a "personal right" that cannot be vicariously asserted.<sup>145</sup>

In the past few years, several technology companies subjected to subpoenas and warrants have filed suits against the federal government in an effort to assert their customers' interests. These companies have not been successful.

In February 2017, Judge James Robart in the Western District of Washington, held that Microsoft did not have standing to assert its customers' Fourth Amendment rights.<sup>146</sup> The lawsuit concerned a practice of the Department of Justice in which the Department obtained warrants for data stored in the cloud and then imposed "gag orders" under section 2705(b) of the SCA, thus preventing Microsoft from disclosing the existence of the warrants to their customers.<sup>147</sup> Microsoft sued on behalf of itself under the First Amendment, as well as on behalf of its customers under the Fourth Amendment. The judge denied a motion to dismiss on the First Amendment claims but dismissed the Fourth Amendment claim for lack of standing. Acknowledging the "difficult situation" the holding creates for "Microsoft's customers [who] will be practically unable to vindicate their own Fourth Amendment rights"<sup>148</sup> because they would never learn about the intrusion in the first place, the court nonetheless held that decades of Fourth Amendment jurisprudence commanded the result. Judge Robart noted that the "conundrum . . . is not unique to the case; it is also true of the victim of an unreasonable search in a stranger's home."<sup>149</sup>

---

143. See generally Pugh, *infra* note 159, at 987–96 (discussing third-party standing in the context of electronic data).

144. See *Alderman v. United States*, 394 U.S. 165 (1968).

145. *Id.* at 174.

146. *Microsoft Corp. v. DOJ*, 233 F. Supp. 3d 887, 915 (W.D. Wash. 2017).

147. See *supra* note 62 and accompanying text.

148. *Microsoft Corp.*, 233 F. Supp. at 916.

149. *Id.* (citing *Alderman*). Note that although this is a good example of a case in which the government is not incentivized to follow the command of the Fourth Amendment, it is actually a bad analogy. The general theory for disallowing third parties to avail themselves of the exclusionary rule is that no right *of those third parties* had been violated, and in fashioning the rule, the Supreme Court did not believe the benefits of additional deterrence outweighed the costs. See *Alderman*, 394 U.S. at 174–75. In *Microsoft*, however, the data sought belonged to customers, who absolutely did have a Fourth Amendment interest in it and would be able to assert that interest had they known about it. The only reason they were not able to

An even more recent case in New York state court reached a similar result. There, the New York District Attorney's office issued 381 search warrants to Facebook in connection with an investigation into a large-scale insurance fraud conspiracy. Facebook moved to quash the warrants. The New York Supreme Court<sup>150</sup> denied the motion, holding that "Facebook could not assert the Fourth Amendment rights of its users."<sup>151</sup> Facebook appealed the order, but the Appellate Division affirmed. It held that New York civil procedure law did not provide for interlocutory review of a denial of a motion to quash a criminal warrant.<sup>152</sup> The New York Court of Appeals affirmed the decision as well.<sup>153</sup>

Facebook's unsuccessful argument rested on a distinction between warrants and subpoenas. Because the warrants at issue did not bear all the indicia of a traditional warrant (for example, unlike a typical warrant executed by law enforcement officers, the warrant here directed a third party to turn over records owned by its customers), the orders, Facebook argued, were more like subpoenas than true warrants.<sup>154</sup> Facebook hoped to make the distinction because under New York civil procedure law, a subpoena would be reviewable on an interlocutory appeal.<sup>155</sup> But it bears more general significance because the SCA specifically provides for government access to customer data by subpoena, which only requires a showing of "*reasonable grounds* to believe that the contents of a wire or electronic communication, or the records or other information sought, are relevant and material to an ongoing criminal investigation."<sup>156</sup>

---

do so is that they did not know about it because of the gag order. Microsoft's inability to assert its customers' Fourth Amendment rights for them is therefore rooted in traditional third-party standing doctrine, and not the special exclusionary rule doctrine from *Alderman*.

150. In New York's court system, the Supreme Court is the trial court, the Appellate Division is the intermediate appellate court, and the Court of Appeals is the court of final review.

151. *In re* 381 Search Warrants Directed to Facebook, Inc. 132 A.D.3d 11, 14 (N.Y. App. Div. 2015).

152. *Id.* at 23–24.

153. *In re* 381 Search Warrants Directed to Facebook, Inc. 78 N.E.3d 141, 142 (N.Y. 2017).

154. *See id.* at 146.

155. *Id.* ("[A] motion to quash a subpoena issued prior to the commencement of a criminal action, even if related to a criminal investigation, is 'civil by nature' . . . [and] an order resolving a motion to quash such a subpoena is a final and appealable order." (emphasis removed)).

156. 18 U.S.C. § 2703(d) (emphasis added). The Appellate Division explained that "[t]his is essentially a reasonable suspicion standard." 132 A.D.3d at 22 n.8.

A warrant, meanwhile, requires a showing of probable cause. The Appellate Division explained that by trying to argue that the warrants at issue were “subpoenas” that it could move to quash on a third-party standing theory, Facebook was trying to have its cake and eat it too:

Facebook cannot have it both ways. On the one hand, Facebook is seeking the right to litigate pre-enforcement the constitutionality of the warrants on its customers’ behalf. But neither the Constitution nor New York Criminal Procedure Law provides the targets of the warrant the right to such a pre-enforcement challenge. On the other hand, Facebook also wants the probable cause standard of warrants, while retaining the pre-execution adversary process of subpoenas. We see no basis for providing Facebook a greater right than its customers are afforded.<sup>157</sup>

Thus, at least the New York courts have concluded that third-party standing for orders issued under the SCA are unavailable because they should be seen as warrants subject to Fourth Amendment third-party standing doctrine. An assisted MITM order would likely face the same fate. Although the case law is highly underdeveloped, the dearth of any positive precedent renders it unlikely that service providers could successfully raise their customers’ privacy rights on their behalf.<sup>158</sup> Until that changes,<sup>159</sup> a service provider would have to rely on a personal injury, such as the First Amendment theory outlined in Part III.

Because companies cannot assert the Fourth Amendment rights of their users, and because, as Judge Robart observed, users are unlikely to know they are being targeted and thus cannot assert

---

157. 132 A.D.3d at 22.

158. In 2008, the Foreign Intelligence Surveillance Court of Review found that Yahoo had standing under the Protect America Act to assert a Fourth Amendment challenge on behalf of its customers to a warrantless surveillance directive. *In re Directives to Yahoo! Inc.* Pursuant to Section 105B of the Foreign Intelligence Surveillance Act, No. 08-01 (Fisa Ct. Rev. 2008), <https://www.dni.gov/files/documents/0909/FISC%20Merits%20Opinion%2020080822.pdf> [<https://perma.cc/4R6A-5GHB>] (approved for public release by the DNI 20140909). However, the court was careful to note that the third-party standing was specially granted by the Protect America Act. *Id.* at 10–11.

159. The *Microsoft* case has already led some academics to argue that third-party standing doctrine should be relaxed because companies are better suited than their customers for challenging warrants seeking data in the cloud. *See, e.g.*, Margot E. Kaminski, *Standing After Snowden: Lessons on Privacy Harm from National Security Litigation*, 66 DEPAUL L. REV. 413, 436–38 (2017); Sarah E. Pugh, Comment, *Cloudy with a Chance of Abused Privacy Rights: Modifying Third-Party Fourth Amendment Standing Doctrine Post-Spokeo*, 66 AM. U. L. REV. 971 (2017).

their own Fourth Amendment rights, Fourth Amendment jurisprudence is unhelpful. The next Section analyzes the user's *First* Amendment rights (again recognizing that the same obstacle of not knowing that one is the subject of an investigation would, as a practical matter, foreclose those claims anyway).

*B. The First Amendment Does Not Protect Users*

Putting aside the fact that users are unlikely to assert their own rights because they would not know that they are the subject of an investigation, the users whose messages are being decrypted would at least have a valid, personal Fourth Amendment claim. The users whose messages are being *forged*, however, would not have any sort of Fourth Amendment claim, since no "search" of their messages is being performed. This Section investigates whether those users could assert a First Amendment injury instead.

To make the hypothetical as friendly as possible, assume that the government is in fact forging messages, rather than simply passing them along.<sup>160</sup> As a threshold matter, the users will not have third-party standing to assert the service providers' First Amendment rights in a case in which the service provider simply consents to falsely attesting about a key that the government supplied. Although the keys are conceptually associated with a user (and it is in that sense that one could say they "belong" to them), they are still data legally owned by the service providers. If that service provider chooses to allow the government to falsely present a key as belonging to a user, the user (or perhaps her counterparty in a conversation) may have a breach-of-contract claim against the service provider, as well as a possible civil claim under the SCA.<sup>161</sup> But because, as I argue in Part III, the service provider has ample ability to assert its own First Amendment rights, a user could not stand in its shoes to assert them.<sup>162</sup>

---

160. In Figure 5 in Part I.C, the message that was being forged by the government to send to Bob as if it came from Alice was in fact the same message Alice sent. But of course, it doesn't have to be that way. For example, suppose Alice is planning a heist with Bob and messages him "take the gun, leave the cannoli." The government could intercept the encrypted message and modify it to read "take the gun and *bring* the cannoli," thus ensuring the availability of a snack when the police execute their sting operation.

161. See 18 U.S.C. § 2707(a) (2012).

162. See generally *Kowalski v. Tesmer*, 543 U.S. 125, 130 (2004) (explaining that a prerequisite to asserting standing of another party's rights is a showing of "a hindrance to the possessor's ability to protect his own interests" (internal quotation marks omitted)).

Thus, the First Amendment claim must be a personal one. One theoretical argument is that if the government is able to impersonate users' speech, it is in a sense compelling them to speak, thus creating a First Amendment injury. Moreover, in accordance with Professor Sacharoff's argument that compelled speech is inconsistent with the First Amendment's role in protecting listeners' interests, recipients of the government's forged messages are being "tricked" into believing they are communicating with someone they are not.<sup>163</sup>

To emphasize, the issue is not just that the government directs a key to be sent in the name of a particular user. The purpose of the assisted MITM attack is to allow the government to forge *arbitrary* messages from the user—that is, it allows the government to construct messages that will appear to be cryptographically signed by the user. It allows the government to do remotely, and at scale, what the police officer in *Hinton* and *Roden* was able to do by happenstance of a search incident to arrest.<sup>164</sup> To use a physical analogy, the scenario is not just a government official signing a letter with a target's name, but rather copying every citizen's handwriting so perfectly that the government can construct any possible missive to look as if it were written by any citizen.

It is difficult to construct an analogous fact pattern to something like this—where the government can routinely send messages and pretend that the messages are coming from somebody else, without any way for the listener to verify the identity of the sender. Cryptography is almost uniquely designed to address this problem. But one idea that comes to mind is the right of prisoners with respect to mail. Under *Thornburgh v. Abbott*,<sup>165</sup> prison guards may censor mail as long as the regulations guiding the censorship are "reasonably related to legitimate penological interests."<sup>166</sup> *Thornburgh* overruled an earlier case that invalidated a mail censorship system under something that resembled an intermediate scrutiny standard.<sup>167</sup> But in dicta, *Thornburgh* pointed out that the earlier case concerned itself with outgoing mail, whereas the regulation at issue in *Thornburgh* dealt with *incoming* mail. Since most censorship regulations are directed at maintaining order and security, which is far more likely to be threatened by incoming mail than outgoing

---

163. See Sacharoff, *supra* note 126.

164. See *supra* note 40.

165. 490 U.S. 401 (1989).

166. *Id.* at 413 (internal quotation omitted) (citing *Turner v. Safley*, 482 U.S. 78, 89 (1987)).

167. See *Procunier v. Martinez*, 416 U.S. 396, 413–14 (1974).

mail, the cases may also be distinguished on those grounds, preserving the previous holding.<sup>168</sup>

In any case, consider a hypothetical fact pattern in which prison officials repeatedly sent mail in the name of an inmate without his or her knowledge or consent. Recipients of the letters would have no way of knowing that the statements are not actually being made by the inmate, and the inmate would have no way of proving otherwise (at least for as long as she was incarcerated). Surely, such a program would raise First Amendment concerns—and in the context of prisons and “legitimate penological interests,” it would be hard to imagine that it would pass even rational basis review. One could construe this either as the denied freedom of speech of the inmate—with the prisoner essentially being compelled to speak to the world—or, as Professor Sacharoff suggests, the injured listener interest of the recipients of the mail.<sup>169</sup> In any case, the injury is substantial.

And yet, notwithstanding the “listener’s injury” formulation, I am hesitant to conclude that, absent an extreme situation like the prison hypothetical just described, there exists a freestanding constitutional right to what amounts to “integrity of identity.” Of course, it may come up incidentally. For example, if the government repeatedly wrote messages on behalf of another user associating her with a political faction, that might violate her freedom of association. But what if the government simply posted an occasional movie review, and digitally signed it as her? Would that be illegal? It is not compelled speech, at least not as that term is usually understood. Absent specific factual circumstances, it does not necessarily interfere with her ability to access or enjoy any protected speech, nor to associate with any group. Furthermore, unlike the prison hypothetical, it does not foreclose the individual distancing herself from the comments with more speech.<sup>170</sup>

---

168. See 490 U.S. at 411–14.

169. See Sacharoff, *supra* note 126.

170. Another intuitive concern with the government sending messages as if they were sent by some individual is a species of privacy invasion. By “privacy,” I mean a broad concept of personal integrity as sometimes applied to the “privacy” rights of bodily autonomy, such as those recognized in *Roe v. Wade* and *Lawrence v. Texas*. See *Roe v. Wade*, 410 U.S. 113, 152–54 (1973) (discussing abortion rights in the context of privacy); see also *Lawrence v. Texas*, 539 U.S. 558, 565 (2003) (discussing the case against the backdrop of *Eisenstadt v. Baird*, 405 U.S. 438, 453 (1972)). The privacy rights in those cases were constitutionally protected through a notion of “penumbras” of the Bill of Rights, see *Roe*, 410 U.S. at 152, but it is unclear how that translates to an injury like a MITM attack, where no restriction whatsoever is being placed on the individual.

And yet, intuitively, it feels very troubling. To be sure, we may demand a good reason for the government's behavior—otherwise it may not survive rational basis review. But such reasons are not hard to imagine. For example, the *Washington Post* has reported that the FBI has, on at least one occasion, deployed malware designed to discover a target's IP address (and thus, location) by creating a website that looked like it was operated by the Associated Press.<sup>171</sup> Had it been able to leverage the website's authentication system,<sup>172</sup> it would have been extremely difficult, if not impossible, to determine that the website was not in fact published by the media company.<sup>173</sup> Similarly, targets would be more likely to install software sent to them by a friend; by forging a message to look like it was coming from a known contact, the government could easily trick a target into installing malware that assists in tracking the target's location or otherwise assists in law enforcement operations. In much the same way, an assisted MITM attack would allow the government to communicate with someone who thinks they are communicating with a known associate, with full confidence of that belief founded on the cryptographic protocols implemented in their messaging services.

---

On the other hand, it is clear that there is *some* recognition of privacy rights for an injury of the kind discussed in this Note. For example, an individual posing as another online may be liable in tort for appropriation of likeness, a species of privacy right. See Bradley Kay, Note, *Extending Tort Liability to Creators of Fake Profiles on Social Networking Websites*, 10 CHI.-KENT J. INTELL. PROP. 1 (2010). The problem, however, is that the government enjoys sovereign immunity that it has not waived for privacy torts. See 28 U.S.C. § 2680 (2012) (listing torts for which sovereign immunity is waived under the Federal Torts Claims Act). Additionally, the Lanham Act waives sovereign immunity for both the federal government and state governments, 15 U.S.C. § 1122(a)–(b) (2012), and so individuals may have remedies for, e.g., false endorsement claim. § 1125(a). Consider, for example, Bruce Springsteen suing the President for playing his music at a White House event. The Lanham Act does not reach misappropriation of likeness claims, however. See also Paul v. Davis, 424 U.S. 693 (1976) (holding that reputation alone is not a constitutionally protected interest).

171. Ellen Nakashima & Paul Farhi, *FBI Lured Suspect with Fake Web Page, but May Have Leveraged Media Credibility*, WASH. POST (Oct. 28, 2014), [https://www.washingtonpost.com/world/national-security/fbi-lured-suspect-with-fake-web-page-but-may-have-leveraged-media-credibility/2014/10/28/e6a9ac94-5ed0-11e4-91f7-5d89b5e8c251\\_story.html](https://www.washingtonpost.com/world/national-security/fbi-lured-suspect-with-fake-web-page-but-may-have-leveraged-media-credibility/2014/10/28/e6a9ac94-5ed0-11e4-91f7-5d89b5e8c251_story.html) [<https://perma.cc/X982-E773>].

172. Websites use authentication systems that operate on the same fundamental principles as those in end-to-end encrypted messaging systems. See generally GLOBALSIGN, *supra* note 30.

173. If the government successfully did that, user's web browsers would display the "green lock" or a similar graphic, making it look as if there was a secure connection to the real website. See MOZILLA, *supra* note 99.

Indeed, one can imagine motivations even outside the law enforcement context. Suppose, for example, that a small town was trying to prop up a local business and did so by writing glowing reviews in the names of its citizens on Yelp. That is certainly a rational basis, but intuitively, it feels troubling. And yet, I know of no constitutional reason that the town (or state or federal) government could not undertake such a program.<sup>174</sup>

But so what? We began this discussion by focusing on authentication. What if the government were only allowed to order the replacement of encryption keys and not verification keys? Then, the government could bypass end-to-end encryption, finally submitting it to tried-and-true Fourth Amendment doctrine. With verification keys not subject to assisted MITM orders, people's identities would remain intact.

The problem is that cryptography does not work that way. Either the government has the ability to falsely associate itself with another user's identity, or it doesn't. If it doesn't, end-to-end encryption is here to stay—recall that a verified channel is all that is needed to ensure a confidential conversation.<sup>175</sup> Users could use the verification keys to share *new* encryption keys to create a new, secure channel, and the government would have no way of accessing the new channel. In other words, an assisted MITM order, if it were allowed, has to be all or nothing.

---

174. As mentioned in note 171, *supra*, the users may be able to collect damages from the federal government in an analogous scenario under the Lanham Act on a false endorsement theory. But even that result is not constitutionally mandated.

This hypothetical brings to mind the plot of the Bollywood movie *Poster Boys*, in which three men discover that their likenesses were used on a poster promoting vasectomies. See *POSTER BOYS* (Sony Pictures 2017). Of course, public service announcements about family planning (in support of either birth control or procreation) are not exclusively the province of fiction. See, e.g., Adam Century, *China's Colorful Family-Planning Propaganda*, *THE ATLANTIC* (Nov. 18, 2013), <https://www.theatlantic.com/china/archive/2013/11/chinas-colorful-family-planning-propaganda/281594/> [<https://perma.cc/FXB3-5TB5>] (providing examples of government messaging promoting family planning in support of the one-child policy); Joshua Keating, *Rap video urges Singaporeans to get busy making babies*, *FOREIGN POLICY* (Aug. 8, 2012 4:23 PM), <http://foreignpolicy.com/2012/08/08/rap-video-urges-singaporeans-to-get-busy-making-babies/> [<https://perma.cc/7WYP-B4T6>] (displaying a government-sponsored rap music video encouraging listeners to procreate). Unlike the example of the Yelp reviews, however, these examples—both real and fictional—do not need to rely on cryptography, since the likeness of any actor would be noticed visually.

175. See *supra* note 31 and accompanying text.

If the order is allowed and verification keys could be forged, society must be comfortable with the reality that the government will sometimes be able to speak in the digital voices of its citizens, without their permission or even knowledge. Courts will have to fashion tests to balance interests rarely considered together—free-floating free speech interests versus law enforcement’s interest in surveillance. Is it acceptable for police officers to digitally pose as a target and express opinions in her name? If so, can it be on any topic? And for how long? This Note does not propose a resolution to this tension. But the tension is an inescapable reality of the cryptographic world we have created.

### CONCLUSION

End-to-end encryption is usually discussed in the context of information privacy, and it has befuddled the government because it operates outside the rules of traditional privacy safeguards. Through Fourth Amendment doctrines and related jurisprudence, American society has struck a balance between privacy interests and law enforcement. We have almost two-and-a-half centuries’ worth of experience in procedures to evaluate that tradeoff. It is no wonder that a sudden wrench in that history feels so disruptive.

