Berkeley

[technology law Journal]

VOLUME 26
NUMBER 3
2011

UNIVERSITY OF CALIFORNIA, BERKELEY
SCHOOL OF LAW
BOALT HALL

Production: Produced by members of the *Berkeley Technology Law Journal*. All editing and layout done using Microsoft Word.

Printer: Joe Christensen, Inc., Lincoln, Nebraska.
Printed in the U.S.A.
The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Library Materials, ANSI Z39.48—1984.

Berkeley Technology Law Journal
U.C. Berkeley School of Law
Student Center, Ste. 3
Berkeley, California 94720-7200
btlj@law.berkeley.edu
http://www.btlj.org

BERKELEY TECHNOLOGY LAW JOURNAL

VOLUME 26 NUMBER 3 SYMPOSIUM 2011

TABLE OF CONTENTS

SYMPOSIUM: TECHNOLOGY: TRANSFORMING THE REGULATORY ENDEAVOR

ARTICLES

NOTES

SUBSCRIBER INFORMATION

The *Berkeley Technology Law Journal* (ISSN1086-3818), a continuation of the *High Technology Law Journal* effective Volume 11, is edited by the students of the University of California, Berkeley School of Law (Boalt Hall) and is published four times each year (March, July, September, December) by the Regents of the University of California, Berkeley. Periodicals Postage Rate Paid at Berkeley, CA 94704-9998, and at additional mailing offices. POSTMASTER: Send address changes to Journal Publications, University of California, Berkeley Law—Library, LL124 Boalt Hall—South Addition, Berkeley, CA 94720-7210.

Correspondence. Address all correspondence regarding subscriptions, address changes, claims for non-receipt, single copies, advertising, and permission to reprint to Journal Publications, University of California, Berkeley Law—Library, LL 124 Boalt Hall—South Addition, Berkeley, CA 94705-7210; (510) 643-6600; JournalPublications@law.berkeley.edu. *Authors*: see section titled Information for Authors.

Subscriptions. Annual subscriptions are $65.00 for individuals and $85.00 for organizations. Single issues are $30.00. Please allow two months for receipt of the first issue. Payment may be made by check, international money order, or credit card (MasterCard/Visa). Domestic claims for non-receipt of issues should be made within 90 days of the month of publication; overseas claims should be made within 180 days. Thereafter, the regular back issue rate ($30.00) will be charged for replacement. Overseas delivery is not guaranteed.

Form. The text and citations in the *Journal* conform generally to the THE CHICAGO MANUAL OF STYLE (16th ed. 2010) and to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 19th ed. 2010). Please cite this issue of the *Berkeley Technology Law Journal* as 26 BERKELEY TECH. L.J. ___ (2011).

BTLJ ONLINE

The full text and abstracts of many previously published *Berkeley Technology Law Journal* articles can be found at http://www.btlj.org. Our site also contains a cumulative index, general information about the *Journal*, and the Bolt, a collection of short comments and updates about new developments in law and technology written by members of BTLJ.

INFORMATION FOR AUTHORS

The Editorial Board of the *Berkeley Technology Law Journal* invites the submission of unsolicited manuscripts. Submissions may include previously unpublished articles, essays, book reviews, case notes, or comments concerning any aspect of the relationship between technology and the law. If any portion of a manuscript has been previously published, the author should so indicate.

Format. Submissions are accepted in electronic format through the Express*O* online submission system. Authors should include a curriculum vitae and resume when submitting articles, including his or her full name, credentials, degrees earned, academic or professional affiliations, and citations to all previously published legal articles. The Express*O* submission website can be found at http://law.bepress.com/expresso.

Citations. All citations should conform to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 19th ed. 2010).

Copyrighted Material. If a manuscript contains any copyrighted table, chart, graph, illustration, photograph, or more than eight lines of text, the author must obtain written permission from the copyright holder for use of the material.

DONORS

The *Berkeley Technology Law Journal* and the Berkeley Center for Law & Technology acknowledge the following generous donors to Berkeley Law's Law and Technology Program:

Partners

COOLEY LLP

FENWICK & WEST LLP

ORRICK, HERRINGTON & SUTCLIFFE LLP

Benefactors

COVINGTON & BURLING LLP

FISH & RICHARDSON P.C.

KASOWITZ BENSON TORRES & FRIEDMAN LLP

KIRKLAND & ELLIS LLP

LATHAM & WATKINS LLP

MCDERMOTT WILL & EMERY

MORRISON & FOERSTER LLP

SKADDEN, ARPS, SLATE, MEAGHER & FLOM LLP & AFFILIATES

WEIL, GOTSHAL & MANGES LLP

WHITE & CASE LLP

WILMER CUTLER PICKERING HALE AND DORR LLP

WILSON SONSINI GOODRICH & ROSATI

WINSTON & STRAWN LLP

Members

ALSTON + BIRD LLP

BAKER BOTTS LLP

BINGHAM MCCUTCHEN LLP

DEWEY & LEBOEUF LLP

DURIE TANGRI LLP

FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP

GUNDERSON DETTMER STOUGH VILLENEUVE FRANKLIN & HACHIGIAN, LLP

HAYNES AND BOONE, LLP

HICKMAN PALERMO TRUONG BECKER, LLP

KEKER & VAN NEST LLP

KILPATRICK TOWNSEND & STOCKTON LLP

KNOBBE MARTENS OLSON & BEAR LLP

MORGAN, LEWIS & BOCKIUS LLP

MUNGER, TOLLES & OLSON LLP

ROPES & GRAY LLP

SCHWEGMAN LUNDBERG WOESSNER

SIDLEY AUSTIN LLP

VAN PELT, YI & JAMES LLP

WEAVER AUSTIN VILLENEUVE & SAMPSON, LLP

Patrons

BAKER & MCKENZIE

BTLJ ADVISORY BOARD

BOARD OF EDITORS 2011–2012

Executive Committee

Editor-in-Chief
TARAS M. CZEBINIAK

Managing Editor
MICHELLE MA

Senior Articles Editors
AARON MACKEY
MILES PALLEY
ARIELLE SINGH

Senior Executive Editor
WYATT GLYNN

Senior Annual Review Editors
RYAN IWAHASHI
BRITT LOVEJOY

Senior Scholarship Editor
ALEXANDER REICHER

Editorial Board

Submissions Editors
DANIEL KAZHDAN
ZACHARY MARKARIAN

Production Editors
JARAD BROWN
LAUREN SIMS

Bluebook Editors
JILLIAN FEINBERG
CONRAD GOSEN
MIKE SHEEN

Annual Review Editors
ROSS BARBASH
REZA DOKHANCHY

Notes & Comments Editors
COURTNEY BOWMAN
WINNIE HUNG

Symposium Editors
CIARA MITTAN
KILEY WONG

Web Content Editor
MICHAEL SOBOLEV

Information Management Editor
ANDREW FONG

Web & Technology Editor
ANDREA YANKOVSKY

External Relations Editor
ARIANA GREEN

Publishing Editor
NICK WOLOSZCZUK

Member Relations Editor
JEN SPENCER

Articles Editors

LAUREN ESCHER
AMY HAYDEN
KAREN KOPEL
RUBINA KWON
YVONNE LEE
JANE LEVICH
ANGELA MAKABALI
TAYLOR MARGOT
NIKHIL MATANI
HANNAH MINKEVITCH
SONYA PASSI
DAVID ROSEN
JOE SEXTON