But “end-to-end encryption” is little more than a marketing term. The sort of confidentiality it promises are consequences of basic mathematical facts, whose efficacy is based not on any sort of privacy concerns but on the knowledge that individuals know whom they’re speaking with. That is the foundation we must be willing to disturb if we decide that end-to-end encryption is too dangerous to allow. As it stands, we have entrusted third parties to accurately connect us to our counterparties, and for as long as that trust is respected, traditional First Amendment doctrine protects those parties from conscription into government surveillance. If those parties do not object to an assistance order, however, citizens are without recourse. There is no question that the Framers did not anticipate this conundrum when they were drafting the Bill of Rights. Now that the reality is upon us, however, we should determine if we should allow it. If it is a step we are willing to take, we ought to take it deliberately.

**CRAMDOWN INTEREST RATES IN THE  
COMMERCIAL CREDIT CONTEXT:  
ARGUMENTS FOR THE TWO-  
STEP APPROACH IN  
CHAPTER 11**

*MATTHEW WIENER\**

Introduction .....	224
I. Cramdown Plans and Interest Rates Under Chapter 11 .....	228
A. Interest Rates, Present Value, and Market Value .	229
B. Judicial Methods for Determining Cramdown Interest Rates .....	233
1. Formula Rates .....	233
2. Forced Loan or Coerced Loan.....	235
3. Cost of Funds .....	235
4. Contract Rate.....	237
C. The <i>Till</i> Decision and Adoption of the Formula Rate Approach .....	238
D. <i>Till</i> 's Famous Footnote 14 and the Two-Part Approach.....	239
II. <i>Momentive</i> Performance Materials and the Extension of the Formula Rate Approach into Chapter 11.....	240
A. <i>Momentive</i> Facts and Arguments.....	240
B. Judge Drain's Reasoning in <i>Momentive</i> and Adoption of the Formula Rate Approach .....	243
C. The Second Circuit's Recent Opinion Overturned <i>Momentive</i> and Implemented a Two-Step Approach.....	245
III. The Business of Banking .....	246
A. Bank Debt .....	247
B. Capital Markets Debt .....	249
C. Interest Rates in Commercial Credit Markets ....	250
D. Present Value and "Profit" .....	251

---

\* J.D., 2018, New York University School of Law; B.A., 2009, Vassar College. I would like to thank Marcel Kahan, Michael Walsh, and Paul Wiener for their thoughtful and incisive comments, as well as the New York University Annual Survey of American Law for providing the opportunity to publish this Note.

- E. The Market Value of the *Momentive* First Lien Notes Decreased as a Result of the Use of the Formula Rate Approach ..... 254
- IV. Arguments in Favor of the Formula Rate Approach .. 256
- V. Arguments Against the Formula Rate Approach and Its Use in the Chapter 11 Context ..... 259
  - A. The Formula Rate Approach Reduces Certainty and Predictability Which Will Negatively Affect Credit Markets and Increase Interest Rates Across the Board ..... 260
  - B. The Formula Rate Approach Fails to Compensate Secured Creditor Claims in the Risky Bankruptcy Context ..... 264
  - C. Because Chapter 11 Is Fundamentally Dissimilar to Chapter 13, the Formula Rate Approach Does Not Apply ..... 268
  - D. The Formula Rate Approach Inappropriately Creates an Option and Provides Negotiating Leverage for Chapter 11 Debtors ..... 270
- VI. The Two-Step Approach Adopted by the Second Circuit is Better Suited for Chapter 11 and the Commercial Credit Context ..... 271
  - A. The Two-Step Approach Is Easily Administered and Reduces Evidentiary Burdens on the Court and Debtor ..... 272
  - B. The Two-Step Approach Compensates Creditors for Risk Without Unduly Burdening Debtors..... 274
  - C. The Two Step Approach Facilitates Outcome Certainty and Protects the Justified Expectations of Creditors and Debtors ..... 275
- Conclusion ..... 276

INTRODUCTION

In *In re MPM Silicones, LLC*<sup>1</sup> (hereinafter “*Momentive*”), the Bankruptcy Court for the Southern District of New York addressed certain elements of the Bankruptcy Code’s “cramdown”<sup>2</sup> exception to the requirement that a plan of reorganization be approved by all

---

1. *In re MPM Silicones, LLC*, No. 14-22503-RDD, 2014 WL 4436335 (Bankr. S.D.N.Y. Sept. 9, 2014), *aff’d*, 531 B.R. 321 (S.D.N.Y. 2015).

2. Under the Bankruptcy Code, a “cramdown” refers to the confirmation of a debtor’s plan of reorganization over the objections of its creditors, provided certain requirements are met. *See infra* Part I for further explanation of the cramdown provisions of the Bankruptcy Code.

classes of claims and interests.<sup>3</sup> In particular, the *Momentive* Court was asked to determine the interest rate required on deferred payments made by a debtor pursuant to a cramdown plan in order to give its secured creditors, who voted against the plan, the present value of their secured claims. The methodology for determining such interest rates selected by the court was highly controversial and became the subject of significant litigation and academic discourse. Not only did it have a negative impact on the value of the secured claims at issue in *Momentive*, but it also had the potential to broadly affect the likelihood of plan confirmation by Chapter 11 debtors,<sup>4</sup> and the relative rights and bargaining power of creditors

---

3. See 11 U.S.C. § 1129(a)(7)(A)(i) (2010) (“The court shall confirm a plan only if . . . [w]ith respect to each impaired class of claims or interests . . . each holder of a claim or interest of such class . . . has accepted the plan . . .”).

4. Debtors may file for protection under various chapters of the Bankruptcy Code, though the three most commonly used are Chapters 7, 11, and 13. While Chapter 7 is not directly implicated in this Note, it is nonetheless significant, since a plan may not be approved under Chapters 11 or 13 unless the proposed payments under such plans are equal to the amount a creditor would receive under Chapter 7. See LYNN M. LOPUCKI & ELIZABETH WARREN, SECURED CREDIT: A SYSTEMS APPROACH 95 (7th ed. 2011). Under Chapter 7, a debtor’s nonexempt assets are surrendered to a bankruptcy trustee and liquidated. The proceeds of the liquidation are distributed to the debtor’s creditors, after which the debtor’s obligations may be discharged. *Id.* at 94; see also 7 COLLIER ON BANKRUPTCY ¶ 700.01 (Alan N. Resnick & Henry J. Sommers eds., 16th ed. 2010). Under Chapter 11, a debtor is given the opportunity to reorganize its business or to liquidate all or a portion of its assets in an orderly manner. *Id.* at ¶ 1100.01. The debtor (or in the case of an involuntary bankruptcy, the debtor’s creditors) proposes a plan of reorganization, which may include extension or modification of its obligations, such as reductions in interest rates and the principal amount owing. *Id.* During a Chapter 11 bankruptcy, the debtor remains in control of its assets, which it may operate, utilize, or sell. If the debtor’s plan is confirmed, the debtor ‘emerges’ from bankruptcy. Its obligations under the plan replace its obligations prior to bankruptcy and the debtor’s remaining debts, if any, are discharged. See LOPUCKI & WARREN at 95. There are few statutory limits on who may file for bankruptcy under Chapter 11. Chapter 13 is similar to Chapter 11, in that it permits a debtor, who remains in control of its assets, to propose a payment plan for its existing secured and unsecured creditors, usually over the course of three or five years. *Id.* In exchange for maintaining possession of its assets, the debtor promises to repay its pre-bankruptcy obligations out of future income. BARRY E. ADLER, DOUGLAS M. BAIRD, & THOMAS H. JACKSON, BANKRUPTCY: CASES, PROBLEMS, AND MATERIALS 38 (4th ed. 2007). Chapter 13 is only available to debtors with regular income and unsecured and secured debts of \$394,725 and \$1,184,200, respectively. See *id.*; see also 7 COLLIER ¶ 1300.14[3]. Thus, most large corporations are not eligible to file for bankruptcy under Chapter 13. If a debtor completes its Chapter 13 plan, its obligations are discharged and it remains in possession of its property. See 7 COLLIER ¶ 1300.01; see also LOPUCKI & WARREN at 95.

and debtors generally.<sup>5</sup> Fortunately, an October 2016 opinion by the Court of Appeals for the Second Circuit resolved some of this controversy.<sup>6</sup> It rejected the bankruptcy court's methodology and instead adopted one based, more appropriately, on ascertainable market rates of interest. Regardless, the issue remains highly polemical and subject to significant disagreement among courts, academics, and commentators.

Though much has been written about *Momentive* and despite the significant financial impacts of the decision on secured creditors, few have fully considered *Momentive's* impact on creditors in the commercial lending context. In fact, most who have addressed the subject have advocated for approaches to determine cramdown interest rates that are largely favorable to debtors but have failed to also consider the broader impacts on commercial credit markets. In this Note I will discuss the *Momentive* decision and critique the methodology it used to assign a below-market interest rate on replacement notes issued to objecting holders of secured claims under a cramdown plan. I will argue bankruptcy courts confronted with the determination of cramdown interest rates should instead follow the approach adopted by the Second Circuit, which better reflects prevailing market rates and more accurately provides crammed down creditors with the present value of their claims as required by the Bankruptcy Code.

Part I introduces the provisions of the Bankruptcy Code that allow confirmation of a plan over the objections of secured creditors and provides background on interest rates, present value, and market value. It also describes four methodologies courts have considered in determining which interest rate mostly fairly compensates creditors who receive replacement loans under such plans: the formula rate, forced loan, cost of funds, and contract rate approaches.<sup>7</sup> Part I then discusses the United States Supreme Court's decision in *Till v. S.C.S. Credit Corp.*,<sup>8</sup> which adopted the formula rate approach as the appropriate standard for determining cramdown interest rates in Chapter 13 bankruptcy cases.

Part II examines *Momentive's* extension of the formula rate approach from the Chapter 13 to the Chapter 11 context, beginning

---

5. See Waltraud S. Scott, *Deferred Cash Payments to Secured Creditors in Cram Down of Chapter 11 Plans: A Matter of Interest*, 63 WASH. L. REV. 1041, 1042 (1988).

6. *In re MPM Silicones, LLC*, 874 F.3d 787 (2d Cir. 2017).

7. This Note will use the term "loan" to describe all instances in which an entity "lends" money to a borrower, including credit facilities, such as term loans and revolving loans, and debt securities, including bonds and notes.

8. 541 U.S. 465 (2004).

with a discussion of the facts and background of that case, followed by an analysis of the court's reasoning. It then details the Second Circuit's holding, which overruled the bankruptcy court and opted instead for market (rather than court-improvised) valuations of the appropriate interest rate in Chapter 11 cases.

Part III describes the business of banking, the types of debt accessible to commercial borrowers, and the methodology of determining interest rates in commercial credit markets. It then argues courts should follow the Second Circuit and look to prevailing market rates when determining cramdown interest rates and that the exclusion of "profit" from such rates misconstrues the nature of interest rates. This effectively precludes secured creditors from receiving the present value of their claims. Part III then demonstrates the extent to which the replacement notes under *Momentive* fail to provide the value to holders of secured claims required by the cramdown provisions of the Bankruptcy Code.

Part IV and Part V consider, respectively, arguments for and against the formula rate approach, each focusing on administrability, evidentiary burdens on both creditors and debtors, reasonable compensation for risk, and outcome certainty in the bankruptcy process.

Finally, Part VI argues the formula rate approach should be abandoned in Chapter 11 in favor of the market-based analysis adopted in the Second Circuit. Had the Second Circuit's two-step approach been used in *Momentive*, the debtor's secured noteholders would have received an interest rate on their notes better keyed to prevailing market rates and avoided the steep losses in the market value of those securities post-confirmation. Unlike the formula rate approach, the two-step method is more easily administered because it relies, in the first instance, on ascertainable market rates and therefore reduces the likelihood of litigating disputes. Furthermore, in the event of litigation, the factual issues are likely to be narrow, thereby reducing the evidentiary burden as well as any costs associated with the production of such evidence. Moreover, this method is the single methodology which compensates creditors for their risk of lending to debtors emerging from Chapter 11 without unduly burdening such debtors who would have to seek financing from the market regardless. It also protects the expectations of both creditors and debtors with regard to their negotiations over contractual terms and best balances their interests according to a visible market.

I.  
CRAMDOWN PLANS AND INTEREST RATES  
UNDER CHAPTER 11

Section 1129(b) of Title 11 of the United States Code (the “Bankruptcy Code”) permits the confirmation of a plan over the dissent of one or more classes of impaired claims or interests<sup>9</sup> (hereinafter, all references to section 1129 refer to the Bankruptcy Code). While section 1129(b) requires a plan proponent to meet all the other requirements of section 1129(a),<sup>10</sup> including, for instance, feasibility,<sup>11</sup> so-called “cramdowns” are a powerful tool in the hands of a debtor.<sup>12</sup> In order to protect the interests of dissenting classes, section 1129(b) states that for a cramdown to be approved, the plan must not “discriminate unfairly”<sup>13</sup> and must be “fair and equitable”<sup>14</sup> to all impaired classes of claims. Section 1129(b)(2) provides three mechanisms by which a bankruptcy court may find a plan fair and equitable as to dissenting or impaired claims, any of which may independently satisfy this requirement. Importantly, a debtor may comply with section 1129(b)(2) by

---

9. See 11 U.S.C. § 1129(b) (2010) (“[T]he court, on request of the proponent of the plan, shall confirm the plan notwithstanding the requirements of [Section 1129(a)] if the plan does not discriminate unfairly, and is fair and equitable, with respect to each class of claims or interests that is impaired under, and has not accepted, the plan.”).

10. See *id.* (“[A]ll of the applicable requirements of subsection (a) of [section 1129] other than paragraph (8) [must be] met”); see also *In re DBSD N. Am., Inc.*, 419 B.R. 179, 204 (Bankr. S.D.N.Y. 2009) (stating Section 1129(b)(1) is a fundamental provision of the Bankruptcy Code, which “provides that if all applicable requirements of section 1129(a) are met other than section 1129(a)(8) (which requires that every impaired class vote in favor of the Plan), a plan may be confirmed so long as the requirements set forth in section 1129(b) are satisfied.”) (emphasis in original).

11. See 11 U.S.C. § 1129(a)(11) (2010).

12. See Daniel R. Wong, *Chapter 11 Bankruptcy and Cramdowns: Adopting a Contract Rate Approach*, 106 NW. U. L. REV. 1927, 1933 (2012).

13. 11 U.S.C. § 1129(b)(1). The term “unfair discrimination” means, in general, that no class of dissenting claims may receive value different from the value given to other similarly situated claims. See *In re Young Broad., Inc.*, 430 B.R. 99, 139–40 (Bankr. S.D.N.Y. 2010) (stating that as defined in Section 1129(b)(1), “a plan unfairly discriminates when it treats similarly situated classes differently without a reasonable basis for the disparate treatment.”).

14. 11 U.S.C. § 1129(b)(1); see also David Griffiths, *Momentous Decision in Momentive Performance Materials: Cramdown of Secured Creditors – Part I*, WEIL BANKRUPTCY BLOG (Sept. 9, 2014), <https://business-finance-restructuring.weil.com/chapter-11-plans/momentous-decision-in-momentive-performance-materials-cram-down-of-secured-creditors-part-i/> [<https://perma.cc/B8HT-BMWV>].

making payment to holders of secured claims gradually over time, even if junior claims are paid sooner.<sup>15</sup>

Section 1129(b)(2)(A)(i) is most relevant to the present discussion of *Momentive*, as it essentially permits a plan proponent to propose new loans for dissenting secured creditors,<sup>16</sup> provided those creditors retain liens securing their claims<sup>17</sup> and receive deferred cash payments with a value at least equal to the value of their interests in the collateral securing their claims on the confirmation date.<sup>18</sup> Within the debate over cramdowns, perhaps the only point of agreement is that the language of section 1129(b)(2) requires bankruptcy courts to conduct a present value analysis in assessing cramdown interest rates.<sup>19</sup>

#### A. *Interest Rates, Present Value, and Market Value*

An understanding of interest rates is fundamental to a discussion of the present value analysis required by the Bankruptcy Code. Interest rates reflect the price of money.<sup>20</sup> The real rate of interest is exclusive of inflation.<sup>21</sup> In contrast, the nominal interest rate reflects the current or inflation-adjusted price of money.<sup>22</sup> Nominal interest rates on defaultable securities are comprised of three ele-

---

15. See 7 COLLIER, *supra* note 4, at ¶ 1129.04[2]; see also *Matter of Penn Cent. Transp. Co.*, 458 F. Supp. 1234, 1283 (E.D. Pa. 1978) (“[T]he absolute priority rule does not require sequential distributions (i.e., cash payment in full to senior creditors before any distribution is made to junior creditors), but merely that the values represented by the higher-ranking claims are fully satisfied by the values distributed under the Plan.”).

16. See 7 COLLIER, *supra* note 4, at ¶ 1129.04[2][a].

17. See 11 U.S.C. § 1129(b)(2)(A)(i)(I).

18. See 11 U.S.C. § 1129(b)(2)(A)(i)(II) (“[T]hat each holder of a claim of such class receive on account of such claim *deferred cash payments* totaling at least the allowed amount of such claim, *of a value, as of the effective date of the plan*, of at least the value of such holder’s interest in the estate’s interest in such property . . . .”) (emphasis added).

19. See *United Sav. Ass’n of Tex. v. Timbers of Inwood Forest Assocs.*, 484 U.S. 365, 377 (1988) (“Under § 1129(b) a secured claimant has a right to receive under a plan the present value of his collateral. This entitlement arises . . . from the provision of § 1129(b)(2)(A)(i)(II) . . . .”). The legislative history of the Bankruptcy Code also indicates Congress explicitly intended present value to be analyzed in the confirmation of a plan. See H.R. Rep. No. 595, at 413-18 (1978), *reprinted in* 1978 U.S.C.C.A.N. 6370-71 (“[P]roperty is to be valued as of the effective date of the plan, thus recognizing the time value of money . . . .”).

20. See DAVID W. PEARCE, *THE DICTIONARY OF MODERN ECONOMICS* 370 (2nd ed. 1983).

21. See *id.* at 313 (stating nominal is “the opposite of real.”).

22. See *id.* (defining nominal as “the measurement of economic magnitude in current prices.”).

ments, each of which relates to a different aspect of the delayed use of money: the expected rate of inflation, the real rate of interest, and a default or credit risk premium.<sup>23</sup>

Inflation, or an increase in price levels,<sup>24</sup> decreases the spending power of money over time. If inflation occurs while a loan is outstanding, the money a lender provides to a borrower will be worth less when it is repaid. To protect lenders against this loss, interest rates incorporate the market's expectation of the average rate of inflation over the life of the loan.<sup>25</sup> The real rate of interest (to be distinguished from a real interest rate, which is the opposite of a nominal interest rate) is the interest in excess of inflation on a riskless security.<sup>26</sup> It is a form of compensation to induce a lender to postpone their consumption of funds between the present and a future date.<sup>27</sup> The real rate of interest and inflation should be identical for all borrowers, though both measures vary with the term of the loan.<sup>28</sup>

Variations in interest rates result from borrower-specific default (or credit risk).<sup>29</sup> Lenders price such risk, not all of which can be compensated for in any instance.<sup>30</sup> The greater the borrower's perceived risk, the greater the inducement must be for the lender to provide them with funds.<sup>31</sup> The interest rates on United States Treasury securities are some of the lowest in the market<sup>32</sup> because

23. See Scott, *supra* note 5, at 1046; see also J. FRED WESTON & EUGENE F. BRIGHAM, *ESSENTIALS OF MANAGERIAL FINANCE*, 345-46 (6th ed. 1982). A simple method for determining the real rate of interest is to subtract the inflation factor from the nominal interest rate. *Id.*

24. See PEARCE, *supra* note 20, at 209.

25. See Scott, *supra* note 5, at 1046; EUGENE F. BRIGHAM & JOEL F. HOUSTON, *FUNDAMENTALS OF FINANCIAL MANAGEMENT* 128-29 (8th ed. 1998); WESTON & BRIGHAM, *supra* note 23, at 345-46. The rate of inflation has been positive since World War Two. BUREAU OF LABOR STATISTICS, *CONSUMER PRICE INDEX - ALL URBAN CONSUMERS*, <https://data.bls.gov/pdq/SurveyOutputServlet> (last visited January 13, 2017, 2:20 PM).