BERKELEY CENTER FOR LAW & TECHNOLOGY 2011–2012

Executive Director
ROBERT BARR

Faculty Directors

PETER MENELL
ROBERT MERGES
DEIRDRE MULLIGAN
PAMELA SAMUELSON
PAUL SCHWARTZ
SUZANNE SCOTCHMER
MOLLY VAN HOUWELING

Associate Director
LOUISE LEE

Assistant Director
JULIA TIER

Affiliated Faculty and Scholars

AARON EDLIN
JOSEPH FARRELL
RICHARD GILBERT
BRONWYN HALL
THOMAS JORDE
MICHAEL KATZ
DAVID MOWERY
DAVID NIMMER
DANIEL RUBINFELD
ANNALEE SAXENIAN
JASON SCHULTZ
HOWARD SHELANSKI
CARL SHAPIRO
MARJORIE SHULTZ
LON SOBEL
TALHA SYED
DAVID TEECE
JENNIFER M. URBAN
HAL R. VARIAN
DAVID WINICKOFF

FOREWORD: TECHNOLOGY'S TRANSFORMATION OF THE REGULATORY ENDEAVOR

Kenneth A. Bamberger†

Both the practicalities of governance and our understandings of it have come a long way since the articulation of the insight that code "regulates;"[1] that the choices embedded in technology for a whole variety of reasons (and none at all) have normative implications; and that the computer code of California's Silicon Valley—"West-Coast Code"—operates on a very different logic than the Beltway variety: the "East-Coast Code" of statutes and regulations.[2]

While those insights revolved the lens through which we view policy issues, reality is even more complicated, muddled, and less differentiated than these original important dichotomies suggest. For the technology form of code is not simply an additional mode of regulation; rather, it infuses, grounds, and enables legal regulation and governance itself—just as it does all aspects of our lives. Technology is part and parcel of management and decision making, of action and inaction.

Indeed, regulators have taken to heart the cyberspace lesson that "[i]f code is law . . . 'control of code is power,' "[3] enlisting technological capacity in the pursuit of policy aims. Digital computing, communication, and information management offer tools of extraordinary strength. Technology permits forms of regulation and enforcement and a capacity for both concentration and diffusion of power and authority that have never before existed. It further creates possibilities for governance in contexts heretofore thought ungovernable.

---

† Professor of Law, University of California, Berkeley School of Law.

1. *See* LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 60 (1999) ("How the code regulates . . . [is a] question[ ] that any practice of justice must focus in the age of cyberspace."); *see also* Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 TEX. L. REV. 553, 554 (1998) (noting that technological capabilities and system design choices provide sources of rulemaking).

2. LESSIG, *supra* note 1, at 53.

3. *Id.* at 60 (quoting WILLIAM J. MITCHELL, CITY OF BITS: SPACE, PLACE, AND THE INFOBAHN 112 (1996)).

At the same time, the fusion of technology and regulation introduces normative inputs into governance, creating particular consequences. Those consequences—intended or not, visible or opaque—must be made the subject of searching inquiry, as they implicate foundational assumptions of accountability, fairness, and reliability on one hand, and the effectiveness of governance and its fidelity to rules adopted by democratic and constitutional processes on the other.

Perhaps because of its breadth across substantive contexts, the scope of technology's role as a regulatory instrument and the implications of that role have largely eluded systematic inquiry.[4]

To that end, in March 2011, the Berkeley Center for Law and Technology ("BCLT") convened academics and policymakers to address these issues from a variety of lenses and perspectives. The structure of BCLT's symposium, "Technology: Transforming the Regulatory Endeavor," reflected the belief that answering the big picture question of how technology is transforming the art and science of governance requires both scholarly inquiry that drills down into particular examples and analysis identifying themes that emerge across context, and recognition of the importance of contextual difference.

The symposium panels included discussions of both specific cases and general themes, which are incorporated in the collected essays in this volume. The symposium speakers reflected a rough typology of four distinct ways that technology has transformed the regulatory endeavor: technology's use in (1) *making individualized decisions* about government benefits; (2) *assessing and managing governance risks*; (3) *monitoring regulatory compliance*; and (4) *forcing compliant behavior* by regulated parties.

The first, individualized decision making, was reflected in the work of symposium panelist Danielle Citron.[5] The second, technological risk assessment and management, was explored by a panel that included Nuclear Regulatory Commissioner George Apostolakis, earthquake researcher Patricia Grossi, and legal economist Eric Talley, discussing the use of technology to measure and regulate nuclear, natural disaster, and financial

4. One exception is REGULATING TECHNOLOGIES: LEGAL FUTURES, REGULATORY FRAMES AND TECHNOLOGICAL FIXES (Roger Brownsword & Karen Yeung eds., 2008) (containing papers presented at a 2007 conference on both "Technology as a Regulatory Tool" and "Technology as a Regulatory Target").

5. *See* Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249 (2008).

risks, respectively.[6] While these three panelists focused at the time on the similarities of method and challenge in regulating in complex contexts, the kinship between these three types of risk was underscored just days thereafter by the events following Japan's 2011 Tōhoku earthquake and the consequent damage to the Fukushima nuclear power plant and the regional and national economies.

Examples of the third and fourth categories, regarding the use of technology as a regulatory instrument, are explored in this volume by Molly K. Macauley and Nathan Richardson on environmental monitoring, and by Ira S. Rubinstein on "privacy-by-design."

Macauley and Richardson's *Seeing the Forests and the Trees: Technological and Regulatory Impediments for Global Carbon Monitoring* discusses the way that increased capacity to monitor forests through remote sensing instruments carried on aircraft or satellites can permit the use of forest carbon offsets in climate policy, an as-yet unexplored option in environmental governance.[7] Their work offers important direction for regulators, given the reliance on monitoring capacity implicit in major trends regarding not just environmental law, but regulation and governance more broadly.[8] The move away from command-and-control regulatory mandates to a focus on outcomes undergirds the turn towards "new governance" approaches across the range, including performance-based regulation; market-based or market-mimicking models; and regulatory approaches that "adapt" to the changing situations that monitor defects.[9] In light of institutional impediments to monitoring discussed by Macauley, Richardson, and symposium panelist Eric Biber,[10] the role of technology in monitoring becomes increasingly important.

Rubinstein's *Regulating Privacy by Design*, in turn, explores perhaps the most ambitious use of technology as a regulatory instrument: embedding technology into product design in ways intended to direct behavior towards compliance with regulatory norms.[11] As Rubinstein details, regulators on

6. These themes are, as well, reflected generally in recent work such as Kenneth A. Bamberger, *Technologies of Compliance: Risk and Regulation in a Digital Age*, 88 TEX. L. REV. 669, 714 (2010); Erik F. Gerding, *Code, Crash, and Open Source: The Outsourcing of Financial Regulation to Risk Models and the Global Financial Crisis*, 84 WASH. L. REV. 127, 179 (2009).

7. Molly K. Macauley & Nathan Richardson, *Seeing the Forests and the Trees: Technological and Regulatory Impediments for Global Carbon Monitoring*, 26 BERKELEY TECH. L.J. 1387 (2011).

8. *See generally id.*

9. *See generally* Kenneth A. Bamberger & Deirdre K. Mulligan, *New Governance, Chief Privacy Officers, and the Corporate Management of Information Privacy in the United States: An Initial Inquiry*, 33 LAW & POL'Y 477, 480–82 (2011) (describing "new governance" approaches to regulation).

10. *See* Eric Biber, *The Problem of Environmental Monitoring*, 83 U. COLO. L. REV. 1 (2011).

11. Ira S. Rubinstein, *Regulating Privacy by Design*, 26 BERKELEY TECH. L.J. 1409 (2011).

both sides of the Atlantic have settled on frameworks that encourage the use of both "Privacy Enhancing Technologies" and default settings that favor privacy. These technological solutions often serve as not just complements to but also substitutes for data protection and privacy laws.[12]

Against this background, the remainder of the works in this volume explore thematic questions raised by these technological and regulatory developments. Philosopher Helen Nissenbaum casts the analytic net most broadly in her Keynote Address, asking the basic question: "If technology regulates, why do we need regulation (and vice versa)?"[13] Drawing from a range of work in both law and science and technology studies, Nissenbaum reflects on the difficulties of translating policy prescriptions into code.[14] As she describes, the exercise is "not quite as straightforward as simply plugging values into a technology and then believing that you have immediately had some positive and protracted impact on society."[15] Rather, with reference to both privacy-protecting technology and digital rights management ("DRM") technologies—perhaps the two most developed examples of the use of technology as behavior-forcing regulatory instruments—she posits a number of reasons for the continued salience of legal regulation.[16] The most straightforward might be as a corrective, when "regulation by technology contradicts societal values."[17] Yet even when such a corrective is not needed, Nissenbaum suggests, the coexistence of both law and technology is necessary because it permits the "handoff" between two regulatory systems.[18] Such a handoff not only allows law to compete with technology but also provides an alternate means "to shape how people s[ee], underst[an]d, and interpret[] prevailing" technologies that might otherwise be believed to be simply natural, neutral, or "regular."[19]

From different perspectives, the Notes by Krzysztof Bebenek and April Elliott expand on the themes of interaction between different regulatory tools. In *Strong Wills, Weak Locks: Consumer Expectations and the DMCA*

12. *Id.* at 1410–14.

13. Helen Nissenbaum, *From Preemption to Circumvention: If Technology Regulates, Why Do We Need Regulation (and Vice Versa)?*, 26 BERKELEY TECH. L.J. 1367 (2011). More specifically, Nissenbaum asks, "if technology embodies values, and if technology is capable of regulation, what role is left for law and regulation?" *Id.* at 1368.

14. *See generally id.*

15. *Id.* at 1370.

16. *Id.* at 1374–79.

17. *Id.* at 1374.

18. *Id.* at 1380.

19. *Id.*

Anticircumvention Regime, Bebenek again draws on the DRM context.[20] He cautions against an overemphasis on either law or technology as normatively determinative and suggests the importance of a third "regulator" in governing behavior: consumer norms. The power of these norms, he suggests, not only tempers technology's regulatory effectiveness but also must be considered in shaping the law if legal regulation is to be effective.[21] In turn, Elliott, in *Medicare as Technology Regulator: Medicare Policy's Role in Shaping Technology Use and Access*, considers the phenomenon of legal regulation often ignoring its impact on technology choices, and the implications for policy.[22]

Finally, in *Lost in Translation: Legality, Regulatory Margins, and Technological Management*, Roger Brownsword addresses squarely the implications of the "sea change in the regulatory environment" when "technologies are used to manage conduct in a way that assures a patterned outcome."[23] Specifically, he identifies important governance transformations that occur when legal regulation is replaced by "techno-regulation."[24] Such a substitution, he argues, diminishes regulation's moral component, the traditional notion that salient (legal) constraints embody shared notions of what behavior is "legitimate."[25] This presupposition of regulation as "an inclusive attempt to articulate the community's best interpretation of its moral commitments" is, in turn, replaced by a signal that everything that is (technically) possible is permissible, and vice versa: "if the door will not open without the required biometric confirmation, there is no way in."[26] By this account, the handoff from law to technology shifts regulation's pitch from the "normative . . . to the non-normative register."[27]

In this light, Brownsword joins the other symposium authors in structuring important framing questions for the emerging research agenda in

---

20. Krzysztof Bebenek, Note, *Strong Wills, Weak Locks: Consumer Expectations and the DMCA Anticircumvention Regime*, 26 BERKELEY TECH. L.J. 1457 (2011).

21. *Id.* at 1475–86.

22. April Elliott, Note, *Medicare as Technology Regulator: Medicare Policy's Role in Shaping Technology Use and Access*, 26 BERKELEY TECH. L.J. 1489 (2011).

23. Roger Brownsword, *Lost in Translation: Legality, Regulatory Margins, and Technological Management*, 26 BERKELEY TECH. L.J. 1321, 1323 (2011).

24. *Id.*; Roger Brownsword, *What the World Needs Now: Techno-Regulation, Human Rights and Human Dignity*, *in* 4 GLOBAL GOVERNANCE AND THE QUEST FOR JUSTICE: HUMAN RIGHTS 203 (Roger Brownsword ed., 2004).

25. *See generally* Brownsword, *supra* note 23 (describing how technological regulation can decrease opportunities for community participation in the law's creation and moral self-determination).

26. *Id.* at 1324.

27. *Id.* at 1326.

techno-regulation: the relative capacities of competing regulatory instruments; the ways in which each of those multiple instruments are intertwined, and can deepen, illumine, or undermine the others; and the manner in which fundamental governance values such as regulatory legitimacy—reflecting not only "the purposes pursued by regulators" but also "the means that they use to implement their purposes"[28]—might be translated for the technological age.

28. *Id.* at 1325.