26. See BRIGHAM & HOUSTON, *supra* note 25, at 128.

27. See WESTON & BRIGHAM, *supra* note 23, at 345.

28. See Scott, *supra* note 5, at 1047.

29. See *id.*

30. See generally Philip E. Strahan, *Borrower Risk and the Price and Nonprice Terms of Bank Loans*, *Fed Res Bank of N.Y. Staff Reports*, 1 (1990), [https://www.newyorkfed.org/medialibrary/media/research/staff\\_reports/sr90.pdf](https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr90.pdf) [<https://perma.cc/V7PK-ZBQS>].

31. See Scott, *supra* note 5, at 1047; see also Strahan, *supra* note 30, at 7 (“[L]oans more likely to end in default and loans where collection in default is more difficult carry higher interest rates than other loans.”).

32. The interest rate on a 30-year U.S. Treasury Bond was 2.87% as of October 2017. In comparison, the interest rate on high quality corporate bonds (i.e.,

they are considered to be risk free.<sup>33</sup> Riskier debt, such as consumer loans and certain commercial loans, bear higher interest rates to compensate lenders for the increased likelihood of default.<sup>34</sup>

The time value of money is the idea that a dollar today is worth more than a dollar tomorrow. When a creditor makes a loan, it is not repaid immediately. Rather, it may be paid in a lump sum at the end of the loan or according to a repayment schedule. During this period, the debt declines in value from a time preference perspective. Money decreases in value over time due to the non-eliminable risk of default, inflation, and the opportunity costs associated with the inability of the creditor to use the money right away, whether for investment or any other profitable purpose.<sup>35</sup> The longer the period of time, the lower will be the value of the money when repaid.<sup>36</sup>

Present value captures the worth of future money in terms of its value now, taking into account time value. The market value of a loan is a function of the present value of the payments it will generate over time.<sup>37</sup> The face value of that loan is independent of its market value.<sup>38</sup> When a debt is repaid over time, “a creditor receives the ‘present value’ of its claim only if the total amount of the deferred payments includes the amount of the underlying claim

---

rated AAA, AA, or A) was 4.21% in October 2017. *Compare Daily Treasury Yield Curves*, U.S. DEPT. OF THE TREASURY, <https://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yield> (last visited Oct. 4, 2017) [<https://perma.cc/C9ZT-85K2>] with *Corporate Bond Yield Curve Papers and Data*, U.S. DEPT. OF THE TREASURY, <https://home.treasury.gov/data/treasury-coup-on-issues-and-corporate-bond-yield-curve/corporate-bond-yield-curve> (last visited September 18, 2018).

33. ALAN N. RECHTSCHAFFEN, CAPITAL MARKETS, DERIVATIVES AND THE LAW: EVOLUTION AFTER CRISIS 109 (2d ed. 2014) (“Treasuries are backed by the full faith and credit of the U.S. government and have little or no credit risk . . . .”); *U.S. Treasury Securities*, FINRA <http://www.finra.org/investors/us-treasury-securities> (last visited October 10, 2017) [<https://perma.cc/XBN2-2BAU>] (describing U.S. Treasury securities as “the safest investments you can make, because all Treasury securities are backed by the “full faith and credit” of the U.S. government” and which bear “[n]o call risk and virtually no liquidity, event or credit and default risk.”).

34. See Scott, *supra* note 5, at 1047.

35. See *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 474 (2004); see also John K. Pearson, Dillon Jackson & Tim Nohr, *Ending the Judicial Snipe Hunt: The Search for the Cramdown Interest Rate*, 4 AM. BANKR. INST. L. REV. 35, 37 (3d ed.1996).

36. See Pearson et al., *supra* note 35, at 37.

37. See ASWATH DAMODARAN, INVESTMENT VALUATION: TOOLS AND TECHNIQUES FOR DETERMINING THE VALUE OF ANY ASSET 37 (3rd ed. 2012).

38. See *id.* (“[T]he value of debt is unaffected by changes in interest rates during the life of the loan or bond.”).

plus an appropriate amount of interest to compensate the creditor for the decreased value of the claim caused by the delayed payments.”<sup>39</sup> This is accomplished by determining the correct discount rate and then discounting the stream of payments to present value.<sup>40</sup> Market conditions affect the applicable discount rate. When a loan bears a market rate of interest (or when the market rate and the interest rate on the loan are the same), “the market will discount that [loan] at the same rate, and the present value of the [loan] will be its face amount.”<sup>41</sup> However, when a loan bears a below-market rate of interest, the market will discount it at a higher rate, resulting in a present value below its face amount. This is logical, as a rational market participant can achieve a higher rate of return on her investment by purchasing other loans bearing higher interest rates.<sup>42</sup>

To say that a loan has a “below-market interest rate” implies its interest rate is lower than the rate of interest on similar debt incurred by similarly situated borrowers. For instance, in *In Re Texas Grand Prairie Hotel Realty*, the rate of interest under the cramdown was below-market because it was lower than what other market participants would have charged for debt issued to a similarly situated debtor under similar terms. The court approved a 5% rate for a crammed down loan in reliance on the formula rate approach, consisting of a 1.75% risk adjustment over prime, when “the market was charging rates in excess of 5% on smaller, over-collateralized loans to comparable hotel owners” and “no willing lender would have extended credit on the terms it was forced to accept under the § 1129(b) cramdown plan . . . .”<sup>43</sup> The rate of interest under the cramdown was below-market because it was lower than what other market participants would have charged for debt issued to a similarly situated debtor under similar terms.

As noted above, the relationship between a crammed down interest rate and the market rate of interest can dramatically affect

---

39. *Rake v. Wade*, 508 U.S. 464, 472 n.8, (1993); see also LOPUCKI & WARREN, *supra* note 4, at 124 (“[T]he amount of money that must be paid at some later time to have a present value of \$X as of the effective date of the plan is \$X plus interest . . . .”).

40. *Rake*, 508 U.S. at 472 n.8; see also PEARCE, *supra* note 20, at 347.

41. 7 COLLIER, *supra* note 4, at ¶ 1129.03[2][b].

42. See STEPHEN A. ROSS, RANDOLPH W. WESTERFIELD & JEFFREY JAFFE, *CORPORATE FINANCE* 235-36 (9th ed. 2010). Conversely, where the market rate of interest is lower than the interest rate on the debt instrument, the security will be discounted at a lower rate, and its present value will exceed its face value.

43. *In re Tex. Grand Prairie Hotel Realty, LLC*, 710 F.3d 324, 336 (5th Cir. 2013).

the present value of a loan. The only way to provide creditors with the present value of their claims is to use prevailing market rates of interest. To the extent a below-market interest rate is assigned to a loan, it is virtually guaranteed that the present value of the loans will be below its face value.

*B. Judicial Methods for Determining Cramdown Interest Rates*

In conducting the present value analysis requisite for the approval of a cramdown under section 1129(b)(2), the bankruptcy court must determine an appropriate rate of interest, which when added to the debtor's deferred stream of payments is equivalent to the present value of the creditor's secured claim, determined as of the effective date.<sup>44</sup> Secured claims will not be deemed to have received fair and equitable treatment under a cramdown if deferred payments bear "insufficient interest to discount the payment to the allowed amount of the claim."<sup>45</sup> In such cases, the debtor's plan will not be confirmed. Unfortunately, the Bankruptcy Code does not provide a clear statutory directive on how bankruptcy courts should calculate appropriate cramdown interest rates.<sup>46</sup> Prior to *Till*,<sup>47</sup> the seminal case on judicial determination of cramdown interest rates in the Chapter 13 context, four approaches predominated, though the proper analysis of interest rates was contentious and subject to confusion.<sup>48</sup>

1. Formula Rate

The "formula rate" approach begins with a base rate and adjusts it upwards to accommodate the risk of non-payment by a specific debtor.<sup>49</sup> While the Supreme Court in *Till* used the national prime rate<sup>50</sup> as a base rate, other courts have relied instead on the

---

44. See Wong, *supra* note 12, at 1934; see also 7 COLLIER, *supra* note 4, at ¶ 1129.05[2][c].

45. 7 COLLIER, *supra* note 4, ¶ 1129.03[4][a][i][C]; see also Pearson, Jackson & Nohr, *supra* note 35, at 38 ("If the implied or proposed rate is less than the rate determined by the court to be fair and equitable, the plan cannot be confirmed.").

46. See *In re Am. Homepatient, Inc.*, 420 F.3d 559, 565 (6th Cir. 2005).

47. *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 474 (2004).

48. See Chaim J. Fortgang & Thomas Moers Mayer, *Valuation in Bankruptcy*, 32 UCLA L. REV. 1061, 1119 (1985) ("[F]ew bankruptcy issues have met with as much confusion as the determination of a proper discount rate.").

49. See *Till*, 541 U.S. at 479.

50. The prime rate generally refers to the interest rate used by commercial banks in lending to their most credit-worthy customers. See Ronald F. Greenspan & Cynthia Nelson, "UnTill" We Meet Again, *Why the Till Decision Might Not Be the Last Word on Cramdown Interest Rates*, AM. BANKR. INST. J., 48, 49 (Dec. 2004/ Jan. 2005)

treasury rate.<sup>51</sup> Regardless, the use of prime or any other easily and objectively ascertainable rate<sup>52</sup> was viewed in *Till* as facilitating rapid, expedient, and objective determination of cramdown interest rates.<sup>53</sup>

The size of the risk adjustment under the formula rate approach depends on “the circumstances of the estate, the nature of the security, and the duration and feasibility of the reorganization plan.”<sup>54</sup> Generally, the risk adjustment is between 1% and 3%.<sup>55</sup> In *Till*, the Court approved a risk premium over the prime rate of 1.5%,<sup>56</sup> resulting in a total interest rate of 9.5%.<sup>57</sup> Theoretically, no risk adjustment would be necessary if the court determined with certainty the debtor was not at risk of defaulting on its obligations.<sup>58</sup> However, such an outcome is highly, if not entirely, unlikely for corporate debtors<sup>59</sup> as only governments, with their ability to

---

(describing the national prime rate as “the base rate on corporate loans posted by at least 75 percent of the nation’s 30 largest banks.”).

51. See *Till*, 541 U.S. at 479 (using the “national *prime* rate”) (emphasis added); cf. *GMAC v. Valenti*, 105 F.3d 55, 64-65 (2d Cir. 1997) (remanding a Chapter 13 “case to the bankruptcy court for a recalculation of the interest rate based upon the *treasury rate* plus an additional risk premium.”) (emphasis added). The treasury rate refers to the return on investment, expressed as a percentage, on investments in U.S. government debt obligations.

52. See *In re Smith*, 178 B.R. 946, 952 (Bankr. D. Vt. 1995) (listing other base rates, including the “[f]ederal tax rate, federal or state legal rates, rates on consumer loans, Federal Land Bank rates” or an “average of several of these rates.”); see also Greenspan & Nelson, *supra* note 50, at 49 (describing other base rates, including LIBOR (London Interbank Offering Rate) and the Federal Home Loan Bank District Cost of Funds Index (COFI)). The use of these rates in any individual instance “depends on the type of loan, its duration (length) and lender circumstances. . . .” *Id.*

53. See *Till*, 541 U.S. at 479.

54. *Id.*

55. See *id.* at 480.

56. See *id.*

57. See *id.* at 471.

58. See *id.* at 479 n.18 (“[I]f the court could somehow be certain a debtor would complete his plan, the prime rate would be adequate to compensate any secured creditors forced to accept cramdown loans.”).

59. See Peter Schulman & C. Brad Peterson, *Understanding Discounting in Litigation*, COLO. LAW., Apr. 2007, at 23, 26 (stating while corporate securities cannot be considered risk free, certain “[i]nvestment-grade corporate bonds can yield rates of return that are only slightly higher than “risk-free” rates.”); see also Aswath Damodaran, *Estimating Risk Free Rates*, STERN SCH. OF BUS., <http://people.stern.nyu.edu/adamodar/pdfiles/papers/riskfree.pdf> (last visited Jan. 7, 2016) (stating the inability to rule out default risk means securities issued by private firms, regardless of their size or perceived safety, contain some degree of risk).

print money, may be considered risk free.<sup>60</sup> Recent history shows that even governments are not immunized against default.<sup>61</sup>

## 2. Forced Loan or Coerced Loan

Under the “forced loan” or “coerced loan” approach, a court sets the cramdown rate in order to bring the creditor to the level it would have obtained had it “foreclosed and reinvested the proceeds in loans of equivalent duration and risk.”<sup>62</sup> In other words, a creditor must be put in the same economic position it would have been had it liquidated the collateral securing its claim up front and made new equivalent loans to the reorganized debtor.<sup>63</sup> It assumes this rate is best determined by comparison to what a creditor would charge in the “open market for a loan of similar duration and risk as to the post-cramdown loan.”<sup>64</sup> Unlike the formula rate approach, the forced loan approach uses evidence of rates from actual credit markets.<sup>65</sup>

## 3. Cost of Funds

The “cost of funds” approach relies on the rate a creditor pays when it borrows for itself.<sup>66</sup> This approach makes two critical assumptions: first, that a creditor will cover any losses resulting from deferred payments by seeking new funds in the market, and second, that the creditor has access to additional funding.<sup>67</sup> Under this approach, the creditor must receive a rate equivalent to its cost of replacing the funds “locked up” in the replacement note issued

---

60. See Schulman & Peterson, *supra* note 59, at 26 (noting only U.S. treasuries are considered “risk-free” given the recognition by most investors of the extremely low likelihood of government default); see also Damodaran, *supra* note 59 (stating since governments control the printing of money, and therefore can always, at least theoretically, generate sufficient cash to meet their obligations, their securities are viewed as risk free).

61. For an example of a recent government default, see Nick Timiraos, Heather Gillers & Matt Wirz, *Puerto Rico’s Debt Crisis Deepens as Government Misses Payment*, WALL ST. J. (May 2, 2016), <http://www.wsj.com/articles/puerto-ricos-debt-crisis-turns-up-the-heat-on-congress-1462219483> [<https://perma.cc/WCE6-TLT5>] (describing the failure of the Puerto Rican government to honor a nearly \$400 million principal payment).

62. *Koopmans v. Farm Credit Servs. of Mid-Am, ACA*, 102 F.3d 874, 875 (7th Cir. 1996); see also *GMAC v. Valenti*, 105 F.3d 55, 63 (2d Cir. 1997).

63. See *Valenti*, 105 F.3d at 63 (quoting *Gen. Motors Acceptance Corp. v. Jones*, 999 F.2d 63, 69 (3d Cir. 1993)).

64. Wong, *supra* note 12, at 1936.

65. See *id.*

66. See *Valenti*, 105 F.3d at 63.

67. See Wong, *supra* note 12, at 1940.

under the cramdown plan,<sup>68</sup> or one that approximates the creditor's cost of funds.<sup>69</sup> Since the funds are no longer available to the creditor as a result of the cramdown plan, the creditor, theoretically, will be forced to obtain them elsewhere. Similarly to the formula rate approach, the cost of funds approach excludes creditor profits and administrative costs.<sup>70</sup> It posits the creditor is entitled only to the time value of money and should not be compensated for the other factors that ordinarily are included in an interest rate.<sup>71</sup>

Various criticisms have rendered the cost of funds approach the least widely accepted method for determining cramdown interest rates.<sup>72</sup> For instance, the cost of funds approach emphasizes the creditor's credit-worthiness over the debtor's.<sup>73</sup> It focuses on the former's credit history, business structure, capitalization, and operational efficiency, while ignoring the debtor's risk of default.<sup>74</sup> As a result, a poorly capitalized and inefficient creditor may receive a higher cramdown interest rate than a better capitalized and more efficient competitor with a lower cost of funds.<sup>75</sup> The variation among creditors and their abilities to borrow at specific rates would also require bankruptcy courts to engage in evidentiary proceedings to assess a given creditor's cost of funds on a case-by-case basis.<sup>76</sup>

The cost of funds approach also incorrectly assumes creditors have access to an unlimited supply of credit. Moreover, if a creditor

---

68. See *Gen. Motors*, 999 F.2d at 67 (“[W]e believe that it would be inappropriate to attempt to exclude consideration of ‘profit’ from a determination of the § 1325 interest rate. Moreover, we perceive no unfairness in including the profit component that anyone lending money in a commercial context expects.”).

69. See Todd J. Zywicki, *Cramdown and the Code: Calculating Cramdown Interest Rates Under the Bankruptcy Code*, 19 T. MARSHALL L. REV. 241, 254 (1994) (quoting 5 COLLIER ON BANKRUPTCY ¶ 1325.03 at 1325-46 (Lawrence P. King ed., 15th ed. 1993)).

70. See Michael Elson, *Say “Ahhh!”: A New Approach for Determining the Cram Down Interest Rate After Till v. SCS Credit*, 27 CARDOZO L. REV. 1921, 1934 (2006); see also *In re Ivey*, 147 B.R. 109, 117 (M.D.N.C. 1992) (describing an advantage of the cost of funds approach as the fact that “it does not include profit to the creditor.”).

71. See Zywicki, *supra* note 69, at 254. For further discussion on the components of interest rates, see section I.B, *infra*.

72. See Ry Ellison, *The Application of Till to Chapter 11 Cases in the Fifth Circuit*, 1 BUS. & BANKR. L.J. 28, 32 (2014) (claiming “the cost of funds approach is perhaps the least accepted of the various cramdown rate calculation methods.”).

73. *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 478 (2004).

74. See David G. Epstein, *Don't Go and Do Something Rash About Cram Down Interest Rates*, 49 ALA. L. REV. 435, 450-51 (1998).

75. See *id.* at 451.

76. See *GMAC v. Valenti*, 105 F.3d 55, 64 (2d Cir. 1997).

has limited credit to draw upon and a maximum amount of profit to generate through use of that credit, the creditor will suffer a loss as a result of borrowing merely to replace capital locked up in a replacement loan. That loss is equal to the interest rate the lender could enjoy on a loan to borrowers similar to the debtor minus transaction costs and cost of funds. The cost of funds approach fails to compensate the creditor for this loss by assessing a cramdown rate based merely upon the replacement of one pool of funds with another.<sup>77</sup>

#### 4. Contract Rate

Finally, the “contract rate” approach establishes a rebuttable presumption that the correct cramdown rate is the original interest rate negotiated in the pre-bankruptcy agreement between the creditor and debtor,<sup>78</sup> either of whom may provide evidence to challenge the presumptive cramdown rate in favor of one that is higher or lower.<sup>79</sup> For instance, evidence of changes in market conditions or the debtor’s risk profile as a result of a successful restructuring may be sufficient to convince the court to adjust the presumptive contract rate.<sup>80</sup>

The contract rate approach is often rightly criticized in that its starting point is the “stale” rate under the pre-petition loan. The factors that led the parties to agree to that rate may be completely different at the time of the confirmation of the plan. For example, there may be less debt on the debtor’s balance sheet, litigation claims against the debtor may have been eliminated, and the debtor’s collateral may be more or less valuable. In addition, market factors may have changed, especially if there is a long delay between the contract date and when the debtor emerges from Chapter 11.<sup>81</sup> Beginning with an “outdated” rate and introducing

---

77. See *United Carolina Bank v. Hall*, 993 F.2d 1126, 1130 (4th Cir. 1993) (arguing the cost of funds approach is flawed because “[w]hen it is recognized that every secured creditor has a limited amount of credit on which to draw, then it follows that utilizing some of that borrowing capacity without providing the secured creditor with the usual return on its capital produces a loss for the secured creditor.”).

78. See Wong, *supra* note 12, at 1938.

79. See *Till v. S.C.S. Credit Corp.* 541 U.S. 465, 473 (2004).

80. See Wong, *supra* note 12, at 1938.

81. See Aaron J. Bell, *Making Cramdown Palatable: Post-Confirmation Interest on Secured Claims in a Chapter 11 Cramdown*, 23 WILLAMETTE L. REV. 405, 419-20 (1987) (criticizing the failure of a court to take into account the twenty months between the date the contract was entered into and the date of confirmation, since “[i]t is doubtful that the twenty-month-old rate was an accurate indication of the market

expert testimony as to why it should be adjusted seems arbitrary. Additionally, when the post-petition creditor is a healthier and more credit-worthy company, the contract rate initially overcompensates the creditor and then obligates the debtor to provide evidence to force the interest rate to a more appropriate and lower level.