LOST IN TRANSLATION: LEGALITY, REGULATORY MARGINS, AND TECHNOLOGICAL MANAGEMENT

Roger Brownsword[†]

TABLE OF CONTENTS

† Roger Brownsword is a Professor of Law at King's College London, and he was the founding director of TELOS (a newly created research center focusing on technology, ethics, law, and society). He also maintains a long-standing link with the University of Sheffield as an Honorary Professor in Law.

I. INTRODUCTION

The concept of law is contested in many ways. Some jurists argue that law must be understood as an essentially moral enterprise. Others insist on a strict separation of the concepts of law and morals.[1] Some, rather narrowly, identify law with the operations of highly institutionalized legislative assemblies and courts—law, on this view, is hard and high. Others see law everywhere, in the codes and guidance that are associated with much less formal regulation and governance.[2] However, on one point, all protagonists are agreed: whatever our particular conceptual understanding of law, it is a normative phenomenon that we are trying to frame. As formal high law shades into regulation and governance, even into ethics and morals, it remains normative. The enterprise is still one, as Lon Fuller famously expressed it, of seeking to subject human conduct to the governance of rules.[3]

In a time of rapid technological change,[4] how well does our existing conceptual apparatus serve us? Arguably, foundational concepts such as human rights and human dignity represent precisely the intellectual anchoring points that we need to preserve if we are to maintain a critical distance between emergent technologies and what we judge to be their progressive (and regressive) applications and practices.[5] By contrast, some concepts that were crafted in an earlier time—for example, privacy[6] and

1. *See* H.L.A. Hart, *Positivism and the Separation of Law and Morals*, 71 HARV. L. REV. 593 (1958); Lon L. Fuller, *Positivism and Fidelity to Law—A Reply to Professor Hart*, 71 HARV. L. REV. 630 (1958).

2. *See, e.g.*, JAN KLABBERS ET AL., THE CONSTITUTIONALIZATION OF INTERNATIONAL LAW 11 (2009) (examining what a constitutional international legal order could look like); THEORIZING THE GLOBAL LEGAL ORDER (Andrew Halpin & Volker Roeben eds., 2009) (exploring a range of vexed issues concerning global law, legal pluralism, and the judicial role); *see also* Roger Brownsword, *Framers and Problematisers: Getting to Grips with Global Governance*, 1 TRANSNAT'L LEGAL THEORY 287 (2010) (elaborating on the idea of regulatory cosmopolitanism).

3. *See* LON L. FULLER, THE MORALITY OF LAW (1969).

4. *See* PIERRE BALDI, THE SHATTERED SELF (2002) (discussing how technological advancements which manipulate genomes are creating a new concept of what defines humans).

5. *See, e.g.*, ROGER BROWNSWORD, RIGHTS, REGULATION, AND THE TECHNOLOGICAL REVOLUTION (2008) (discussing the challenge affecting regulation of fast developing technological and scientific advancement); Roger Brownsword, *What the World Needs Now: Techno-regulation, Human Rights and Human Dignity*, *in* 4 GLOBAL GOVERNANCE AND THE QUEST FOR JUSTICE: HUMAN RIGHTS 203 (Roger Brownsword ed., 2004) (explaining how the integration of modern technology into globalization creates challenges for the regulatory framework supposed to manage these developments).

6. *See, e.g.*, GRAEME LAURIE, GENETIC PRIVACY (2002); DANIEL J. SOLOVE, UNDERSTANDING PRIVACY (2008); Roger Brownsword, *Consent in Data Protection Law:*

property[7]—seem to need to be re-crafted for our technological age. In this context, the question arises: do we need to rethink our concept of law (and, concomitantly, our valuation of legality and the rule of law) at a time when technology is set to bear in on our root assumption that law is a normative enterprise?

Regulatory theorists have taught us to think of the channeling function of law as having three phases: first, setting the rule or standard; second, monitoring compliance; and, third, correcting for non-compliance.[8] While there is more to the legal enterprise than channeling conduct, as a channeling instrument, law involves direction, detection, and correction. Clearly, technologies of various kinds are already being employed at all three phases of the legal enterprise.[9] However, some of the most debated technologies (particularly CCTV, DNA profiling, RFID implants, and so on) are employed to reinforce the rules and to encourage compliance. Such technological reinforcement amplifies the law's prudential signal (the thought is that, with the likelihood of detection being increased, it is not in one's interest to break the rule), and this might be a significant shift away from whatever moral signals the law otherwise gives. However, even with this drift from the moral, we are still dealing with a normative enterprise.

The sea change in the regulatory environment takes place when technologies are used to manage conduct in a way that assures a patterned outcome. When this happens the enterprise is no longer normative because the environment is controlled so that it is no longer possible to act in certain ways or so that we cannot act otherwise than we do. The signals shift from being prudential (this ought, or ought not, to be done because it is, or is not, in one's interest to do it) or moral (this ought, or ought not, to be done

Privacy, Fair Processing and Confidentiality, *in* REINVENTING DATA PROTECTION? 83 (Serge Gutwirth et al. eds., 2009); Roger Brownsword, *Regulating Brain Imaging: Questions of Privacy and Informed Consent*, *in* I KNOW WHAT YOU ARE THINKING: BRAIN IMAGING AND MENTAL PRIVACY (Sarah J.L. Edwards et al. eds., forthcoming 2012); Helen Nissenbaum, *Privacy as Contextual Integrity*, 79 WASH. L. REV. 119 (2004).

7. *See, e.g.*, JAMES BOYLE, SHAMANS, SOFTWARE, AND SPLEENS (1996) (discussing the problems with determining who owns what in the new technological age); F. Gregory Lastowka & Dan Hunter, *The Laws of the Virtual Worlds*, 92 CALIF. L. REV. 1 (2004) (analyzing whether virtual objects constitute legal property).

8. *See, e.g.*, BRONWEN MORGAN & KAREN YEUNG, AN INTRODUCTION TO LAW AND REGULATION 74–75 (2007).

9. *See* Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1258 (2008) (discussing the use of technology at the stage of rule making); *see also, e.g.*, Isaac B. Rosenberg, *Involuntary Endogenous RFID Compliance Monitoring as a Condition of Federal Supervised Release—Chips Ahoy?*, 10 YALE J.L. & TECH. 331, 333 (2008) (discussing the use of technology at the stage of correction).

because it is, or is not, in line with respecting the *legitimate* interests both of oneself *and of others*) to indicating what is reasonably practicable or possible (or not reasonably practicable or impossible). For example, if the door will not open without the required biometric confirmation, there is no way in.

One of the concepts that seems to be lost in the translation from a traditional legal order to a technologically managed order is normativity—ought and ought not becomes can and cannot. In the latter kind of order, to the extent that the regulatory environment is managed in this way, agents are unable to act on their own judgments of what ought to be done, whether for the sake of self-interest or for the sake of the moral interests of oneself or others. As lawyers, clinging on to the idea of law as a normative enterprise, what should we make of such technological changes to the mechanisms of social ordering? What are the implications of regulating by design?[10]

Jurists might decline to engage with this new world. They might declare that their cognitive interest is limited to law understood as a normative phenomenon. That is, it is fine for others to take an interest in technological management but, as jurists, the question remains to identify the essential nature of the normative enterprise that is law. While it might be defensible to insist that the concept of law should be confined to normative forms of ordering, it surely is not sensible to limit the horizons of juristic inquiry in this way. If law (as a normative enterprise) assumes a shrinking significance in technologically managed regulatory environments, its conceptual relevance seems less obvious. Why should social scientists treat law as an important organizing concept when social order relies less on normative mechanisms? Moreover, if we think that the real interest in law lies less with its normative structure and form and more with its commitment to legality—due process and the like—then jurists need to work on the articulation of these commitments in non-normative regulatory environments.[11]

For jurists who are prepared to engage with the world of "techno-regulation"[12]—not the most attractive of terms, admittedly—the question is whether the increasing reliance on technological regulatory instruments is a cause for concern. Does it matter that there is an amplification of prudential signals; is there a challenge here to moral community? And, when both prudential and moral signals are overtaken by non-normative signals, what

10. For an important engagement with this question, see Karen Yeung, *Towards an Understanding of Regulation by Design*, *in* REGULATING TECHNOLOGIES 79 (Roger Brownsword & Karen Yeung eds., 2008).

11. *See* Mireille Hildebrandt & Bert-Jaap Koops, *The Challenges of Ambient Law and Legal Protection in the Profiling Era*, 73 MOD. L. REV. 428 (2010).

12. Brownsword, *supra* note 5, at 203.

does this signify for the possibility of making and acting on one's own prudential and moral judgments? This leads us to consider: in a non-normative regulatory environment, who is exercising power? Who is in control? Who is accountable? Once we discern the trajectory of technological regulatory tools, we might be so concerned as to favor a highly precautionary approach and to think that the regulatory line must be held against any loss of prudential and moral self-determination. However, if we are prepared to concede that some examples of technological management are appropriate, we must consider what the criteria are for assessing whether such non-normative management is appropriate.

On my understanding, we should conceive of law as an essentially moral enterprise.[13] However, just as importantly, we need to frame our inquiries in a way that both brings in the larger regulatory environment and highlights the importance of regulatory legitimacy.[14] On my reading, although legality does not presuppose *foreground* normative signals (the norms can be in the background), it does presuppose an inclusive attempt to articulate the community's best interpretation of its moral commitments. Regulators have the responsibility to act as stewards for the conditions that make moral community possible (and morally meaningful) and to facilitate the participation of regulatees in setting the terms for the ordering of public life as well as in endorsing the particular regulatory registers and technologies to be employed for such ordering purposes. In other words, regulatory legitimacy is to be tested not only in relation to the purposes pursued by regulators (what they are trying to achieve) but also to the means that they use to implement their purposes (that is, how they regulate). And, what is more, reliance on techno-regulation needs to be open to review, not only in relation to a particular regulatory intervention, but also in the light of the overall balance of normative and non-normative instruments of social ordering.

The Article is in four principal Parts. The first Part sketches the idea of a regulatory environment, drawing out in particular the three key registers (or signals)—namely, moral, prudential, and practicable/possible—that regulators employ. Relative to these three registers, the Article will identify two significant movements associated with the use of technology as a regulatory instrument: first, the movement from the moral to the prudential;

13. I mean this in a strong sense: moral reason is focal for practical reason and hence for both legal and regulatory reason. *See* DERYCK BEYLEVELD & ROGER BROWNSWORD, LAW AS A MORAL JUDGMENT (1986) (arguing that a legal idealist conceptual framework has superior theoretical credentials to that of legal positivism).

14. *See* BROWNSWORD, *supra* note 5, at 10.

and, second, the movement from the normative (whether moral or prudential) to the non-normative register.

The second Part considers the implications of technology being deployed in ways that amplify the prudential signals in the particular regulatory environment. Such amplification might have some impact on those regulatees who tend to reason prudentially; the self-interested reasons for compliance might now outweigh the self-interested reasons for non-compliance. However, there may be unintended side effects, such as the erosion of conditions for moral community. I suggest that developing the concept of a "regulatory margin"[15] would provide a critical doctrinal opening and a benchmark for review of changes to the complexion of the regulatory environment, in response to these side effects.

The third Part examines a technologically managed environment where the regulatory signals are no longer either prudential or moral. When the normative signals are no longer the primary register, this shift appears to reduce both prudential and moral self-determination. Once again, but now in order to maintain the conditions for moral community as well as to preserve space for prudential self-determination, I argue that we must develop the idea of a regulatory margin.

Finally, the fourth Part returns briefly to the question of legality. On any view of law (even hard-nosed legal positivism), intelligent regulation presupposes some engagement by regulators[16] with both the prudential preferences and the moral commitments of their regulatees. If we follow the Fullerian view that law itself is essentially a reciprocal enterprise, then such engagement is necessary in a definitional sense. Such engagement is precisely what needs to be carried across from the old to the new. This engagement

15. Changes in the complexion of the regulatory environment can go unnoticed and unchallenged. The function of the "regulatory margin" is to raise the consciousness of both regulators and regulatees that such changes may be occurring, to give regulatees a way of compelling review of regulatory action by reference to such changes, and to provide a doctrinal space for a jurisprudence to develop that establishes which changes are acceptable and which are not.

16. The terms "regulators" and "regulatees" drip with ambiguity. However, for my purposes, "regulators" are those who put in place the signaling features of the regulatory environment and "regulatees" are those to whom such signals are directed. If we are thinking about a part of the regulatory environment that is dominated by law-like modes of regulation, then the lawmakers are the regulators and the law-subjects are the regulatees. However, this presupposes a rather hierarchical relationship between regulators and regulatees representing just one of several types of regulatory environment. In those environments that are the product of self-regulatory activities (as is the case, for example, with much of the regulation of the Internet), those who act in the capacity of "regulators" are also very obviously "regulatees".

needs to be carried across in a way that enables communities to debate not only particular proposals for the use of techno-regulation but also the bigger picture of the kind of regulatory environment that is constructed. That is to say, there need to be debates not only about regulatory purpose and content but also about the complexion and character of the regulatory environment.

II. THE NATURE OF THE REGULATORY ENVIRONMENT

This Part sketches the salient features of the "regulatory environment." Regulatory environments can be articulated in many different forms; there is no standard pattern. However, this Part highlights the kinds of action-guiding signals that regulators may employ. It does so because, quite simply, the key questions in this Article concern the significance of the kinds of signals that are employed and, concomitantly, the changing complexion of our regulatory environments.