C. *The Till Decision and Adoption of the Formula Rate Approach*

In *Till*, a plurality of the Supreme Court rejected the forced loan, cost of funds, and contract rate approaches in the Chapter 13 bankruptcy context. Instead, the Court held the formula rate approach “best comports with the purposes of the Bankruptcy Code” and should therefore serve as the proper standard for determining cramdown interest rates.<sup>82</sup> In adopting the formula rate approach, the plurality reasoned the three discarded tests were too complicated, created expensive evidentiary proceedings, and prioritized making the creditor whole over ensuring an appropriate present value for the debtor’s future payments.<sup>83</sup>

By comparison, the formula rate approach took “its cue from ordinary lending practices.”<sup>84</sup> It used as a reference point the national prime rate, “which reflects the financial market’s estimate of the amount a commercial bank should charge a creditworthy commercial borrower.”<sup>85</sup> The *Till* Court believed bankruptcy courts would be able to easily determine the risk adjustment component of the formula rate approach, since much of the relevant evidence would already be contained in the debtor’s bankruptcy filings.<sup>86</sup> Finally, by starting with a low interest rate and adjusting upward, the plurality reasoned that the creditor, rather than the debtor, would suffer the brunt of any evidentiary burden. The plurality preferred this outcome because the creditor presumably had better access to any information absent from the debtor’s filings, such as evidence pertaining to the “liquidity of the collateral market.”<sup>87</sup>

---

at the time the plan became effective.”); *see also In re Mitchell*, 39 B.R. 696, 702 (Bankr. D. Or. 1984) (“It is obvious that a contract rate of interest is not an adequate indicator of current market conditions over an extended period of time.”).

82. *Till*, 541 U.S. at 480.

83. *Id.* at 477.

84. *Id.* at 478.

85. *Id.* at 479.

86. *See id.*

87. *Id.*

D. *Till's Famous Footnote 14 and the Two-Part Approach*

Despite *Till's* adoption of the formula rate approach, its footnote 14 suggests the approach for determining cramdown interest rates best suited for Chapter 13 may not be equally suited for Chapter 11.<sup>88</sup> The footnote distinguishes the limited market for lending to Chapter 13 debtors from the market for lending to Chapter 11 debtors.<sup>89</sup> In the latter, “numerous lenders advertise financing for chapter 11 debtors in possession.”<sup>90</sup> As a result, the footnote suggests, “when picking a cramdown rate in a chapter 11 case, it might make sense to ask what rate an efficient market would produce.”<sup>91</sup>

Courts relying on this footnote have adopted a two-step approach for determining an interest rate in Chapter 11 cramdowns. In the first instance, “the market rate should be applied in Chapter 11 cases where there exists an efficient market.”<sup>92</sup> However, in the absence of any efficient market for a Chapter 11 debtor, “the Bankruptcy Court should employ the formula approach endorsed by the *Till* plurality.”<sup>93</sup> Courts that have adopted this rule have found efficient markets exist where they offer loans “with a term, size, and collateral comparable to the forced loan contemplated under the cramdown plan.”<sup>94</sup> This analysis has allowed courts to continue relying on market rates in Chapter 11, where markets are efficient and supported by sufficient evidence. In Chapter 13, by contrast, the absence of an adequate or efficient market for lending to debtors implies there is no data on which to base a market-referenced

---

88. *See id.* at 476 n.14.

89. *See id.* (stating in Chapter 13, “there is no free market of willing cramdown lenders” but that this is not true in Chapter 11).

90. *See id.*; *see also* Griffiths, *supra* note 14.

91. *Till*, 541 U.S. at 476 n.14.

92. *In re Am. Homepatient, Inc.*, 420 F.3d 559, 568 (6th Cir. 2005); *see also In re 20 Bayard Views, LLC*, 445 B.R. 83, 111 (Bankr. E.D.N.Y. 2011); *Mercury Capital Corp. v. Milford Conn. Assocs., L.P.*, 354 B.R. 1, 12 (D. Conn. 2006) (on remand, the bankruptcy court was required to consider whether “an efficient market rate [existed] for the type of loan Mercury [was] forced to give the debtor under the competing plans.”); *In re Brice Rd. Developments, LLC*, 392 B.R. 274, 280 (B.A.P. 6th Cir. 2008) (“[I]n this circuit, in a chapter 11 case where an ‘efficient market’ exists, the market rate should be applied, and where no ‘efficient market’ exists, the formula approach endorsed by the Supreme Court . . . should be employed.”).

93. *Homepatient*, 420 F.3d at 568; *see also In re Bayard Views*, 445 B.R. at 111; *Mercury Capital* 354 B.R. at 13; *In re Brice*, 392 B.R. at 280 (“[I]n this circuit, in a chapter 11 case where an ‘efficient market’ exists, the market rate should be applied, and where no ‘efficient market’ exists, the formula approach endorsed by the Supreme Court . . . should be employed.”).

94. *In re Tex. Grand Prairie Hotel Realty, LLC*, 710 F.3d 324, 336 (5th Cir. 2013).

cramdown interest rate, such that the formula rate approach is a more desirable, if not the only option.

## II.

### MOMENTIVE PERFORMANCE MATERIALS AND THE EXTENSION OF THE FORMULA RATE APPROACH INTO CHAPTER 11

#### A. Momentive *Facts and Arguments*

In *Momentive*, Judge Drain extended the Supreme Court plurality's endorsement of the formula rate approach in Chapter 13 to a Chapter 11 case. The debtor, a silicon and quartz producer, filed for Chapter 11 protection in April 2014, at least in part as a result of a \$3.8 billion take private transaction sponsored by Apollo Global Management.<sup>95</sup> At issue were two classes of senior secured notes executed in 2012: \$1.1 billion in First Lien Notes and \$250 million in 1.5 Lien Notes (collectively the "Senior Secured Notes").<sup>96</sup> Both the First and 1.5 Lien notes were due to mature in October 15, 2020, and bore interest rates of 8.875% and 10%, respectively.<sup>97</sup>

The debtor's plan included a "deathtrap" provision.<sup>98</sup> If the Senior Secured Notes voted in favor of the plan, they would receive all outstanding principal and accrued interest in cash on the effective date. However, no make-whole premium<sup>99</sup> would be permitted under this option.<sup>100</sup> Alternatively, if the Senior Secured Noteholders voted against the plan, the debtor would instead provide them with replacement notes, which matured in seven, and seven and a half years depending on lien priority, with interest rates to be deter-

---

95. See Patrick Fitzgerald, *Apollo's Momentive Performance Materials Makes Chapter 11 Bankruptcy Filing*, WALL ST. J. (Apr. 13, 2014), <https://www.wsj.com/articles/apollos-momentive-performance-materials-makes-chapter-11-bankruptcy-filing-1397439652> [<https://perma.cc/5SEX-Q7L3>].

96. See *In re MPM Silicones, LLC*, 531 B.R. 321, 325 (Bankr. S.D.N.Y. 2015); David Griffiths, *Momentous Decision in Momentive Performance Materials: Cramdown of Secured Creditors – Part II*, WEIL BANKRUPTCY BLOG (Sept. 10, 2014), <https://business-finance-restructuring.weil.com/chapter-11-plans/momentous-decision-in-momentive-performance-materials-cramdown-of-secured-creditors-part-ii/>.

97. See *In re MPM Silicones, LLC*, 531 B.R. at 325.

98. See Griffiths, *supra* note 96.

99. A make-whole premium is a provision in a debt instrument allowing the issuer to retire or pay off the remaining debt early. The borrower must then pay the holder a lump sum based on the present value of the future payments not made as a result of the early retirement. See *Make-Whole Call Provisions*, CAPITAL ADVISORS GROUP (Oct. 31, 2004), <https://capitaladvisors.com/wp-content/uploads/2016/12/Make-Whole-Call-Provisions.pdf>.

100. See *In re MPM Silicones, LLC*, 531 B.R. at 326.

mined under the cramdown provisions in section 1129(b)(2) and by using the approach endorsed by the Supreme Court in *Till*.<sup>101</sup> In effect, the deathtrap offered the Senior Secured Noteholders one of two options: run the risk of receiving a below-market interest rate on the replacement notes, or forgo the make-whole premium and their right to receive the value of any uncollected future interest payments.

When the Senior Secured Notes voted against the plan,<sup>102</sup> the debtor attempted to cram it down over their objections. In response, the Senior Secured Notes argued the interest rates on the replacement notes violated section 1129(b)(2)'s fair and equitable requirement. They interpreted the *Till* plurality opinion to require the use of the formula rate approach only after a determination by the court that an efficient market did not already exist<sup>103</sup> or that a market rate of interest should be used to calculate a cramdown interest rate, when available. The Senior Secured Noteholders claimed such a market was in fact present. They pointed to the fact that the *Momentive* debtor had secured, and the bankruptcy court had approved, exit and bridge financing.<sup>104</sup> *Momentive* was the only case, post *Till*, in which the debtors had secured exit financing commitments prior to attempting to cramdown its secured creditors.<sup>105</sup> The Senior Secured Notes also provided evidence of a "robust and readily available" leveraged loan market.<sup>106</sup> The First Lien

---

101. *See id.*

102. *See In re MPM Silicones, LLC*, 531 B.R. at 326.

103. *See* Cramdown Objection of Wilmington Trust, National Association, as Indenture Trustee, to Confirmation of Debtors' Proposed Joint Chapter 11 Plan of Reorganization at 39, *In re MPM Silicones, LLC*, 2014 WL 4436335 (Bankr. S.D.N.Y. Sept. 9, 2014) (No. 7:14-BK-22503); Objection of Bokf, Na, as First Lien Trustee, to the Debtors' Joint Chapter 11 Plan and Confirmation of the Plan with Respect to the Terms of the Replacement First Lien Notes at 38, *In re MPM Silicones, LLC*, 2014 WL 4255115 (Bankr. S.D.N.Y. August 12, 2014) (No. 7:14-BK-22503). Certain courts have adopted this approach, or a near variant. *See, e.g., In re Prussia Assocs.*, 322 B.R. 572, 589 (Bankr. E.D. Pa. 2005) (arguing a caveat exists under *Till* footnote 14, such that the formula rate approach should not be applied where an efficient market exists). While the *Prussia Associates* court found a sufficiently efficient market existed, the totality of evidence did not permit it to conclude with certainty the appropriate market rate of interest, and as a result, it fell back on the formula rate approach. *Id.* at 590.

104. *See* Objection of Bokf, Na, *supra* note 103, at 30.

105. *See id.* at 39.

106. *Id.* ("[T]he current leveraged loan market is strong with over \$459.6 billion in leveraged loans and \$196.7 billion in high yield debt issued year-to-date (representing a total of \$656.3 billion in issuances) and provides an indication of a robust and readily available financing market."); *see also In re MPM Silicones, LLC*, 531 B.R. at 332; *In re MPM Silicones, LLC*, 2014 WL 4436335, at \*24.

Noteholders contended the court should therefore look to the “rate, tenor and terms” of the exit financing,<sup>107</sup> which served as an adequate market check and risk assessment of the debtors and their business, and no further.<sup>108</sup> In addition, the exit financing was to be secured by the same collateral and receive the same priority as the replacement First Lien Notes, yet under the formula rate approach, the First Lien Notes would be priced lower than the exit financing. Perversely, the 1.5 Lien Noteholders would be required to lend to the debtors “at an interest rate lower and on less favorable terms than the exit lenders under the Exit Facilities—despite the fact that the exit lenders will enjoy a superior (i.e., less risky) security position . . . .”<sup>109</sup>

Even if forced to concede the use of the formula rate approach, the Senior Secured Noteholders argued the bankruptcy court erred in its application of the method through its use of the then-current seven-year national treasury rate as a base rate rather than the national prime rate.<sup>110</sup> Since the debtors had relied on the risk-free treasury rate, the First Lien Trustees claimed the risk adjustment component of the formula rate approach was undervalued and should be adjusted higher.<sup>111</sup> In the words of the First Lien Trustees: “the incorrect use of the Treasury Rate as the base rate without any corresponding increase in the risk premium would effectively deprive the First Lien Noteholders of more than an additional full percentage point that they are entitled to.”<sup>112</sup>

Despite these arguments, Judge Drain agreed with the debtor’s application of the formula rate approach based on the national treasury rate. He approved a 4.1% interest rate for the First Lien Noteholders, and a 4.85% interest rate for the 1.5 Lien Noteholders, forcing the secured creditors to accept below-market interest rates on their securities.<sup>113</sup> Judge Drain only added 0.5% to the cramdown interest rates to reflect the use of the treasury rate,<sup>114</sup> in

---

107. See *Objection of Bokf, Na*, *supra* note 103, at 31.

108. See *id.* at 74-75.

109. Cramdown *Objection of Wilmington Trust*, *supra* note 103, at 33.

110. See *Objection of Bokf, Na*, *supra* note 103, at 52 (arguing the *Till* plurality “was clear that in utilizing the formula approach, the starting point is a base rate equal to the prime interest rate.”).

111. See *id.* at 54.

112. *Id.*

113. *In re MPM Silicones, LLC*, No. 14-22503-RDD, 2014 WL 4436335, at \*32 (Bankr. S.D.N.Y. Sept. 9, 2014).

114. At this juncture, it is important to address briefly the confusion between a cramdown interest rate based on prime and one based on treasury. Courts have incorrectly used these rates interchangeably, but their meaning and significance in

spite of the fact that at the time of the holding, the prime rate of 3.25% was more than 1% higher than the then-applicable treasury rate of 2.14%.<sup>115</sup>

*B. Judge Drain's Reasoning in Momentive and Adoption of the Formula Rate Approach*

In resolving the bankruptcy dispute, Judge Drain followed the Supreme Court's opinion in *Till*, as well as *Valenti*, a Second Circuit opinion.<sup>116</sup> He relied on starkly pro-debtor policy factors in addition to the argument that *Till* was binding in both Chapter 13 and Chapter 11. First, Judge Drain noted the purpose of a cramdown interest rate is “not to put the creditor in the same position that it would have been in had it arranged a ‘new’ loan”<sup>117</sup> Rather, the objective is to restore the creditor to “the same economic position it would have been in had it received the value of its allowed claim immediately.”<sup>118</sup> This distinction is significant and highlights Judge Drain's view that present value must be determined as to the value of the creditor's *claim*, which does not include any degree of profit,<sup>119</sup> not as to the value of the creditor's *loan*, which does.

Having rationalized the blanket exclusion of profits from cramdown interest rates, Judge Drain claimed market factors (such as those relied upon in the forced loan approach) are not applicable in “court-administered and court-supervised cramdown loans.”<sup>120</sup> Moving further, Judge Drain found market-based evidence, if relevant at all, was only applicable in setting a proper risk premium.<sup>121</sup> In instances where a bankruptcy court determined the

---

the cramdown context differ. In *Till*, for instance, the plurality used prime plus a credit risk adjustment in the range of 1 to 3%. However, the prime rate contemplates borrower default and incorporates compensation for credit risk (and therefore cannot be said to be truly risk free). Thus, in *Till*, the overall risk adjustment based on real interest rates appears to *exceed* the risk adjustment contemplated by the reasoning of the court. *Momentive*, by comparison, used a risk-free treasury rate, which matched the duration of the loan in question. Though Judge Drain suggested that in some instances, no risk adjustment would be necessary, in *Momentive*, he added .5% for the first lien replacement notes, and .75% percent for the 1.5 lien replacement notes.

115. See Objection of Bokf, Na, *supra* note 103, at 54.

116. See *id.* at 24; *GMAC v. Valenti*, 105 F.3d 55 (2d Cir. 1997).

117. *In re MPM Silicones, LLC*, 2014 WL 4436335, at \*25 (emphasis in original) (*quoting Valenti*, 105 F.3d at 63-4).

118. *Id.*

119. See *id.* For further discussion on the inclusion of profit in cramdown interest rate determinations, see Section I.A, *infra*.

120. *Id.* at \*26 (*quoting Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 477 (2004)).

121. *Id.*

debtor would repay its obligations without risk, no premium on the replacement notes would be necessary.<sup>122</sup>

Unlike in the ordinary credit markets, *Momentive* also stated creditors who issue loans in the cramdown context, albeit under judicial coercion, need not be indifferent between immediate foreclosure and a discounted and deferred stream of payments.<sup>123</sup> Judge Drain accepted *Till's* argument that the very concept of a cramdown precludes creditor indifference, since presumably any creditor would prefer immediate foreclosure to the alternative.<sup>124</sup> In determining a fair and equitable cramdown interest rate on the replacement notes, the debtor was not obligated to provide the creditor a rate adjusted to reflect the lost earnings resulting from its inability to foreclose immediately and re-lend any proceeds in the market.<sup>125</sup>

Finally, *Momentive* summarily rejected arguments (included those forwarded by the Senior Secured Noteholders) stemming from *Till* footnote 14.<sup>126</sup> To Judge Drain, there was “no meaningful difference” between Chapter 11 and Chapter 13, and no justification for restricting the *Till* formula rate approach to the latter.<sup>127</sup> Judge Drain strongly discounted the footnote’s persuasive value, since it equated loans imposed on objecting lenders subject to cramdowns with debtor-in-possession financing.<sup>128</sup> He argued this comparison was incorrect, in part, because debtor-in-possession lenders want to make loans, whereas lenders subject to a cramdown do not.<sup>129</sup> Moreover, debtor-in-possession financing takes place at the beginning of Chapter 11 proceedings, while cramdown rates are determined at the end, when the debtor is “more stable and restructured.”<sup>130</sup> In addition, Judge Drain pointed out that the auto

---

122. *See id.*

123. *See id.*

124. *Id.* at \*25; *see also*. *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 476-77 (2004).

125. *In re MPM Silicones, LLC*, 2014 WL 4436335, at \*25.

126. *Id.* at \*28.

127. *Id.* at \*27.

128. *Id.* Indeed, footnote 14 appears to incorrectly include debtor-in-possession lenders within the “free market of willing cramdown lenders.” *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 476 n.14 (2004). However, it is not clear whether the Supreme Court viewed debtor-in-possession lenders as the *exclusive* evidence of such a market or whether it intentionally excluded post-petition or exit-financing. An equally plausible and more purposive interpretation is that where any such market exists, evidence of debtor-in-possession lenders may be used to determine cramdown interest rates in the Chapter 11 context and the name of the market or type of credit instrument is less relevant.

129. *Id.* at \*25.

130. *Id.*

loan market at issue in *Till* contained a greater number of active lenders and borrowers, all with access to substantial publicly available data, compared to that which existed for lenders and borrowers with regard to debtor-in-possession or exit financing in Chapter 11.<sup>131</sup> Judge Drain found it significant that the *Till* Court did not view the auto loan market as sufficiently competitive to support an approach other than the formula rate approach. Given the fact that only three exit lenders were available to lend to the *Momentive* debtors, Judge Drain claimed “footnote 14 is a very slim reed indeed on which to require a market-based approach. . . .”<sup>132</sup>

C. *The Second Circuit’s Recent Opinion Overturned Momentive and Implemented a Two-Step Approach*

Judge Drain’s holding in *Momentive*, including the generalized exclusion of so-called “profit” from crammed down interest rates, caused a grave injustice to the Senior Secured Noteholders. Fortunately, this portion of his opinion was recently overturned in the Second Circuit, which adopted instead a “two-step approach” relying primarily on market rates of interest. The Second Circuit refused to read the *Till* plurality as ignoring the relevance of efficient markets in determining Chapter 11 cramdown interest rates and stated that to do so would be a “major departure from long-standing precedent” which viewed value as best determined by analysis of the market.<sup>133</sup> It suggested courts should be limited in their ability to make valuation judgments independent of competitive choice, particularly when market valuations are available.<sup>134</sup> Pointing to the expert testimony provided by the Senior Secured Noteholders, the Second Circuit determined that a market rate for the cramdown debt did exist. As discussed, under the debtor’s deathtrap offer, the Senior Secured Noteholders could have accepted immediate payment for their notes. To fund this lump-sum cash-out payment, which was ultimately never required, the debtor sought market financing and was quoted interest rates between 5 and over 6 percent. Thus, the Second Circuit stated:

[W]here, as here, an efficient market may exist that generates an interest rate that is apparently acceptable to sophisticated parties dealing at arm’s-length, we conclude, consistent

---

131. *Id.*

132. *Id.*

133. *In re MPM Silicones, LLC*, 874 F.3d 787, 800 (2d Cir. 2017).

134. *See id.* (quoting *Bank of Am. Nat’l Trust and Sav. Ass’n v. 203 N. LaSalle St. P’Ship*, 526 U.S. 434, 457-58 (1999)).

with footnote 14, that such a rate is preferable to a formula improvised by the court.<sup>135</sup>

Ultimately, the Second Circuit refused to categorically reject the probative value of market rates of interests in the cramdown interest rate context and remanded the case to the bankruptcy court for re-determination of the appropriate rate.