What are we to understand by the concept of "a regulatory environment?" Stated shortly, we should understand it as an action-guiding environment in which regulators direct the conduct of regulatees with a view to achieving a particular regulatory objective. In response to the regulatee's question, "What should I do?", the regulatory environment will signal that particular acts are permitted (even required) or prohibited, that they will be viewed positively, negatively, or neutrally, that they are incentivized or disincentivized, and so on. In technologically-managed regulatory environments, the signals are rather different, indicating whether the performance of a particular act is reasonably practicable or even a possible option.[17] In such a regulatory environment, instead of regulatees asking what they ought to do, their question is, "What can I do?"

Whilst some environments are regulated in a top-down fashion (with regulators clearly distinguishable from regulatees), others are more bottom-up (in the sense that they are self-regulatory). Whereas, in top-down regulatory environments, there is likely to be a significant formal legal presence; in bottom-up self-regulatory environments, this is less likely to be the case (here, as some would have it, it is "governance" that rules). Moreover, while some regulatory environments are reasonably stable and well formed, others are unstable, overlapping, conflicting, and so on.

If we employ this idea of a regulatory environment, then we frame our inquiries in a distinctive way. Crucially, we do not assume that the only regulatory signals are of a formal legal character and nor do we assume that

17. *But see* Roger Brownsword & Han Somsen, *Law, Innovation and Technology: Before We Fast Forward—A Forum for Debate*, 1 LAW INNOVATION & TECH. 1, 4 (2009).

they are necessarily normative. Following Lawrence Lessig's seminal work on the range of regulatory modalities,[18] the first of these assumptions will not be contentious; but it is worth adding a few words in relation to the second of the assumptions. In traditional regulatory environments, both legal and social rules are designed to convey normative signals. Even market signals can speak to what ought (or ought not) to be done, not so much as a matter of respect for others but simply what ought (or ought not) to be done in one's own interest. For example, where a "green" tax is added to the price of larger cars or to fuel, we might reason that we ought to drive a smaller car because larger cars are expensive and put a strain on our personal finances. However, if the price of larger cars is increased beyond our means, our reasoning shifts from the normative mode to the non-normative mode of practicability—it is not so much that we ought not to buy a large car as a matter of self-interest but that we simply cannot (afford to) do so.

When the regulatory modality is that of architecture or code, we might well find that the signal is one of (non-normative) practicability or possibility. However, as with market signals, there might be elements of both normativity and non-normativity—witness, for example, Mireille Hildebrandt's important distinction between "regulative" (normative) and "constitutive" (non-normative) technological features.[19] So, for example, if a car is equipped with sensors that can detect alcohol in the driver, it might be designed to respond normatively (by advising that it is not safe for the driver to proceed) or non-normatively (by immobilizing the car).

To be sure, distinguishing between the way that regulators intend a signal to be understood and the way that (some or all) regulatees actually understand it could problematize this analysis. There might well be some interesting signaling failures. However, for present purposes we can keep things simple by assuming that, in general, regulatees interpret the signal in the way that regulators intended.

Formally, we can say that regulators might attempt to engage the practical reason of their regulatees by using one or more of the following three signaling registers:

> (1) the *moral* register: here regulators signal that some act, *x*, categorically ought or ought not to be done relative to standards of

18. *See* LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 85–100 (1999); Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501, 507–14 (1999).

19. Mireille Hildebrandt, *Legal and Technological Normativity: More (and Less) Than Twin Sisters*, 12 TECHNE: RES. PHIL. & TECH., no. 3, 2008, at 169, *available at* http://works.bepress.com/mireille_hildebrandt/13/.

> right action (as in retributive articulations of the criminal law where the emphasis is on the moral nature of the offence); or
>
> (2) the *prudential* register: here regulators signal that some act, *x*, ought or ought not to be done relative to the prudential interests of regulatees (as in deterrence-driven articulations of the criminal law where the emphasis is on the sanction that will be visited on offenders); or
>
> (3) the register of *practicability* or *possibility*: here regulators signal that it is not reasonably practicable to do some act, *x*, or even that *x* simply cannot be done—in which case, regulatees reason, not that *x* ought not to be done, but that *x* cannot be done (either realistically or literally).

In an exclusively moral environment, the primary normative signal (in the sense of the reason for the norm) is always moral; but the secondary signal, depending upon the nature of the sanction, might be more prudential. In traditional criminal law environments, the signals are more complex. The primary normative signal to regulatees can be either moral (the particular act should not be done because this would be immoral, or the act would be harmful to others) or paternalistically prudential (the act should not be done because it is contrary to the interests of the regulatee). The secondary signal represented by the deterrent threat of punishment, however, is prudential.[20]

As the regulatory environment relies more on technological assistance and management, we can detect two key shifts of emphasis. First, there is a movement from the moral register to the prudential register. We see this, for example, where regulators rely on CCTV, DNA profiling, tracking and monitoring devices, and so on.[21] Here, the strength and significance of the

---

20. Alan Norrie highlights three broad developments in recent British criminal law and justice, namely:

> (i) an increasing emphasis on notions of moral right and wrong and, concomitantly, on individual responsibility ("responsibilisation"); (ii) an increasing emphasis on dangerousness and, concomitantly, on the need for exceptional forms of punishment or control ("dangerousness"); and (iii) an increasing reliance on preventative orders and new forms of control ("regulation"). While the first of these developments is in line with the aspirations of moral community, it is the second and the third that such a community needs to monitor with care. In this light, see, in particular, Lucia Zedner, 'Fixing the Future? The Pre-emptive Turn in Criminal Justice' in McSherry, Norrie, and Bronitt (eds), op cit, 35.

Alan Norrie, *Citizenship, Authoritarianism and the Changing Shape of the Criminal Law*, *in* REGULATING DEVIANCE 13, 20 (Bernadette McSherry et al. eds., 2009).

21. *See* Mark A. Rothstein & Meghan K. Talbott, *The Expanding Use of DNA in Law Enforcement: What Role for Privacy?*, 34 J.L. MED. & ETHICS 153, 160–61 (2006).

moral signal fades as the prudential signal dominates. Second, there is a movement from the normative to the non-normative registers. For example, although some rules and regulations are displayed at international airports (about the rights of passengers if flights are delayed, about not leaving bags unattended, and the like) the regulatory environment is largely architectural and non-normative. The signal that greets passengers in the arrivals hall at the airport is that the only way to board the plane is by following the track that leads from check-in to the boarding gate and that, along the way, passes through security that involves ever more intrusive scanning of person and property.[22] In an environment in which technology and physical architecture regulate, moral and prudential signals drop out of sight to be replaced by signals and structures that—the shopping area apart—leave the passenger with little room for either moral or prudential maneuver. Thus the question for regulatees becomes not what ought to be done but only what can and cannot be done.

In what follows, the Article considers the significance of two critical movements in the character or complexion of the regulatory environment: first, when there is a shift from the moral to the prudential register; and then when there is a rise of non-normative techno-regulation.

III. THE FIRST MOVEMENT: TECHNOLOGIES THAT AMPLIFY PRUDENTIAL SIGNALS

This Part sketches answers to the following two questions. First, should regulators be concerned that there is a movement in the regulatory environment from moral to prudential signals? Second, should they exercise restraint in resorting to new regulatory technologies that serve to amplify prudential signals?

In this context, a reasonable opening question for regulators would be to ask what impact the use of CCTV, DNA profiling, lie-detection technologies, and the like might have on individual decision-making. Is the increase in prudential noise interfering with the ability of agents to try to act morally? This, however, examines only a slice of life in a moral community and regulators would not act responsibly unless they also asked whether the amplification of prudential signals was damaging to moral community more generally.

22. *See* Bert-Jaap Koops, *Technology and the Crime Society: Rethinking Legal Protection*, 1 LAW INNOVATION & TECH. 93 (2009) (discussing how technology facilitates greater criminalization via regulation and constant surveillance).

This Part then introduces the idea of a moral regulatory margin (and, concomitantly, of marginal considerations) that might focus minds on the maintenance of moral community. Finally, this Part offer a radical re-reading of the issues raised by the *Marper* case[23] (in which there was a human rights challenge to the legal provisions in England and Wales authorizing the taking and retention of DNA samples and profiles for criminal justice purposes) to underline the full extent of the responsibilities of regulators.

Two other points should be noted. First, moral philosophers must contend with hypothetical amoralists who, having no interest in or inclination towards doing the right thing, are liable to spoil the party. However, for the purposes of our discussion, this Article side-steps amoralism[24] and assumes a community with moral aspirations. To be sure, this is not to imply that amoralism can be side-stepped in all contexts, particularly where the coherence of moral aspirations is challenged. However, to the extent that there are communities with such aspirations (as, of course, there are), amoralism is irrelevant to the question of whether any tuning down of the moral regulatory signals is significant for such communities.

Secondly, this Article will assume that, within such a community, it is recognized that sovereign regulators have a responsibility, *inter alia*, to act as stewards for the conditions that make it possible to function as a moral community.

A. THE IMPACT OF TECHNOLOGICAL REGULATION ON INDIVIDUAL DECISION-MAKING

In this Part of the Article, the focal question is whether the use of regulatory technologies that amplify prudential signals comes at any cost to moral community. This Part proposes a possible litmus test to ascertain how such a change in the regulatory environment impacts the (morally aspirant) reasons and actions of individual agents.

---

23. *See* R v. Chief Constable of S. Yorkshire Police (*ex parte* LS & Marper), [2004] UKHL 39 (appeal taken from Eng.), *available at* http://www.publications.parliament.uk/pa/ld200304/ldjudgmt/jd040722/york-1.htm (holding that it is lawful, in England and Wales, for the police to retain the DNA samples and profiles of persons who are arrested but who are *not* convicted of an offense); *see also* S & Marper v. United Kingdom (*Marper*), app. nos. 30562/04 & 30566/04, (2009) 48 E.H.R.R. 50 (Eur. Ct. H.R. Dec. 4, 2008), 2008 WL 5044408.

24. For such amoralists, it is only one's own needs and preferences that matter; the only relevant interest is self-interest; the only signals that count are those that are prudential; and, for such agents, the fading of the moral register would be immaterial. For a community of amoralists (if this is not a contradiction in terms), the amplification of prudential signals might be a cause of some concern, but not because it corrodes or challenges the possibility of *moral* community.

To pursue such an inquiry, we might develop four ideal-typical agents as follows:

> Type 1 agents who act only and always on *moral* reasons;
>
> Type 2 agents who act only and always on *prudential* reasons;
>
> Type 3 agents who act on a mix of *moral and prudential* reasons; and
>
> Type 4 agents who are *erratic*, sometimes acting on moral reasons, sometimes on prudential reasons, and sometimes on mixed reasons.

If the prudential signal is amplified, how does this affect the way that individual agents reason and act? For example, if speed cameras are fixed to a section of highway, either monitoring the speed of vehicles at a particular point or their average speed over a longer distance, how does this affect motorists? On the face of it, the presence of cameras reinforces the rules of the road and signals to motorists that, if they do not observe the speed limits, they will be detected. But, how do motorists respond to this amplification of the prudential signal? Amongst criminologists, it is trite that, generally speaking, prudential calculation is more responsive to an increased likelihood of detection than to an increase in the penalties for the particular offense.[25] However, the responses of individuals are not uniform and, in the case of speed cameras, research suggests that (not surprisingly) the responses of motorists vary.[26] For whatever reason, some motorists always observe the speed limits, irrespective of whether they are driving through areas covered by speed cameras. Others slow down, sometimes to accelerate again as they exit the controlled area. Undoubtedly, still others exceed the speed limit, taking little or no notice of the presence of cameras.

For many researchers, the question will be simply whether the use of some particular regulatory technology (such as speed cameras or CCTV) "works"—namely, whether it is effective in assisting the regulators' purposes. That is, how would the amplification of prudential signals impact such agents? For Type 1 agents, unless moral reason offers some optionality (as where various actions are morally permissible), prudential signals, whether amplified or not, are irrelevant. For such agents, prudential considerations only operate within the interstices of moral reason. For Type 2 agents (whose

25. JOHANNES ANDENAES, PUNISHMENT AND DETERRENCE 960 (1974) ("Even the simplest kind of common sense indicates that the degree of risk of detection and conviction is of paramount importance to the preventive effects of the penal law. Very few people would violate the law if there were a policeman on every doorstep.").

26. *See* Claire Corbett & Isabel Caramlau, *Gender Difference in Response to Speed Cameras: Typology Findings and Implications for Road Safety*, 4 CRIMINOLOGY & CRIM. JUST.: INT'L J. 411 (2006).

prudential mind-set will be treated as pathological in an aspirant moral community), the amplification of prudential signals will not change the general way that they reason but, in some cases, it might alter their conduct. For example, motorists who reason in this prudential way might slow down on a road when speed cameras are introduced, reasoning that the introduction of cameras tips the balance of self-interested considerations towards compliance. However, while this would be relevant to understanding the effectiveness of particular regulatory technologies, it would not speak to concerns about damage to moral community.

The remaining categories, Type 3 and Type 4 agents, are probably characteristic of many agents in an aspirant moral community. Here, there are highly relevant questions about the impact of the amplification of prudential signals. Does this change in the regulatory environment affect the way in which these agents reason, reducing the occasions when they reason morally? And, does it interfere with them acting on moral reasons? In principle, there could be some significant alterations in both the reasoning and the conduct of Type 3 and 4 agents. Finally, the potential presence of serious concern about such alterations requires an inquiry to establish their prevalence and their significance.