### III. THE BUSINESS OF BANKING

In light of the foregoing, I will next discuss why the Second Circuit's two-step approach should be adopted as the standard for determining cramdown interest rates under Chapter 11 of the Bankruptcy Code, taking into account the operation of commercial credit markets as well as how such markets determine prevailing interest rates. Commercial credit takes many forms, depending on, among other factors, the debtor's credit quality,<sup>136</sup> size,<sup>137</sup> and leverage;<sup>138</sup> whether the debtor has proprietary information;<sup>139</sup> and prevailing market conditions.<sup>140</sup> Those seeking external financing may borrow from banks and other financial institutions or the capital markets. Often, borrowers prefer to issue debt securities, such as bonds and notes, and view banks as lenders of last resort. Banks impose stricter requirements than bonds on borrowers, which may

---

135. *Id.* at 801. The Second Circuit also distinguished Chapter 11 from the a "sub-prime loan in the Chapter 13 context" where "'value' can be elusive because the market is not necessarily efficient and the borrower is typically unsophisticated." *Id.*

136. See David D. Jenis & Vassil T. Mihov, *The Choice Among Bank Debt, Non-Bank Private Debt, and Public Debt: Evidence from New Corporate Borrowings*, 70 J. FIN. ECON. 3, 5 (2003) (finding "the primary determinant of the choice of debt instrument is the credit quality of the issuing firm.").

137. See Joel Houston & Christopher James, *Bank Information Monopolies and the Mix of Private and Public Debt Claims*, 51 J. FIN. ECON. 1863, 1864 (1996) (finding reliance on bank financing decreases with firm size and leverage); Shane A. Johnson, *An Empirical Analysis of the Determinants of Corporate Debt Ownership Structure*, 32 J. OF FIN. AND QUANT. ANAL. 47, 48 (1997) (stating "bank debt is concentrated among smaller firms.").

138. See Houston & James, *supra* note 137, at 1864.

139. See generally Jayant R. Kale & Costanza Meneghetti, *The Choice Between Public and Private Debt: A Survey*, 23 IIMG MGMT. REV. 5, 6-7 (2011) (reviewing the impact of proprietary information on the decisions of borrowers to obtain financing from public or private sources).

140. For instance, see Mark T. Leary, *Bank Loan Supply, Lender Choice, and Corporate Capital Structure*, 64 J. FIN. 1143, 1180 (2009) (finding expansions and contractions in the supply of bank loans cause variations in firms' debt placement structures).

take the form of more restrictive covenants or more covenants in general.<sup>141</sup> I focus below on the types of debt that may be negatively impacted by below-market cramdown interest rates.

#### A. *Bank Debt*

Loans are rarely executed on a bilateral or individual basis. Instead, loans are more frequently syndicated, or shared among a group of lenders.<sup>142</sup> Syndication has the benefit of spreading the funding obligations and associated risks among a group of institutions such that no entity bears the burden exclusively.<sup>143</sup> There exists a market for origination of syndicated loans among banks, finance companies, and institutional investors.<sup>144</sup> When a company wishes to take on debt, it retains an arranger, often an investment or commercial bank, to build out a “book” of potential creditors.<sup>145</sup> Creditors bid by offering to fund a percentage of the requested principal at a given interest rate.<sup>146</sup> Many credit structures contain market flex provisions, which allow the arranger to modify the terms of the loan within a certain set of parameters to make them more attractive to the credit market in the event the debt is under-subscribed.<sup>147</sup> The terms and rates on the debt are therefore the product of a competitive, market-driven bidding process.

---

141. See William W. Bratton, *Bond Covenants and Creditor Protection: Economics and Law, Theory and Practice, Substance and Process*, 7 EUR. BUS. ORG. L. REV. 39, 20 (2006) (noting “substantial protection,” as measured by covenant incidence and stringency, is the rule in credit agreements between borrowers and banks, but that when one looks to bond markets, there is a “reduction in the incidence and intensity of covenants, even for borrowers of the same risk class as the bank borrowers.”); Arthur E. Wilmarth, Jr., *The Transformation of the U.S. Financial Services Industry, 1975-2000: Competition, Consolidation, and Increased Risks*, 2002 U. ILL. L. REV. 215, 227-231 (discussing a continuum of borrowers, where borrowers with the strongest reputations and financial profiles borrowers are able to tap into the capital markets, but where more “opaque” borrowers are forced to lend from banks and incur more “restrictive covenants and collateral requirements”); R. STAFFORD JOHNSON, *DEBT MARKETS AND ANALYSIS* 273 (1st ed. 2013) (stating “[l]arger corporations, whose credit standings are often strong, prefer to finance their long-term and intermediate-term assets by selling corporate bonds and notes . . .”).

142. See STEVEN MILLER & WILLIAM CHEW, *A GUIDE TO THE LOAN MARKET* 7, STANDARD & POOR’S (Sept. 2011) [<https://perma.cc/S5ZK-9PTF>].

143. See *id.* at 8.

144. See *id.* at 10. Loan mutual funds may also provide retail investors with exposure to the syndicated loan market. *Id.* Thus, below-market cramdown interest rates have the potential to impact not only large banks and financial institutions, but individual retail investors saving for retirement or other investments.

145. See *id.* at 8.

146. See *id.*

147. See *id.*

A credit agreement is the arrangement between a syndicate of financial institutions and a borrower to provide a certain sum of money. Credit agreements can provide financing in many ways.<sup>148</sup> For long-term or intermediate-term financing needs, borrowers request a term loan.<sup>149</sup> A term loan is the extension of credit for a fixed period. The borrower must repay the principal sum advanced, either periodically or in a lump sum at the end of the term (often referred to as a balloon or bullet payment) as well as interest at pre-determined intervals, based on either a fixed or variable rate.<sup>150</sup> Individual borrowers may have both a Term Loan A and Term Loan B. Term Loan As have progressive repayment schedules and are normally syndicated to banks. Term Loan Bs are carved out for syndication to nonbank institutional investors.<sup>151</sup>

Another credit structure is a revolving line of credit, sometimes referred to as a revolver or credit line. Revolvers act like corporate credit cards and allow a borrower to borrow, repay, and borrow again up to the amount committed by its creditors.<sup>152</sup> Credit lines are commonly used for short-term financing purposes<sup>153</sup> and may be tied to a borrowing-base. In such instances, the amount of money the creditor may borrow fluctuates with its current assets, such as inventory or cash and cash receivables.<sup>154</sup>

Such forms of credit are sold and traded among investors on a robust secondary market.<sup>155</sup> Many creditors originate loans with the intent of selling them on a secondary market immediately or shortly

---

148. See JONATHAN GOLIN & PHILIPPE DELHAISE, *THE BANK CREDIT ANALYSIS HANDBOOK* 97 (2d Ed. 2013).

149. See JOHNSON, *supra* note 141, at 23.

150. GOLIN & DELHAISE, *supra* note 148, at 99; see also STRAHAN, *supra* note 30, at 5 (describing a term loan as a facility where “[a] borrower receives a lump sum at the beginning of the contract and pays off the loan, plus interest, over time.”).

151. See MILLER & CHEW, *supra* note 142, at 16.

152. See *id.* at 16; GOLIN & DELHAISE, *supra* note 148, at 101-02.

153. See JOHNSON, *supra* note 141, at 23.

154. See MILLER & CHEW, *supra* note 142, at 16.

155. See *id.* at 21 (“Secondary sales occur after the loan is closed and allocated, when investors are free to trade the paper.”); Blaise Gadanecz, *The Syndicated Loan Market: Structure, Development and Implications*, BIS Q. REV. at 83-84, <http://ssrn.com/abstract=1967463> (stating “[s]yndicated credits are increasingly traded on secondary markets” and the “biggest secondary market for loan trading is the United States, where the volume of such trading amounted to \$145 billion in 2003.”); Robert M. Bushman & Regina Wittenberg-Moerman, *Does Secondary Loan Market Trading Destroy Lenders’ Incentives?*, Chicago Booth Research Paper No. 09-45 at 9, <https://perma.cc/3UKV-LA9S> (stating the size of the market for secondary loans was \$510 billion in 2008.).

thereafter.<sup>156</sup> Below-market interest rates can affect the value of loans trading on secondary markets in two ways. Firstly, the originators of the debt or investors in secondary markets who choose to sell the loans may be forced to do so at a discount, thereby taking a loss. In the alternative, creditors who hold the debt to maturity will utilize their capital in a sub-optimal manner, which may in turn affect their ability to bid competitively on other credit facilities or to meet their capital and funding requirements.

### B. Capital Markets Debt

As an alternative to borrowing from a bank or other institutional investors, companies may borrow from the capital markets. In the first instance, companies issue debt securities in the primary markets.<sup>157</sup> The process for issuing bonds is extremely complex from a regulatory perspective and is beyond the scope of this Note.<sup>158</sup> Of greater significance to a discussion of capital markets debt and cramdowns are secondary markets, or markets “for the buying and selling of existing assets and financial claims.”<sup>159</sup> The United States debt capital market is extremely large; the total value of corporate debt securities outstanding is in excess of \$8 trillion.<sup>160</sup>

The types of debt issued by companies and traded by investors vary widely by maturity, interest payment, covenant packages, and principal, and include convertibles, callable or redeemable securities, and credit sensitive securities.<sup>161</sup> While the exact number of years may vary, bonds are securities with longer maturities, whereas notes, such as those at issue in *Momentive*, typically have shorter maturities than bonds.<sup>162</sup> Entities that purchase securities become

---

156. See Bushman & Wittenberg-Moerman, *supra* note 155, at 3 (stating “at loan origination, lenders anticipate that a given loan will ultimately be sold in the secondary market.”).

157. See JOHNSON, *supra* note 141, at 25 (“[T]he primary market . . . is the market in which new securities are sold for the first time.”).

158. For a more detailed explanation of the regulatory framework and process for issuing debt securities, see RECHTSCHAFFEN, *supra* note 33, at 47-48.

159. JOHNSON, *supra* note 141, at 25. The volume of trading on secondary markets far exceeds the volume of trading on primary markets. *Id.*

160. See Jake Liebschutz & Brian Smith, *Examining Corporate Bond Liquidity and Market Structure*, U.S. DEP’T OF THE TREASURY (May 7, 2016) [<https://perma.cc/MX4T-DFQQ>].

161. See generally JOHNSON, *supra* note 141, at 274-291.

162. See *id.* at 273 (stating notes mature in under five years, whereas bonds mature in over five years); RECHTSCHAFFEN, *supra* note 33, at 51 (stating notes mature in under 10 years, whereas bonds mature in over 10 years).

creditors to the issuer<sup>163</sup> and include pensions, insurance companies, banks, other financial institutions, and households.<sup>164</sup> In exchange for lending to a company by purchasing its securities, creditors receive interest payments and the principal amount of their investment when the security matures.<sup>165</sup> Much like loans, creditors who receive below-market cramdown interest rates will take losses as their securities trade at discounts.

### C. *Interest Rates in Commercial Credit Markets*

In commercial credit markets, the terms of debt, including the interest rate, are the product of arms-length negotiations between willing borrowers and sellers. Such terms are specific to a debtor's financial condition, including the quality of the collateral (if any) securing the debt and also reflect the lender's desire to transact with that borrower.<sup>166</sup> The financial markets play an important role in this process by facilitating the movement of savings, or surplus capital, from those who possess it to those who do not.<sup>167</sup> Holding risk constant, during periods of easy money or when the supply of capital is high, demand for the savings held by any individual entity will be low, given its availability elsewhere. Conversely, if the money supply is low, borrowers must pay a higher price to access excess capital held by others. Thus, market rates of interest also depend on the interaction between supply and demand.<sup>168</sup> As a result, mar-

---

163. See James Woepking, *Private Capital and Development: Challenges Facing International Financial Institutions in a Globalized Economy*, 9 *TRANSNAT'L L. & CONTEMP. PROBS.* 233, 235 (1999); see also RECHTSCHAFFEN, *supra* note 33, at 47 ("From an issuer's perspective, issuing debt means borrowing money from the investor.").

164. See Goldman Sachs Global Investment Research, *Why Market Liquidity has Deteriorated*, 37 *TOP OF MIND* 1, 7 (2015), <http://www.goldmansachs.com/our-thinking/pages/macroeconomic-insights-folder/liquidity-top-of-mind/pdf.pdf> [<https://perma.cc/FP83-BVLW>] (listing pensions, insurance companies, banks and credit unions, other financial companies, broker-dealers, mutual funds and ETFs, households & non-profits, and the rest of the world as corporate bond in a chart titled "Changing Hands.").

165. See Woepking, *supra* note 163, at 235.

166. See Pearson et al., *supra* note 35, at 39; see also Scott, *supra* note 5, at 1051 (stating the underlying collateral also affects the risk component).

167. JAMES C. VAN HORNE, *FINANCIAL MARKET RATES AND FLOWS* 2 (5th ed. 1998) ("[T]he function of financial market is to facilitate the flow of savings from savings-surplus units to savings-deficit ones."); see JOHNSON, *supra* note 141, at 25 (reiterating the role of the financial markets in transferring funds, and stating businesses are almost always "deficit units," in that their current expenditures exceed their income from current projects and production; they are net borrowers).

168. See VAN HORNE, *supra* note 167, at 29.

ket rates float<sup>169</sup> in response to macroeconomic factors including the state of the economy, government policies (both monetary and fiscal), the current and expected levels of inflation, and international factors such as exchange rates and the flow of foreign capital.<sup>170</sup>

Commercial lenders provide a crucial intermediation function in financial markets.<sup>171</sup> Lenders borrow (or purchase) funds at one interest rate, whether from depositors or other commercial lenders, and then lend (or sell) those funds at a higher interest rate to borrowers.<sup>172</sup> Interest rates are therefore a source of revenue to commercial lenders, as well as a cost. The rate at which commercial lenders borrow is referred to as their “cost of funds.”<sup>173</sup> Since most commercial lenders are extremely creditworthy, the interest rates they must pay to their lenders are low.<sup>174</sup> In exchange for acting as intermediaries, commercial lenders earn a “spread,” or the margin between their funding costs and lending rates to borrowers.<sup>175</sup>

#### D. Present Value and “Profit”

Despite their agreement over the need to determine present value as part of a cramdown analysis, courts remain divided over whether present value should include profits for the secured creditor. As described above, an interest rate must be applied to a deferred stream of payments to ensure they are equivalent to the present value of a secured creditor’s claim. Many bankruptcy courts, such as the *Momentive* court, have adopted the view that the “value of a creditor’s allowed claim does not include any degree of profit”<sup>176</sup> and that the interest rate used to determine present value should not reflect the expectation of such an inclusion.<sup>177</sup> Various

---

169. Roger S. Cox, *Bankruptcy & Creditors’ Rights*, 48 SMU L. REV. 875, 909 (1995) (stating it is common practice in commercial loan transactions to use variable interest rates).

170. See JOHNSON, *supra* note 141, at 147.

171. See GOLIN & DELHAISE, *supra* note 148, at 87.

172. See *id.* at 88.

173. See TRACY G. HERRICK, *BANK ANALYST’S HANDBOOK* 163 (1st ed. 1973).

174. As of October 2017, the 1 month, 3 month, 6 month, and 1 year LIBOR was below 2%. See *LIBOR, other interest rates*, BANKRATE (last visited Oct. 4, 2017), <http://www.bankrate.com/rates/interest-rates/libor.aspx> [<https://perma.cc/C2PW-HDKZ>].

175. See GOLIN & DELHAISE, *supra* note 148, at 88.

176. *GMAC v. Valenti*, 105 F.3d 55, 64 (2d Cir. 1997).

177. See Elson, *supra* note 70, at 1932; *In re Smith*, 178 B.R. 946, 953 (Bankr. D. Vt. 1995) (“[T]he Bankruptcy Code simply does not require that reorganization plans provide secured creditors with their contractual profits, protection from risk, collection costs, etc. Rather, it requires only that creditors receive the present value

rationales underlie this perspective. For instance, the rate at issue in a cramdown attaches to a secured claim, not a loan.<sup>178</sup> Present value is therefore equivalent only to the value of the collateral on which the creditor may not foreclose, rather the value of the loan itself.<sup>179</sup> In addition, the source of the obligation to pay interest differs between a claim and loan: “a loan includes a contractual obligation to pay interest” whereas “the obligation to pay interest on a claim . . . is statutory, not contractual.”<sup>180</sup> When addressing secured creditor claims, the amount of interest should not be determined by reference to a contract or the market, but should instead rely on the appropriate statute and relevant judicial interpretations.<sup>181</sup> The bankruptcy court in *In re Dingley*, among others, has also opined that the Bankruptcy Code protects the creditor’s claim or interest in the property collateralizing the debtor’s pre-bankruptcy obligation, “not the creditor’s interest in the profit it had hoped to make on the loan.”<sup>182</sup> Using a rate that includes transaction costs, costs associated with advertising and overhead, and profit<sup>183</sup> would generate a stream of payments to the secured creditors whose present value was greater than the value of its claim.<sup>184</sup>

Yet, the foregoing arguments misunderstand “profit” in the context of interest rates and financial markets generally. Moreover, courts are not clear on what is meant by “profit” in a cramdown interest rate.<sup>185</sup> By reducing pre-petition bankruptcy rates based on some ambiguous notion of illegitimate “profit,” courts will consistently assign below-market interest rates to creditors, virtually ensuring they will not receive the present value of their loans. In reality, “profit” is the real interest rate component of an interest rate, or the cost to the creditor of “deferring present consumption until a

---

of their claims.”); *In re Dingley*, 189 B.R. 264, 269 (Bankr. N.D.N.Y. 1995) (arguing the plain language of Chapter 13 does not indicate any intent to preserve the “lender’s contract ‘market’ rate of interest i.e., profit” and that had Congress wished to provide creditors with their original interest rates, it would have drafted the statute accordingly).

178. *See Smith*, 178 B.R. at 951 (arguing claims “arise out of loans which went bad, but they are in no sense of the word ‘loans.’”).

179. *See Zywicki*, *supra* note 69, at 255.

180. *Smith*, 178 B.R. at 951.

181. *See id.*

182. *Dingley*, 189 B.R. at 269 (quoting *In re Hudock*, 124 B.R. 532, 534 (Bankr. N.D. Ill. 1991)).

183. *See Elson*, *supra* note 70, at 1932.

184. *See In re Cellular Info. Sys.*, 171 B.R. 926, 939 (Bankr. S.D.N.Y. 1994).

185. *See Scott*, *supra* note 5, at 1053 (stating “[c]ourts have concluded that interest rates should not include a ‘profit’ element. The meaning of that assertion is unclear.”).

later time” and “the normal return on capital without which loans would not be made.”<sup>186</sup> Even in a hypothetical risk-free and inflation-less environment, there would still be a positive real rate of interest. Otherwise, lenders would have no inducement to agree to an “exchange between current and future consumption.”<sup>187</sup> There is no “profit” in delayed consumption of funds.<sup>188</sup> Instead, “profit” is better viewed as a cost of capital requisite to entice the deferred use of money or the opportunity cost of having those funds locked up for a given period of time. By denying “profit” to creditors, courts deprive them of compensation for the taking of their present use of capital.<sup>189</sup>

As discussed, the real rate of interest and inflationary expectation should be identical for all borrowers. Interest rates vary depending on the riskiness of an individual borrower.<sup>190</sup> The risk premium is “the return to capital which compensates the owners of capital for the risk involved in its use in business ventures.”<sup>191</sup> Commercial lenders, with large capital cushions and often-implicit governmental backing, are low risk borrowers. Little compensation is required to induce the transfer of funds between them. Thus, lenders borrow funds from one another at a very low rate of interest, such as LIBOR, and extend credit to comparatively risky entities at higher interest rates.<sup>192</sup> The business model of commercial lending is predicated upon the assumption of this risk, a function for which lenders are necessarily compensated by capturing the interest rate spread. Spread, however, represents gross revenue, from which lenders must cover overhead, rent, salaries, and other costs of doing business. Lenders do not “profit” as a result of capturing the spread, in the sense of taking advantage of a borrower. Rather, they are compensated for providing a vital—and risky—economic function.

---

186. Zywicki, *supra* note 69, at 261; *see also* Monica Hartman, *Selecting the Correct Cramdown Interest Rate in Chapter 11 and Chapter 13 Bankruptcies*, 47 UCLA L. REV. 521, 535 (1999) (“Profit cannot be clearly distinguished from the break-even point or real rate of return that lenders require in order to make any loan.”).

187. Scott, *supra* note 5, at 1047.

188. Zywicki, *supra* note 69, at 261.

189. *Id.* at 261-62.

190. *See* section, I.A. *supra*.

191. PEARCE, *supra* note 20, at 388.

192. *See* Simon Constable, *What Is Net Interest Margin?*, WALL STREET J. (May 7, 2017), <https://www.wsj.com/articles/what-is-net-interest-margin-1494209342> (“A bank usually bases what it charges borrowers on long-term interest rates, while the interest it pays to get funds is based on short-term rates.”).

E. *The Market Value of the Momentive First Lien Notes Decreased as a Result of the Use of the Formula Rate Approach*

To demonstrate the decrease in value of the First Lien Notes resulting from the use by the *Momentive* court of the formula rate approach, I make the following assumptions: firstly, the market's valuation of the debt may be inferred by reference to the loan commitments obtained by the debtors to fund the cash-out alternative for the First and 1.5 Lien holders under the Plan's deathtrap provision.<sup>193</sup> The rate for the first lien committed facility was 4% plus LIBOR (with a floor on LIBOR of 1%).<sup>194</sup> The appropriate discount rate, therefore, was 5%, at a minimum. Secondly, interest payments of \$22,550,000 were made semi-annually, or the \$1.1 billion face value multiplied by 2.05%, half the stated 4.1% coupon interest rate under the cramdown. The present value of the First Lien Notes was approximately \$1,042,129,985, or \$57,870,015 below face value. The formula below demonstrates this result:

$$\begin{aligned} & \textit{First Lien Replacement Notes} \\ & = \$22,550,000 * \frac{\left[1 - \left[\frac{1}{(1.025)^{14}}\right]\right]}{.025} + \frac{\$1,100,000,000}{(1.025)^{14}} \end{aligned}$$

The debtor emerged from bankruptcy on October 24.<sup>195</sup> The present value of the replacement First Lien Notes, which began trading on October 31, was diminished even further. At this time, the market rate of interest was 6.026%<sup>196</sup> while the coupon rate on the replacement notes decreased from 4.1% to 3.88%.<sup>197</sup> The decreased coupon rate was likely a result of changes in rates on seven-year treasuries between September 9, 2014, the date of Judge Drain's bench decision, and late October, when the replacement

193. *In re* MPM Silicones, LLC, No. 14-22503-RDD, 2014 WL 4436335, at \*29 (Bankr. S.D.N.Y. Sept. 9, 2014).

194. *See id.*

195. Michael Vitti, *Taking a Deeper Look Into Momentive, Part I*, NATIONAL ASSOCIATION OF CERTIFIED VALUATORS AND ANALYSTS (Dec. 22, 2014), <http://quickreadbuzz.com/2015/12/22/taking-a-deeper-look-into-momentive-part-1/> [<https://perma.cc/KC6B-77X9>].

196. The market rate of interest is inferred from data obtained from Bloomberg Finance L.P.

197. *See Daily Treasury Yield Curve Rates*, U.S. DEPARTMENT OF TREASURY, <https://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldYear&year=2014> [<https://perma.cc/D9SE-9U5J>] (last visited Jan. 15, 2017).

notes began trading. It is not clear why the market rate of interest for *Momentive's* securities increased over the time period in question. However, the market's valuation is consistent with the B3 rating Moody's assigned to the replacement notes<sup>198</sup> and market interest rates on comparable corporate bonds at the time.<sup>199</sup> The replacement notes were valued by the market at approximately \$956,839,728, or \$143,129,984 below face value.<sup>200</sup>

*First Lien Replacement Notes*

$$= \$21,340,000 * \frac{\left[1 - \left[\frac{1}{(1.031)^{14}}\right]\right]}{.031} + \frac{\$1,100,000,000}{(1.031)^{14}}$$

Evidently, the market value of the notes decreased considerably, particularly after *Momentive's* confirmation hearings. The *Momentive* First Lien Noteholders' claim that "the face amount of the Replacement Notes will equal the amount of the Allowed Claims [and] the true value of the Replacement Notes will be far less"<sup>201</sup> is indeed correct. The graph below demonstrates the market's valuation of the notes over time and comports with the present valuation analyses above.<sup>202</sup>

---

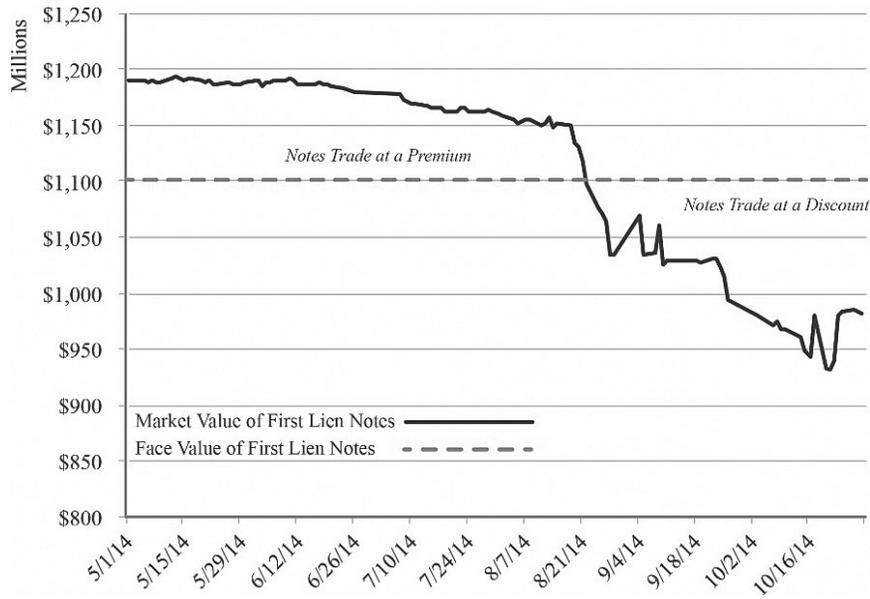
198. See Anthony Hill, *Rating Action: Moody's Assigns B3 to Momentive Performance's First-Lien Notes; Outlook Stable*, MOODY'S INVESTOR SERVICE (Jan. 16, 2015), [https://www.moody's.com/research/Moodys-assigns-B3-to-Momentive-Performances-First-Lien-Notes-outlook—PR\\_316652](https://www.moody's.com/research/Moodys-assigns-B3-to-Momentive-Performances-First-Lien-Notes-outlook—PR_316652) [<https://perma.cc/W4VS-5QH5>].

199. The BofA Merrill Lynch US High Yield B Effective Yield indicates a 6.06% rate of interest for October 30, 2014. See ICE Benchmark Administration Ltd., *ICE BofAML US High Yield B Effective Yield*, FEDERAL RESERVE BANK OF ST. LOUIS, <https://fred.stlouisfed.org/series/BAMLH0A2HYBEY> (last visited Jan. 15, 2017) [<https://perma.cc/SL8B-AG6G>].

200. Market value is inferred from data obtained from Bloomberg Finance L.P.

201. Cramdown Objection of Wilmington Trust, *supra* note 103, at 2-3.

202. Market value is inferred from data obtained from Bloomberg Finance L.P.



Prior to the onset of the confirmation hearings, the First Lien Notes traded at a premium.<sup>203</sup> By the date of the bench ruling, their market value had decreased, and the securities traded instead at a significant discount to face value, forcing the Momentive secured creditors to take heavy losses.

#### IV. ARGUMENTS IN FAVOR OF THE FORMULA RATE APPROACH

The formula rate approach, as applied in *Momentive* and *Till*, has engendered support for a variety reasons, including its potential to decrease unpredictability and the evidentiary costs on debtors and courts in Chapter 11. In addition, proponents argue it does not treat creditors unfairly, since creditors contemplate default ex ante and have the ability to offset risk through portfolio diversification and specialized drafting of individual credit agreements or bond indentures.<sup>204</sup>

203. Market value is inferred from data obtained from Bloomberg Finance L.P.

204. *But see* Scott, *supra* note 5, at 1055 (noting that when taken to its logical conclusion, this argument supports denying the creditor the value of its entire claim, since the creditor's initial contract rate contemplated this result, however,

Minimizing uncertainty and costs in bankruptcy is a valuable objective given the considerations a company or its creditors must weigh in determining whether to file for Chapter 11 or seek an alternative strategy. Adherence by courts to the principle of *stare decisis*, in addition to utilizing jurisprudential procedures that generate predictable outcomes, facilitates this objective.<sup>205</sup> The formula rate approach's reliance on treasury or prime, in addition to the view of both *Till* and *Momentive* that a 1-3% risk adjustment would be appropriate in most instances,<sup>206</sup> suggests cramdown interest rates will be contained within a relatively narrow band.<sup>207</sup> Particularly for the debtor, the certainty afforded by an approximate cap on cramdown interest rates aids in the planning required to file for Chapter 11, provides significant leverage in the negotiation of a plan, and facilitates securing additional post-Chapter 11 financing. By comparison, market-based tests can lead to a wider range of potential outcomes and debtors may find themselves saddled with large and unanticipated interest payments.<sup>208</sup>

In addition, proponents of the formula rate approach claim it reduces the evidentiary and financial burdens of the bankruptcy process on both the debtor and courts.<sup>209</sup> Moreover, under the

---

“[i]f courts took such a position, market rates would skyrocket and destroy an efficient credit economy.”).

205. See Thomas S. Green, *An Analysis of the Advantages of Non-Market Based Approaches for Determining Chapter 11 Cramdown Rates: A Legal and Financial Perspective*, 46 SETON HALL L. REV. 1151, 1176 (2016).

206. See *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 480 (2004); *GMAC v. Valenti*, 105 F.3d 55, 64-65 (2d Cir. 1997); *In re MPM Silicones, LLC*, No. 14-22503-RDD, 2014 WL 4436335, at \*26 (Bankr. S.D.N.Y. Sept. 9, 2014).

207. However, the *Till* court declined to explicitly “decide the proper scale for the risk adjustment.” 541 U.S. at 480. Perhaps as a result, not all courts have limited their risk adjustments to this range. See *McDonald v. Credit Acceptance Co.*, No. 11-12508-BC, 2011 WL 6643074, at \*1; (E.D. Mich. Dec. 21, 2011) (holding the “prime-plus” formula under *Till* does not prohibit interest rates exceeding three percent over prime and adding a “little less than twelve percent over prime” in a case involving a particularly high-risk debtor); *In re Burt*, No. 07-14617-DWH, 2008 WL 4365894, at \*1 (Bankr. N.D. Miss. May 1, 2008) (holding “a 4.25% risk factor should be added to the prime rate for an effective repayment interest rate of 9.5%.”).

208. See Green, *supra* note 205, at 1176-77.

209. See *id.* at 1176; see also Evan D. Flaschen, David L. Lawton & Mark E. Dendinger, *Losing Momentive: A Roadmap to Higher Cramdown Interest Rates*, 5 HARV. BUS. L. REV. ONLINE 100, 105 (2015), <http://www.hblr.org/wp-content/uploads/2015/06/Flaschen-Losing-Momentive-2.pdf> [<https://perma.cc/KKH9-3VPE>] (describing in a positive light modified versions of a formula rate approach that utilize a more “rigid standard” to avoid “significant” evidentiary costs that may arise in situations “without clear evidence of an established market for the secured

formula rate approach, there is no need to determine whether an efficient market exists.<sup>210</sup> For debtors, cost reduction is beneficial because it can result in more rapid emergence from Chapter 11 by healthier companies who enjoy greater financial resources.<sup>211</sup> The costs associated with spending more time in Chapter 11 are already numerous. These include losing customers and market share to competitors, needing approval by the court for transactions outside the ordinary course of business, and being required to pay statutory committee expenses.<sup>212</sup>

While each of the non-formula rate approaches may involve legal fees and expert fees to resolve disputes,<sup>213</sup> under the formula rate approach it should be relatively easy for courts to ascertain the current treasury or prime rate,<sup>214</sup> and the debtor's briefing will provide the court with much of the evidence necessary for determining an appropriate risk premium.<sup>215</sup> Courts are presumably comfortable examining this evidence, since they already must do so as part of the feasibility analysis required under section 1129(a)(11).<sup>216</sup> Such familiarity and expertise may save litigation costs, court time, and resources.<sup>217</sup>

Finally, formula rate approach supporters claim that as part of their normal operations, creditors contemplate default. The risk of

---

claim in question, evidence of comparable rates under similar circumstances, and the risk of nonpayment.”).

210. See Green, *supra* note 205, at 1198 (arguing “the potential uncertainty with respect to the efficiency” of markets in the Chapter 11 context means courts should continue to utilize a formula or non-market-based approach to determine cramdown rates).

211. See *id.* at 1175 (“Unanticipated post-petition fees and expenses can be detrimental to the debtor’s business. The evidentiary costs of a hearing on the efficiency of the market for comparable loans create another obstacle for a [debtor] attempting to get a plan confirmed. More financial resources are drained, and the likelihood of the debtor’s emergence from bankruptcy decreases.”).

212. WEIL, GOTSHAL AND MANGES, LLP, REORGANIZING FAILING BUSINESSES: A COMPREHENSIVE REVIEW AND ANALYSIS OF FINANCIAL RESTRUCTURING AND BUSINESS REORGANIZATION 12-8 - 12-10 (2006).

213. See *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 478 (2004).

214. See Elson, *supra* note 70, at 1926.

215. See *Till*, 541 U.S. at 479 (such evidence includes the “circumstances of the estate, the nature of the security, and the duration and feasibility of the reorganization plan.”).

216. See Green, *supra* note 205, at 1176; see also 11 U.S.C. § 1129(a)(11)(i) (2010) (requiring confirmation of a plan only if it is “not likely to be followed by the liquidation, or the need for further financial reorganization, of the debtor or any successor to the debtor under the plan . . .”).

217. See Green, *supra* note 205, at 1176.

bad debt is already partially reflected in the prime rate and may be mitigated further by credit agreement or bond indenture loan covenants. Creditors also hold diversified portfolios of debt to offset risk. Not only are creditors better candidates to bear risk than any single borrower, they are also less likely to be surprised by an unfavorable outcome stemming from a relationship with a debtor.<sup>218</sup> Additionally, and related to the benefits of certainty described above, creditors who expect to receive the formula rate may be able to predict the interest rate they will receive if the borrower seeks bankruptcy protection. Creditors may then “price their loans accordingly to prevent the risk of under-compensation.”<sup>219</sup> In this light, the fact that the formula rate approach can force a creditor to accept a replacement note bearing a below-market rate of interest is acceptable.<sup>220</sup>

V.  
ARGUMENTS AGAINST THE FORMULA RATE  
APPROACH AND ITS USE IN *MOMENTIVE*  
IN THE CHAPTER 11 CONTEXT

Notwithstanding the foregoing, the application of the formula rate approach in the Chapter 11 context is inappropriate. Notably, the American Bankruptcy Institute has explicitly recommended rejecting *Till* and the formula rate approach. It has advocated instead for the use of a rate that takes into account “the cost of capital for similar debt issued to companies comparable to the debtor as a reorganized entity.”<sup>221</sup> First, the formula rate approach increases creditor uncertainty, which negatively affects liquidity in and access to credit markets, particularly for less creditworthy borrowers (see Part A below). Second, application of the formula rate approach in Chapter 11 systematically undercompensates creditors who may effectively be coerced into loans to formerly bankrupt debtors on terms that would never have been approved in other contexts (see Part B below). Third, Judge Drain is incorrect in his assumption

---

218. *See id.* at 1177.

219. Elson, *supra* note 70, at 1927.

220. It is noteworthy that this argument for the formula rate approach essentially invites banks and lenders to increase credit rates on their loans to protect against acknowledged under-compensation. There is good reason to question whether borrowers generally would appreciate this approach of higher rates across the board.

221. AM. BANKR. INST., COMM’N TO STUDY THE REFORM OF CHAPTER 11, FINAL REPORT AND RECOMMENDATIONS 236-37 (2012), <https://abiworld.app.box.com/s/vwrcv5xv83aav14dp4h>.

that *Till* applies in the Chapter 11 context.<sup>222</sup> Not only does the *Momentive* court overlook the functional differences between the parties involved in and policy purposes of Chapter 11 as compared to Chapter 13, but it also fails to recognize that use of the formula rate approach provides debtors with valuable options, as well as significant leverage over their creditors.

A. *The Formula Rate Approach Reduces Certainty and Predictability, Which Will Negatively Affect Credit Markets and Increase Interest Rates Across the Board*

Under the formula rate approach, secured creditors face the uncertain outcome of receiving a below-market interest rate on a replacement note. Some have described the impact of the formula rate approach in *Till* as a “high-stakes game of chance.”<sup>223</sup> Depending on the jurisdiction and the bankruptcy judge, creditors may receive a range of results in cramdown proceedings that diverge from their estimates of an appropriate interest rate.<sup>224</sup> Since the formula rate approach ignores the original contract rate of interest negotiated at arm’s length by the parties pre-bankruptcy, there is no foreseeable way for creditors to contract around the risk of receiving less than full compensation.<sup>225</sup>

The risk adjustment itself also creates unpredictability and litigation risk. In Chapter 11, where cramdown financing may have long maturities, the difference between a 1% and 3% risk premium can have an enormous impact on the value of a loan.<sup>226</sup> *Till* lists

---

222. The debtor’s own briefing in *Till* suggests present value analyses pursuant to Chapter 11 are a “complex endeavor beyond the scope of consumer Chapter 13.” See Petitioners’ Reply Brief on the Merits at 1 n.1, *Till v. SCS Credit Corp.*, 541 U.S. 465 (No. 02-1016), 2003 WL 22873081 at \*2.

223. Jon W. Jordan, *No More Russian Roulette: Chapter 13 “Cram-Down” Creditors Take a Bullet*, 70 MO. L. REV. 1385, 1402 (2005).

224. See, e.g., *In re Texas Grand Prairie Hotel Realty, LLC*, 710 F.3d at 327 (5th Cir. 2013) (where the creditor argued for an 8.8% cramdown interest rate; the court ultimately approved a 5% rate); *In re Pluma*, 289 B.R. 151, 153, 157 (Bankr. S.D. Cal. 2003) (creditor argued for a 10% cramdown interest rate; the court approved 4.26%); *In re Duval Manor Assocs.*, 191 B.R. 622, 630-31 (Bankr. E.D. Pa. 1996) (where the creditor argued for a 12.44% cramdown interest rate; the court approved 7%); *GMAC v.*, 105 F.3d 55 (2d Cir. 1997) (where the creditor argued for 15.7%; the court remanded for determination of an appropriate rate based on treasury plus 1 to 3%).

225. See Jordan, *supra* note 223, at 1402-03.

226. A simplified example demonstrates this result: assume a debtor issues bonds with \$500,000,000 in principal, which mature in 10 years and exist in an economic environment where the national prime rate is 3% and the market rate is 5%. The bonds compound semi-annually and are owned exclusively by a single

four factors to be used in determining a risk premium: “(1) the probability of plan failure; (2) the rate of collateral depreciation; (3) the liquidity of the collateral market; and (4) the administrative expenses of enforcement.”<sup>227</sup> Other factors used by courts include the “quality of the lender’s collateral,” the “health . . . of the debtor’s business,” and “the feasibility and duration of the plan.”<sup>228</sup> Rather than facilitating the court’s goal of creating an objective process for determining cramdown interest rates, analysis of these factors remains inherently subjective, particularly with regard to its application in the case of an individual debtor.<sup>229</sup> The prediction that the formula rate approach in *Till* “will inevitably, and perhaps inappropriately, frame discussions as to the range of this [risk] adjustment”<sup>230</sup> has been borne out. Much of the controversy in cramdown cases now revolves around determination of an appropriate risk adjustment and expert testimony is often required.<sup>231</sup>

---

creditor when the debtor files for Chapter 11. On the first day of the cramdown proceeding, the court finds that the debtor is low risk and assigns a 1% risk premium for a total coupon rate of 4% on the post-confirmation debt. The present value of these bonds is approximately \$461 million, resulting in a \$39 million loss. Two days later, new information arises which convinces the court that the debtor is in fact much riskier. As a result, the court modifies its initial judgment and assigns a 3% risk premium, for a total coupon of 6%. The present value of the bonds is now \$539 billion. The difference between the two scenarios, a sum of \$78 million, demonstrates the significant impact of the risk premium and suggests creditors will be willing to fight to prove a debtor is high-risk.

227. *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 484 (2004).

228. *See Texas Grand*, 710 F.3d at 334.

229. *See Greenspan & Nelson*, *supra* note 50, at 70.

230. *Id.*

231. *See, e.g., Texas Grand*, 710 F.3d at 334-35 (Experts for both the creditor and debtor agreed on the formula rate approach as well as the use of a prime rate of 3.25% as the base rate, but disagreed on the size of the appropriate upward risk adjustment. The creditor’s expert argued for a total rate of 8.8% based on factors including the circumstances of the bankruptcy estate and the plan’s feasibility, while the debtor’s expert argued instead for a 1.75% adjustment.); *Woodmere*, 178 B.R. at 362 (the cramdown interest rate was “derived by applying the six (6) year U.S. Treasury Note plus a risk factor of 225 basis points to derive a percentage of 7.62%. Both parties offered expert testimony as to the appropriate interest rate” but the debtor’s expert witness argued for a risk factor of 225 basis points while the creditor’s expert witness argued for 450 to 500 basis points); *In re Gramercy Twins Assocs.*, 187 B.R. 112, 124 (Bankr. S.D.N.Y. 1995) (finding a 425 basis point risk premium was necessary to provide a creditor with the present value of its claim, despite creditor’s expert testimony for 500 to 575 basis points); *In re Danny Thomas Props. III Ltd. P’ship*, 231 B.R. 298, 301 (Bankr. E.D. Ark. 1999) (where the creditor’s cramdown interest calculation relied on a 30-year treasury obligation, plus a 2.5-3% risk premium, while the debtor instead argued for a base rate using a 10-year treasury obligation and added 2% for risk. Significantly, “[b]oth sides presented expert testimony as to the appropriateness of the respective

Even within a 1-3% band, there is likely to be considerable disagreement and significant evidentiary burdens for either party to support their proposed rate.<sup>232</sup>

The unpredictability described above can reasonably be expected to have a variety of negative effects on credit markets.<sup>233</sup> When creditors cannot identify with certainty and without excessive financial burden those borrowers who will ultimately file for bankruptcy, the creditors will default to treating all debtors the same. When provisions of the Bankruptcy Code shift risk from debtors to creditors, creditors will ex ante increase the rates of interest they require from all debtors in order to offset the risk of receiving a below-market and under-compensatory cramdown interest rate.<sup>234</sup> The dissenting opinion in *Till* underscored this potential outcome, arguing “[i]f subprime lenders are systematically undercompensated in bankruptcy, they will charge higher rates . . . .”<sup>235</sup> Viewed systematically, across the board interest rate hikes will wipe out the apparent benefit to borrowers as a class of the pro-debtor formula rate approach.<sup>236</sup> While individual debtors may receive a windfall in specific Chapter 11 cases, borrowers as a whole will suffer.<sup>237</sup>

Empirical evidence bears out the prediction that interest rates will increase in response to creditor uncertainty. In an analysis of judicial rulings from October 1979 through June 1993, Joshua Goodman and Adam Levitin determined the availability of mortgage cramdowns to Chapter 13 debtors increased interest rates on

---

rate.”); see also Daniel J. Carragher, *What the Supreme Court’s Prime Plus Ruling Means for Chapter 11*, AM. BANKR. INST. (July/Aug. 2004), <https://www.abi.org/abi-journal/what-the-supreme-courts-prime-plus-ruling-means-for-chapter-11> [https://perma.cc/ESC6-MHX6] (stating a “battle of the experts” will be necessary to determine the risk premium in a Chapter 11 cramdown).

232. See *In re Gillikin*, No. 09-60178, 2011 WL 7704353, at \*5 n.5, (Bankr. S.D. Ga. Mar. 3, 2011) (rejecting reliance on the “*Till* rate” of 1-3% as a per se risk adjustment, requiring instead that the debtor justify the proposed cramdown interest rate and then denying the debtor’s proposed disclosure statement for insufficiently disclosing its rationale for a suggested rate of 4%).

233. See, e.g., Wong, *supra* note 12 (describing the benefits of predictability as “lower lending risks” and enabling “lenders to appropriately evaluate the debt market.”).

234. See Zywicki, *supra* note 69, at 263-64; Jordan, *supra* note 223, at 1404 (stating as a result of cramdowns, “lending rates are likely to increase for all borrowers as creditors attempt to compensate for their bankruptcy losses.”).

235. *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 508 (2004) (Scalia, J., dissenting).

236. See Zywicki, *supra* note 69, at 263.

237. See *id.* at 264-65.

home loans by an average of 12 to 16 basis points.<sup>238</sup> The study suggests this effect might have been more pronounced had the total number of cramdowns been greater at the time.<sup>239</sup> Rajashri Chakrabarti and Nathaniel Pattison, focusing on auto loan rates after enactment of the Bankruptcy Abuse Prevention and Consumer Protection Act, found that eliminating cramdowns resulted in an average 15-basis-point decrease in the interest rate on 48-month auto loans. In states with the highest rates of Chapter 13 filings, interest rates decreased by 30 basis points.<sup>240</sup> Admittedly, both of these studies are outside the Chapter 11 context and implicate forms of cramdown outside the purview of this Note. However, they both suggest the commercial lending market will respond to the availability of cramdowns. While the studies do not compare the cramdown interest rate to then-prevailing market rates, it is reasonable to assume that the greater the spread between these rates and the steeper the potential haircut suffered by the creditor, the more pronounced will be the effect on credit markets.

Interest rate increases will make it harder for borrowers to access capital, especially those at the margins,<sup>241</sup> and may increase their likelihood of default if they do.<sup>242</sup> For instance, the elimination of cramdowns may increase the size of auto loans issued, especially among subprime borrowers,<sup>243</sup> and the availability of

---

238. See Joshua Goodman & Adam Levitin, *Bankruptcy Law and the Cost of Credit: The Impact of Cramdown on Mortgage Interest Rates*, 57 J.L. & ECON. 139 (2014).

239. See *id.* at 156 (stating “Chapter 13 bankruptcy filings were relatively uncommon during the period in question, peaking at .6 percent of all residential mortgages, so Chapter 13 cramdown would have been a comparatively rare occurrence.”).

240. See Rajashri Chakrabarti & Nathaniel Pattison, *Auto Credit and the 2005 Bankruptcy Reform: The Impact of Eliminating Cramdowns*, FED. RESERVE BANK OF N.Y., Rep. No 797, at 3 (Oct. 2016).

241. See Jordan, *supra* note 223, at 1404 (arguing higher interest rates will preclude some high-risk debtors from obtaining credit); see also *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 508 (2004) (Scalia, J., dissenting) (“If subprime lenders are systematically undercompensated in bankruptcy, they will charge higher rates or, if they already charge the legal maximum under state law, lend to fewer of the riskiest borrowers. As a result, some marginal but deserving borrowers will be denied vehicle loans in the first place. Congress evidently concluded that widespread access to credit is worth preserving, even if it means being ungenerous to sympathetic debtors.”).

242. See Jordan, *supra* note 223, at 1405 (“By shifting the financial responsibility to non-bankrupt debtors who are able to obtain credit, the additional burden may force more debtors into bankruptcy.”); Zywicki, *supra* note 69, at 264.

243. See Chakrabarti & Pattison, *supra* note 240, at 25 (finding new subprime loans increased by \$142, relative to new prime loans, when cramdowns were prohibited).

mortgage credit.<sup>244</sup> A chilled credit market is also likely to dampen economic growth, since borrowers will not be able to access funds for new projects and investments.<sup>245</sup> Accordingly, as part of its recommendation to reject the formula rate approach, the American Bankruptcy Institute concluded that the method as applied in *Till* could have negative impacts on distressed debt markets and the liquidity they provide to borrowers.<sup>246</sup>

*B. The Formula Rate Approach Fails to Compensate Secured Creditor Claims in the Risky Bankruptcy Context*

The formula rate approach systematically fails to compensate secured creditors for assuming the risks associated with transacting with borrowers<sup>247</sup> and fails to provide secured creditors with the full value of their claims, as required by section 1129(b)(2)(A)(i)(II). In doing so, it represents a wholesale shifting of the risks and costs associated with bankruptcy and default to the secured creditor.<sup>248</sup> This is problematic for three reasons. Firstly, Chapter 11 debtors are at least as risky, if not riskier, than their counterparts who never enter bankruptcy. Secondly, default is more expensive in Chapter 11 than under other circumstances. Finally, cramdown loans are frequently coerced without equity cushions.

Chapter 11 plans are not a guarantee of future financial success or the debtor's ability to repay a replacement note in full, re-

---

244. See Wenli Li, Ishani Tewari & Michelle J. White, *Using Bankruptcy to Reduce Foreclose: Does Strip-Down of Mortgages Affect the Supply of Mortgage Credit?* 4 (Nat'l Bureau of Econ. Research, Working Paper No. 19952, 2014), <http://www.nber.org/papers/w19952> [<https://perma.cc/4BAX-XJNY>] (the elimination of cramdowns increased mortgage credit approval rates "by 0.9 percentage points, or 1%, in affected relative to unaffected regions.").

245. See Robert K. Rasmussen & David A. Skeel, Jr., *The Economic Analysis of Corporate Bankruptcy Law*, 3 AM. BANKR. INST. L. REV. 85 (1995) ("More expensive credit leads to a reduction in economic activity."). The financial crisis is an extreme example of the effects of a severely dampened credit market on the economy. See Eamonn K. Moran, *Wall Street Meets Main Street: Understanding the Financial Crisis*, 13 N.C. BANKING INST. 5, 73 (2009) (describing the financial crisis's reduction in the general availability of credit to businesses, which decreased their ability "to meet payrolls, pay suppliers, and purchase inventory," and significantly disrupted their ability to "borrow and finance spending, investment, and job creation.").

246. See AM. BANKR. INST., COMM'N TO STUDY THE REFORM OF CHAPTER 11: 2012~2014 FINAL REPORT AND RECOMMENDATIONS 236 (2014), <http://commission.abi.org/full-report> [<https://perma.cc/3544-JXZG>].

247. See *id.*; *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 491 (2004) (Scalia, J., dissenting).

248. See Zywicki, *supra* note 69, at 263 (describing the transfer of risk from debtors to creditors as a result of under-compensatory cramdown interest rates).

ardless of court or trustee supervision during the case or a successful restructuring.<sup>249</sup> One estimate of Chapter 11 recidivism, or the percentage of companies that emerge from Chapter 11 only to reenter a few years later (colloquially referred to as “Chapter 22,” or worse, “Chapter 33”), is approximately 15%.<sup>250</sup> Others have observed recidivism rates as high as one third of companies that emerge from Chapter 11.<sup>251</sup> The *Till* dissent points out that since already-bankrupt borrowers have demonstrated histories of financial instability, they are riskier than other debtors and more likely, “or at the very least, not systematically less” likely to default in the future or end up back in Chapter 11.<sup>252</sup> The formula rate approach itself suggests risk in bankruptcy is impossible to eliminate. If future results were guaranteed, the risk-free rate would be sufficient as a cramdown rate of interest and lending to Chapter 11 borrowers would be as safe as lending to the United States government. That a risk premium is almost always added indicates that risk of default is still contemplated after bankruptcy. Yet, the formula rate approach denies creditors full compensation for assuming it.<sup>253</sup>

While default is an expensive proposition to creditors generally, it is more so in the Chapter 11 bankruptcy context. Various empirical analyses have determined that as a proportion of assets, Chapter 11 costs more than both prepackaged bankruptcies and out-of-court restructurings.<sup>254</sup> There are various explanations for

---

249. See *Till*, 541 U.S. at 493-94 (Scalia, J., dissenting).

250. Edward I. Altman, *Revisiting the Recidivism – Chapter 22 Phenomenon in the U.S. Bankruptcy System*, 8 BROOK. J. CORP. FIN. & COM. L. 253, 273 (2014).

251. See Lynn M. Lopucki & Joseph W. Doherty, *Why Are Delaware and New York Bankruptcy Reorganizations Failing?*, 55 VAND. L. REV. 1933, 1934 (2002) (“[Prior] empirical finding[s] [show] that by February 2000, nine of the thirty companies (30%) emerging from bankruptcy reorganization in Delaware from 1991 to 1996 had filed for bankruptcy a second time” and that New York “had a refiling rate almost as high as Delaware’s (23%).”); Edward I. Altman, *Post-Chapter 11 Performance: Avoiding Chapter 22*, 21 J. APPLIED CORP. FIN. 53, 55 (“[R]oughly one-third of those firms emerging as a publicly registered company experience some form of subsequent distressed restructuring again, including the filing of a second (or third, or even more) bankruptcy.”).

252. *Till*, 541 U.S. at 494 (Scalia, J., dissenting). While certainly debatable, the dissent’s comment is supported by the statistics in Section IV.B, *infra*.

253. See *Gen. Motors Acceptance Corp. v. Jones*, 999 F.2d 63, 68 (3d Cir. 1993) (“[I]f the creditor had been allowed to freely construct and extend the loans coerced by the cramdown proceedings, it would have increased the lending rate based on the debtors’ demonstrably poor credit history.”).

254. Elizabeth Tashjian, Ronald C. Lease & John J. McConnell, *Prepacks: An Empirical Analysis of Prepackaged Bankruptcies*, 40 J. FIN. ECON. 135, 143 (1996) (“As a fraction of assets, the fees paid in prepacks lie between the average of 2.8% . . . for traditional Chapter 11 reorganizations and the 0.65% reported . . . for out-of-court

this phenomenon. For instance, the costs of foreclosure are higher in bankruptcy since repossession may be blocked by the automatic stay.<sup>255</sup> Moreover, various disputes may arise during bankruptcy, which, in addition to required disclosures and filings, necessitate costly legal and expert fees.<sup>256</sup> Such fees can easily increase to tens or even hundreds of millions of dollars<sup>257</sup> and serve to decrease the size of the debtor's estate and the potential for full reimbursement of a secured creditor's claim.

In contrast to loans issued in ordinary commercial lending markets, crammed down loans frequently do not contain an equity cushion<sup>258</sup> and the loan-to-value ratio in cramdowns may be close to or exactly 100%.<sup>259</sup> For creditors, equity cushions are important hedging tools, which afford protection against depreciation in the value of the collateral securing their loans.<sup>260</sup> Generally, loans with higher loan-to-value ratios are viewed by creditors as more risky, since the higher the ratio, the greater the creditor's risk of being under-secured if the borrower ultimately defaults.<sup>261</sup> Outside of Chapter 11, a creditor would reasonably increase the rate of interest charged to a borrower to offset the risk for lending without any

---

restructurings."); See Brian L. Betker, *The Administrative Costs of Debt Restructurings: Some Recent Evidence*, FIN. MGMT. 26, 57 (1997) (finding "the direct costs of traditional Chapter 11 cases average 3.93% of pre-bankruptcy total assets. This figure is significantly larger than average direct costs for prepacks (2.85%).").

255. See *Till*, 541 U.S. at 493 (Scalia, J., dissenting).

256. See Hon. Alexander L. Paskay & Frances Pilaro Wolstenholme, *Chapter 11: A Growing Cash Cow Some Thoughts on How to Rein in the System*, 1 AM. BANKR. INST. L. REV. 331, 335 (1993) ("There is no doubt that Chapter 11 has gained the reputation of a 'cash cow' for the professionals involved in the process.").

257. See Rasmussen & Skeel, Jr., *supra* note 245, at 90; Steve H. Nickles & Edward S. Adams, *Tracing Proceeds to Attorneys' Pockets (and the Dilemma of Paying for Bankruptcy)*, 78 MINN. L. REV. 1079 (1994) (describing the Eastern Airlines, Pan American Airlines, and LTV Corporation bankruptcies, where combined costs exceeded \$250 million).

258. An equity cushion refers to the value of collateral securing a loan in excess of the amount of that loan.

259. See Zywicki, *supra* note 69, at 260; see also *In re Gramercy Twins Assocs.*, 187 B.R. 112, 124 (Bankr. S.D.N.Y. 1995) ("[T]he relatively high loan to value ratio in this case, which is approximately 85%, increases the risk factor."); *In re 20 Bayard Views, LLC*, 445 B.R. 83, 112 (Bankr. E.D.N.Y. 2011) (describing a 100% loan-to-value ratio).

260. See *Gen. Motors Acceptance Corp. v. Jones*, 999 F.2d 63, 68 (3d Cir. 1993).

261. See Greenspan & Nelson, *supra* note 50, at 70.

cushion.<sup>262</sup> However, in the cramdown context, creditors are effectively deprived of the ability to do so.<sup>263</sup>

These risk factors suggest crammed down loans are more costly than those issued under ordinary market conditions<sup>264</sup> and Chapter 11 creditors should be compensated through higher rates of interest. However, the formula rate approach does not compensate the creditor for taking on these risks or make it whole with regard to the potential costs incurred. In failing to do so, it transfers value from the secured creditor to the ordinarily subordinate unsecured creditors or equity holders, potentially in violation of the absolute priority rule.

The absolute priority rule prohibits the allocation of property to junior classes, such as unsecured creditors, unless dissenting senior classes first receive the full value of their claims or all of a debtor's reorganization value.<sup>265</sup> Changes to a debtor's capital structure, which are made to benefit the debtor or a class of the debtor's claimants, may reduce the value available to other classes of holders of claims or interests.<sup>266</sup> A company undergoing reorganization has limited assets that it may distribute to its various stakeholders. Where a debtor receives a windfall in the form of an artificially depressed cramdown interest rate and is obligated to make smaller interest payments to a crammed down creditor, assets are concordantly made available to satisfy other claims, including those that are subordinate. Indeed, a criticism of Chapter 11, generally, is the ability of conflicted managers to inappropriately transfer wealth from creditors to equity holders.<sup>267</sup> By failing to

---

262. See Hartman, *supra* note 186, at 537 (claiming "no creditor would make a loan for 100 percent of the value of its collateral.").

263. See *Gen. Motors*, 999 F.2d at 68.

264. See *id.* (stating crammed down loans are more "costly than new loans that the creditor would have voluntarily extended.").

265. See 11 U.S.C. §§ 1129(b)(2)(B)(ii), § 1129(b)(2)(C)(ii); see also 7 COLLIER, *supra* note 4, at ¶ 1129.03[4][a][i] (the absolute priority is well stated as follows: "[a] plan of reorganization may not allocate any property whatsoever to any junior class on account of the members' interest or claim in a debtor unless all senior classes consent, or unless such senior classes receive property equal in value to the full amount of their allowed claims, or the debtor's reorganization value, whichever is less.").

266. See Steven L. Schwarcz, *The Easy Case for the Priority of Secured Claims in Bankruptcy*, 47 Duke L.J. 425, 429 (1997) (describing the Modigliani-Miller hypothesis to secured lending).

267. See Michael Bradley & Michael Rosenzweig, *The Untenable Case for Chapter 11*, 101 YALE L.J. 1043, 1045-46 (1992); see also Allan C. Eberhart, William T. Moore & Rodney L. Roenfeldt, *Security Pricing and Deviations from the Absolute Priority Rule in Bankruptcy Proceedings*, 45 J. FIN. 1457, 1458 (1990) ("[T]he amount paid

compensate secured creditors, the formula rate approach contravenes the absolute priority rule by providing this value to subordinate claimants who should not receive any payout until the secured creditors have been paid in full.

*C. Because Chapter 11 is Fundamentally Dissimilar to Chapter 13, the Formula Rate Approach Does Not Apply*

The rationales for accepting the formula rate approach, including those described in *Till*,<sup>268</sup> are less applicable in the Chapter 11 context as compared to Chapter 13. The competing goals of the Bankruptcy Code are often described as “rehabilitation” and “reorganization”<sup>269</sup> or providing a fresh start to a debtor while simultaneously allocating limited assets to its creditors.<sup>270</sup> A court overseeing a company through bankruptcy must therefore balance the ability of a debtor to successfully emerge from bankruptcy against the protection of the interests of secured creditors and other constituencies.

Chapter 13, which is aimed at insolvent persons, not businesses, clearly indicates a Congressional preference for the debtor-in-possession’s ability to achieve a fresh start.<sup>271</sup> It is designed to provide maximum relief to individuals who are viewed as deserving of the ability to start over again, freed of outstanding obligations.<sup>272</sup> Given this objective, the rights of creditors may be of secondary importance to the right of the debtor to achieve a start fresh. How-

---

to shareholders in excess of that due under strict adherence to the [absolute priority rule] is measured, and this amount is on average 7.6% of the total value paid to all claimants.”).

268. See *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 479-80 (2004) (describing rationales such as straightforwardness, objectivity, familiarity, and cost reduction).

269. Harvey R. Miller & Shai Y. Waisman, *Is Chapter 11 Bankrupt?*, 47 B.C. L. REV. 129, 149 (2005).

270. See Elson, *supra* note 70, at 1944.

271. See *In re Goeb*, 675 F.2d 1386, 1388 (9th Cir. 1982) (acknowledging “Congress’ effort to give as many debtors as possible a fresh start through Chapter 13’s liberal discharge provisions”) (citations omitted); *In re Barnes*, 13 B.R. 997, 999 (D.D.C. 1981) (“The purpose behind Chapter 13 is to give the debtor a fresh start. . . .”); see also 8 COLLIER ON BANKRUPTCY ¶ 1300.02 (Alan N. Resnick & Henry J. Sommers eds., 16th ed.) (describing Congress’ intent for Chapter 13 as “to encourage, but not require, financially overextended *individual* debtors to make greater voluntary use of repayment plans commensurate with each debtor’s abilities, as the most effective means of improving, first, debtor relief and, second, creditor recoveries”) (emphasis added).

272. See Elson, *supra* note 70, at 1944.

ever, Chapter 11, which is intended for businesses<sup>273</sup> (though individuals can and do file for Chapter 11<sup>274</sup>), does not overtly prioritize this goal.<sup>275</sup> In the absence of an explicit rationale prioritizing the interests of debtors over creditors, there is less reason to apply an anti-creditor method, such as the formula rate approach, in determining cramdown interest rates.

The *Till* plurality's concerns regarding a debtor's financial burden are more salient in Chapter 13, where he or she is an individual with limited personal assets and any added burden could create disadvantages in the case's proceedings.<sup>276</sup> Justice Ginsburg's comments during oral arguments in *Till* "evidenced a concern that typical chapter 13 debtors cannot afford to pay experts, so a calculation method that required one would render proceedings increasingly complicated and prohibitively expensive."<sup>277</sup> Yet, the very consideration of multi-year repayment terms in a Chapter 11 case suggests not only that the debtor's business is economically viable, but also that during the pendency of the Chapter 11 proceeding it has been financially able to hire counsel and a valuation expert or financial advisor.<sup>278</sup> Additionally, the court's reluctance to use approaches requiring outside or market evidence is less relevant in the Chapter 11 context, in which judges often and willingly rely on expert testimony to better understand a debtor's business or de-

---

273. See 7 COLLIER, *supra* note 4, at ¶ 1100.01 (stating "Chapter 11 . . . is fashioned primarily for *business* debtors") (emphasis added).

274. See *Toibb v. Radloff*, 501 U.S. 157 (1991) (holding an individual debtor not engaged in business may reorganize under Chapter 11); see also Anne Lawton, *The Individual Chapter 11 Debtor Pre-and Post-BAPCPA*, 89 AM. BANKR. L.J. 455, 456 (2015) ("As the study's findings reveal, individual debtors comprise a sizeable proportion - more than one in five - of the chapter 11 debtors in the adjusted 2004 and 2007 . . . Moreover, the percentage of chapter 11 cases identified as individual filings increased from approximately 23% in 2004 to slightly more than 27% in 2007.").

275. See Wong, *supra* note 12, at 1944; Elson, *supra* note 70, at 1946 ("Any method of cramdown interest determination in Chapter 11 must therefore aid in the rehabilitation of the debtor, but should not prioritize the debtor's fresh start over the rights of creditors."); see also Miller & Waisman, *supra* note 269, at 144 (noting "Congress' intent to balance the interests of all parties involved in the Chapter 11 reorganization process.").

276. See Mark J. Thompson & Katie M. McDonough, *Lost in Translation: Till v. SCS Creditor Corp. and the Mistaken Transfer of a Consumer Bankruptcy Repayment Formula to Chapter 11 Reorganizations*, 20 FORDHAM J. CORP. & FIN. L. 893, 917 (2015).

277. *Id.* at 916.

278. See *id.* at 917; see Elson, *supra* note 70, at 1947 (stating "[p]arties entering into chapter 11 bankruptcy typically possess a much larger pool of assets than those in personal bankruptcy.").

velop valuations based on the financial markets and courts generally review financial statements and projections in the ordinary course.<sup>279</sup>

Indeed, *Till's* footnote 14 may be read as an attempt to differentiate Chapter 13 from Chapter 11. For instance, the *Till* plurality states that while there is a market for willing cramdown lenders in Chapter 11, “the same is *not* true in the chapter 13 context.”<sup>280</sup> The footnote then goes on to state, “[in] the Chapter 13 context, *by contrast* [to Chapter 11], the absence of any such market obligates courts to look to first principles and ask only what rate will fairly compensate a creditor for its exposure.”<sup>281</sup> A fair reading of this language indicates the plurality was attempting to differentiate bankruptcies that occur in Chapter 11 and Chapter 13 rather than mandate the universal application of the formula rate approach in both contexts.<sup>282</sup>

*D. The Formula Rate Approach Inappropriately Creates an Option and Provides Negotiating Leverage for Chapter 11 Debtors*

Under certain circumstances, a debtor in Chapter 11 has a choice when dealing with secured creditors: it can seek debtor-in-possession financing to cash out its existing lenders or utilize the bankruptcy process to extend its existing secured debt at a court-determined cramdown interest rate. In *Momentive*, the debtor posed these alternatives to its creditors. The Senior Secured Notes could opt for the repayment of all outstanding principal and interest on their notes (without a make-whole) or receive replacement notes with an interest rate to be determined under the Bankruptcy Code.<sup>283</sup>

---

279. See Thompson & McDonough, *supra* note 276, at 917; Deborah Langehennig, *Application of the Till Interest Rate*, 68 TEX. B.J. 1022, 1027-28 (2005); see also *In re MPM Silicones, LLC*, 874 F.3d 787, 801 (2d. Cir. 2017) (“[W]e understand that the complexity of the task of determining an appropriate market rate of interest will vary from case to case. . . [b]ut, at the end of the day, we have no reason to believe the task varies materially in difficulty from the myriad tasks which we regularly rely on the expertise of our bankruptcy courts to resolve.”).

280. *Till v. S.C.S. Credit Corp.*, 541 U.S. 465, 476 n.14 (2004) (emphasis in original).

281. *Id.* (emphasis added).

282. See Thompson & McDonough, *supra* note 276, at 931 (Arguing that by distinguishing between “cramdown under the two chapters, the plurality implicitly adopts a perspective . . . that differentiates consumer bankruptcies from business reorganizations.”).

283. *In re MPM Silicones, LLC* 874 F.3d 787, 801-02 (2d. Cir. 2017).

If the debtor knows *ex ante* it will be able to refinance its debt at a lower interest rate than the one provided by a bankruptcy court, it will choose to pay down its existing obligations and incur new ones. Since the debtor must seek external financing to do so the rate of interest on the new obligations will be at market value. However, if the debtor believes a bankruptcy court will assign a cramdown interest rate that is higher than the prevailing market rate of interest (for instance, because market conditions have changed since the initial obligations were incurred), the debtor will seek to avoid litigating a new interest rate and provide its existing creditors with more favorable terms when attempting to confirm its plan. In either instance, the debtor benefits from the difference between the interest rate it would be forced to pay via a cramdown proceeding and the prevailing market rate. Moreover, even if the debtor does not view this option as such, the threat of a below-market interest rate provides the debtor with significant leverage over its creditors. If a secured creditor knows it stands to receive a below-market interest rate determined under the formula approach, the debtor may be able to compel it to accept covenant lite or other undesirable contractual features to avoid that outcome.

Certainly, the availability of external financing for a Chapter 11 debtor is not a given, and secured creditors can interfere with the debtor's ability to exercise this option profitably. However, the foregoing demonstrates that under the formula rate approach, the debtor is provided with additional value. To minimize that value, the appropriate cramdown methodology should be one that eliminates the gap between the market and cramdown rates. By systematically providing creditors with below-market rates, the formula rate approach fails to do so.

## VI.

### THE TWO-STEP APPROACH ADOPTED BY THE SECOND CIRCUIT IS BETTER SUITED FOR CHAPTER 11 AND THE COMMERCIAL CREDIT CONTEXT

In light of the foregoing, the formula rate approach is an inadequate measure for determining cramdown interest rates. Its one-size-fits-all technique does not adequately address the wide range of debtors, risk factors, and market factors that bankruptcy courts properly should consider as part of this analysis. In addition, by assigning below-market interest rates, the formula rate approach ensures creditors will not receive the present value of their loans. The two-step approach adopted by the Second Circuit, particularly in its reliance on current market interest rates, provides a better alterna-

tive and should be adopted as the standard in calculating cramdown interest rates. Unlike the formula rate approach, the two-step approach proceeds directly to determine the proper market-based rate for the specific post-confirmation loan, which in practice is likely to be readily determinable and within a relatively narrow range. In this regard, the two-step approach: should be easily administrable with low evidentiary costs (see Part A below); will properly compensate the creditor for the risk and costs it incurs in lending to debtors emerging from Chapter 11 (see Part B below); and will facilitate outcome certainty for both creditors and debtors (see Part C below). Use of the two-step approach will prevent the negative effects on creditors and credit markets described above and avoid outcomes where economic growth is hindered by creditor reticence or increased interest rates.

A. *The Two-Step Approach is Easily Administered and Reduces Evidentiary Burdens on the Court and the Debtor*

Reducing the evidentiary and financial burden on both the court and debtor is important in order to decrease strain on the judicial system, ensure the time spent in and associated costs of Chapter 11 are kept to a minimum, and increase the likelihood that the reorganized debtor is left with greater resources. As a practical matter, the two-step approach should easily achieve this objective because there exist a variety of sources of loan pricing information, including SEC filings, bankruptcy court filings, and major financial periodicals, upon which the court can rely to determine the appropriate cramdown rate.<sup>284</sup> Certainly, creditors have greater access to loan-pricing information than bankruptcy courts. Creditors may develop loan-pricing models in-house or purchase them from private vendors. As a result, they can assist in the determination of a market interest rate by filing documentation on their loan-pricing methodology and analysis of the specific debtor. Creditors would likely be incentivized to offer reasonable rates, particularly if the alternative was a cramdown well below market value.

The availability of such information to both parties should serve to narrow the range of negotiation and allow for resolution of

---

284. See Zywicki, *supra* note 69, at 259 (stating the “rates and terms available to borrowers for a standard loan are well-known and easily available to the public, through numerous financial reports and daily publications.”); see also Gary W. Marsh & Matthew M. Weiss, *Chapter 11 Interest Rates After Till*, 84 AM. BANKR. L.J. 209, 224 (2010) (“Established commercial interest rates frequently serve as a starting point for bankruptcy courts when they are attempting to discern market rates of interest.”).

the cramdown rate without litigation. In fact, the bid asked in the negotiation may well be narrower than the 1-3% risk adjustment to be resolved under the formula rate approach. However, if negotiation fails, the extent of expert testimony is likely to be limited to the debtor's financial position and prospects derived from the bankruptcy case. That experts may become involved is a weak argument against the forced loan approach, particularly given the heavy reliance on experts to determine the proper risk adjustment under the formula rate approach.<sup>285</sup>

While the goal of any cramdown interest rate approach must be to keep costs to a minimum, it is not always possible or necessary to eliminate them altogether. As compared to Chapter 13, Chapter 11 has a greater likelihood that debtors can afford to incur certain costs, should they choose to do so.<sup>286</sup> On balance, however, the availability of market data means the two-step approach offers the best chance to minimize such costs, when and if they arise.

Much has been made of the need to isolate an "efficient market" in order to determine a cramdown rate under the coerced loan approach.<sup>287</sup> However, those relying on this argument misunderstand the operation of commercial credit markets. Commercial lenders are fully capable of assessing individual borrowers and their risk profiles. As discussed, creditors have access to multiple internal and external resources for pricing debt. Prior to filing for bankruptcy, the debtor presumably felt confident in the ability of the creditor to do so and believed the rate, negotiated at arms-length, was sufficiently fair (otherwise they would not have agreed to the terms provided). That there was no debate over the efficiency of the market on which the creditor assigned such terms at the time of the original negotiation makes it unclear why the efficiency of the market post-confirmation has become a make-or-break factor in the determination of a cramdown interest rate. The rate proposed will

---

285. See *supra* section III.A.

286. See 7 COLLIER, *supra* note 4, at ¶ 1129.05[2][c][i] (stating the rationale of minimizing costs "while not absent from chapter 11 cases, is certainly minimized in larger chapter 11 cases" presumably as a result of the ability of a Chapter 11 debtor to bear any related evidentiary costs). See also Section IV.C., *supra* (discussing the differences between Chapter 11 and Chapter 13 debtors with regard to the former's superior ability to bear evidentiary costs).

287. See Green, *supra* note 205, at 1180-98 (discussing the application of the efficient capital market hypothesis and its application in the *Till* context for determining whether an efficient market exists on which to base coerced loan cramdown interest rates); Wong, *supra* note 12, at 1947-52 (discussing the efficient market approach and criticizing debtor-in-possession financing as a proxy for an efficient market).

also be subject to court review, though for the reasons mentioned above, the extent of that review is likely to pose minimal burden on bankruptcy courts and parties.

*B. The Two-Step Approach Compensates Creditors for Risk Without Unduly Burdening Debtors*

Cramdown rates based on real market data are the only rates that properly compensate creditors for risk, unless the market is distorted at the time. This occurred, for example, in late 2008 and early 2009.<sup>288</sup> Putting aside such exceptional circumstances, creditors expect a margin over their funding costs, which provides a reasonable return based on the credit risk of the borrower.<sup>289</sup> Were the creditor's funds not locked in the crammed down loan, it would presumably use those funds to capture the market rate on a loan to a different borrower.<sup>290</sup> Market-based cramdown rates afford this result without overcompensating the creditor,<sup>291</sup> since they reflect the credit community's appraisal of financing to similarly situated debtors and with similar terms.

At the same time, there is no undue burden on the reorganized debtor if it is required to pay such rates, since it would be forced to pay them were it to seek an alternative loan in the market. The Chapter 11 reorganization process is not intended to give the reorganized debtor a bonus interest rate and thereby a possible competitive advantage over other companies.<sup>292</sup> Perversely, if the below-market rate allowed the debtor to reduce prices, its competi-

---

288. See *S&P/LSTA U.S. Leveraged Loan 100 Index 10 Year Chart*, S&P DOW JONES INDICES, <https://us.spindices.com/indices/fixed-income/sp-lsta-us-leveraged-loan-100-index> [<https://perma.cc/5X5K-TNK8>] (last visited Nov. 15, 2015).

289. See *Gen. Motors Acceptance Corp. v. Jones*, 999 F.2d 63, 67 (3d Cir. 1993) (noting creditors reasonably anticipate a profit on their loans to debtors); *In re Cassell*, 119 B.R. 89, 92 (W.D. Va. 1990) (“The market rate does include ‘profit’ on the loan. Otherwise, no creditor would willingly make such a loan. The court feels that the profit component of the market rate is properly considered part of the cost of capital.”).

290. See Zywicki, *supra* note 69, at 262.

291. Creditor overcompensation would be prohibited under the fair and equitable standard. See 7 COLLIER, *supra* note 4, at ¶ 1129.03[4][a][i][C], n.101.

292. See Ian Dattner, *Chapter 11 Protection: Whom Are We Protecting?*, 38 COLUM. J.L. & SOC. PROBS. 287, 293 (2005) (describing the unfair competitive advantage afforded to a Chapter 11 debtor, particularly in a capital-intensive business, when it is able to reduce its debts and corresponding interest payments); see also Zywicki, *supra* note 69, at 251-252 (stating the Bankruptcy Code tries to preserve “the substantive rights of parties as defined outside bankruptcy” and therefore, “debtors should not see bankruptcy as an opportunity to improve their positions relative to creditors.”).

tors could be forced into Chapter 11. Creditors might also be forced to recognize such extensive impairments that they would also be faced with economic insolvency and ultimately require reorganization.<sup>293</sup>

C. *The Two-Step Approach Facilitates Outcome Certainty and Protects the Justified Expectations of Creditors and Debtors*

The two-step approach will facilitate outcome certainty in bankruptcy proceedings and protect the reasonable expectations of both creditors and debtors. The debtor and creditor entered into the pre-petition agreement on mutually agreeable terms and interest rates in the broader context of a functioning credit market after receiving competitive bids from other lenders where there was public information on market rates. Moreover, creditors undeniably enter commercial lending agreements with the expectation of generating sufficient revenue to at least fund their operations and cover overhead.<sup>294</sup> To deny them undermines their reasonable expectations and the pre-bankruptcy negotiation process entered into by creditors and debtors at arms-length.

In addition, aside from the fact that the post-confirmation financing is involuntary, there is no compelling reason for a different methodology to address cramdown rates. Once the two-step approach is firmly established, both creditors and debtors will have certainty as to the nature of their negotiations and that the ultimate cramdown rate will balance their interests according to a market that is generally visible from public sources. There may be unusual instances where the market information is not available, for instance, if a borrower has unique assets or businesses, but in that case, falling back on the formula rate approach is reasonable.

---

293. For instance, a “Troubled Debt Restructuring” is a form of impairment that occurs when a creditor agrees to modify the terms of a loan for reasons, whether legal or economic, it would not otherwise consider, such as an in a Chapter 11 cramdown. See OFFICE OF THE COMPTROLLER OF THE CURRENCY, BANK ACCOUNTING ADVISORY SERIES, at 23 (2016), [https://www.occ.gov/publications/publications-by-type/other-publications-reports/baas.pdf](https://www OCC.GOV/publications/publications-by-type/other-publications-reports/baas.pdf) [https://perma.cc/FRZ2-CDPY]. Such modifications may be imposed by a court and encompass any “reduction (absolute or contingent) of the stated interest rate for the remaining original life of the debt,” including interest rates which are below market. *Id.*

294. *Gen. Motors*, 999 F.2d at 67 (“Any creditor extending credit anticipates that over the course of the loan it will recover, in interest, its cost of capital and its cost of service . . . . When we focus on the present value of chapter [11] deferred payments to a secured creditor who has been forced to extend this new credit, this fact should not be ignored.”).

## CONCLUSION

Ultimately, the name of the approach for determining a cramdown interest rate is irrelevant.<sup>295</sup> Instead, the salient question is who should be responsible for assessing the terms and riskiness of a commercial credit transaction. The formula rate approach suggests courts are best able to handle this task and that risk in any case is minimal. The two-step approach adopted by the Second Circuit views the markets as superior. Ideally, the two would generate identical outcomes; however, various policy, social, and financial considerations force them to vary.

Significant controversy and lack of clarity persist in the cramdown interest rates context. Rather than clarifying the competing doctrinal views, *Momentive* and *Till* have only exacerbated the issue and both cases remain subject to criticism in relevant legal and professional communities. Fortunately, the Second Circuit's recent opinion overturning the application of the formula rate approach in Chapter 11 has restored sensibility to the Bankruptcy Code, at least within one portion of the country. In this Note, I have attempted to argue that courts must look to market rates to provide creditors with value equal to that of their interests in the collateral securing their claims on the confirmation date. As discussed, artificial rates supplied by courts do not accurately represent the economics of arms-length creditor-debtor relationships and consistently fail to provide commercial creditors with the present value of their claims. As the only cramdown interest rate methodology that relies on market rates of interest, the two-step approach is a superior option. Not only does the two-step approach reduce the evidentiary burden on the courts, debtors, and creditors as a result of the availability of market information by which the court may determine an appropriate cramdown rate, it also best compensates creditors for the risk they assume in lending to post-Chapter 11 debtors. In addition, the two-step approach affords ex ante certainty that the terms of a credit agreement will be honored and ex post certainty that the ultimate cramdown rate will be fair to both the creditor and debtor relative to a public market. As a result, courts should follow the two-step approach in future Chapter 11 cramdown proceedings.

---

295. See *In re Seasons Partners, LLC*, 439 B.R. 505, 520 (Bankr. D. Ariz. 2010) (“Whether one starts with a ‘base rate’ and adds for risk, or just accepts that a proven market rate includes relevant risk (in an appropriate case), the result should not vary by much.”).