That said, some might think there is little risk in the amplification of prudential signals. After all, existing criminal justice regimes employ a mix of moral and prudential signals. The prudential signals are regularly tuned up or tuned down by changes in penalties, by targeting particular offences, and so on, and we detect no obvious change in moral community. Moreover, if the amplification of prudential signals does change the conduct of Type 2 agents so that they cause less harm to the morally protected interests of other agents, there is an element of moral gain without any offsetting loss—at least when assuming that Type 2 agents are incorrigible prudentialists.

In the cases of Type 3 and 4 agents, the moral trade-off is more complex. As with Type 2 agents, the amplification of prudential signals might lead to a reduction in the harm caused to the protected moral interests of others. However if a switch from moral to prudential reason also occurs in the thinking of these agents (even though their conduct is unaltered) this suggests some corrosion of moral community. That is, even though these agents might do what is generally thought to be the right thing, they now do so for prudential rather than moral reasons.

Without further inquiry, we cannot be confident about the impact of the amplification of prudential signals that comes with an increased reliance on some regulatory technologies. So long as such technologies operate at the fringes of a traditional criminal justice system, there is probably little, if any, overall cost to moral community. However, where the regulatory

environment features pervasive surveillance and monitoring technologies, aspirant moral communities should not be so complacent. In a panopticon environment, how likely is it that moral reason will survive, let alone flourish?

Recently, Beatrice von Silva-Tarouca Larsen[27] has suggested that the "general public [might have] not quite woken up to the potential dangers of CCTV."[28] Her principal concern relates to the loss of anonymity (and privacy) in public places. However, putting her finger on precisely the point that is central to this paper, she says:

> Another reason speaks against pervasive recording in public space as a strategy for crime prevention. Increasing the threat of punishment does not deprive punishment of its moral message, and highlighting the detection risk of offending does not have to dilute the deontological condemnation expressed in punishment. *Nevertheless, one should not rule out the possibility that an over-reliance on CCTV, with its emphasis on the instrumental appeal to desist from crime in order to avoid paying the cost, might entail a dilution of the moral reasons for desistence.* This could become a problem, for it is not possible to record and monitor people all the time. It is important that policy makers realise that CCTV can only ever be a small part of the solution for enforcing the criminal law, and that instrumental obedience is no substitute for moral endorsement of criminal prohibitions. Strengthening, communicating and convincing people of the normative reasons for desistence should always remain a priority.[29]

Accordingly, her recommendation is "that policy makers should opt for very selective implementation of public CCTV, within a narrow setting, targeted on particular crimes and a particular type of offender."[30] While such implementation might render CCTV coverage more effective in preventing and detecting crime, this is not really the point. Rather, as Larsen concludes: "Above all, it is important to remember that surveillance can never be a substitute for frontline crime-prevention work in and with the community, for the normative legitimacy of criminal prohibitions *and the moral incentive to abstain from harming others*."[31]

Clearly, to address the question of the significance of amplified prudential signals, we need to think beyond the impact on individual agents

27. BEATRICE VON SILVA-TAROUCA LARSEN, SETTING THE WATCH: PRIVACY AND THE ETHICS OF CCTV SURVEILLANCE (2011).

28. *Id.* at 83.

29. *Id.* at 153–54 (emphasis added).

30. *Id.* at 186.

31. *Id.* (emphasis added).

who are already members of a morally aspirant community and to remind ourselves about the project of moral community.

B. MORAL COMMUNITY: THE PROJECT

This Section presents a thumbnail sketch of what constitutes the project of "moral community." As a project, the emphasis is on process rather than product. It is about how the community organizes its moral deliberations. Moreover, because moral community is being treated as a generic concept, the qualifying condition is that regulators and regulatees are focused on trying to do the right thing, not that they subscribe to a particular school of substantive morality.

There is a distinction between the project of moral community in a generic sense and particular articulations of moral community. The organizing idea for the project is that the community and its members should endeavor to do the right thing relative to the legitimate interests of themselves and others. What counts as a *legitimate* interest, and who counts as an *other*, are deeply contested matters. The way in which these questions are answered will determine how a particular moral community is articulated. So, for example, if we treat the avoidance of pain and distress as the key *legitimate* interest of others, and if we treat *others* as those who are capable of experiencing pain and distress, then the community will articulate along negative utilitarian lines. If we treat an agent's freedom and well being as the relevant *legitimate* interest of others, and if we treat *others* as those who are capable of acting in a purposive way, then the community will articulate along liberal rights-based lines. If we treat human dignity as the key *legitimate* interest, and if we treat all humans as relevant *others*, then the community will articulate as some version of dignitarianism, and so on.[32] These examples could be multiplied many times. However, the point is that these many different articulations are all examples of moral community in the generic sense; and they are all such examples because they start with a commitment to try to do the right thing relative to the legitimate interests of others.

In such an aspirant moral community, the regulatory environment should declare the community's commitment to doing the right thing and it should express its understanding of the guiding principles. At some times and in some places, the process of articulating the community's moral commitments might have been left to an elite group (of philosopher kings or wise men). In that scenario, the commitments so articulated might have been seen as a durable statement (in a world of little change) and the substantive principles

32. *See* Roger Brownsword, *Bioethics Today, Bioethics Tomorrow: Stem Cell Research and the 'Dignitarian Alliance*,*'* 17 NOTRE DAME J.L. ETHICS & PUB. POL'Y 15, 18–19 (2003).

articulated might have been viewed with epistemic certainty. However, the project of moral community as I view it for the twenty-first century is rather different: it is inclusive, constantly under review, and undertaken with a degree of uncertainty.

To start with *inclusiveness*, the project of moral community implies that all voices should be heard, with comprehensive public engagement. This means that, in principle, all members of the community should be able to participate in debates about how the regulatory environment should be articulated if it is to keep faith with the ideal of doing the right thing. On some matters, members of the community might be agreed; and, in all probability, the higher the level of generality at which governing principles are formulated, the easier it will be to agree that these are relevant principles for the guidance of agents who wish to do the right thing. However, there will be many matters that are disputed. Even if the most fundamental of principles are agreed upon, there might be disagreement about the scope and application of a principle in a particular case, about prioritizing competing principles, about where to draw the line between those who are relevant others and those who are not, and so on.[33] So far as is practicable, inclusive deliberations about such matters must occur. Once a decision has been made, a moral community must treat it as provisional and open to *review*.[34] That is, the fact that the balance of argument has favored a particular decision today does not secure it in perpetuity. A moral community must leave open the possibility of revisiting, reviewing, and renewing its decisions. Finally, unless the community claims moral omniscience—which, in the twenty-first century, is hardly a plausible position—it must regard its articulated principles with a degree of epistemic *uncertainty*. This does not have to unravel the project, but it does mean that the current articulation cannot be treated as being set in stone.

To the extent that the public life of such a community focuses on constructing an appropriate regulatory environment, it follows that we cannot assess the impact of an amplification of prudential signals simply by checking the way that regulatees reason and respond to such signals. For, as members of the community, regulatees have a role to play in debating the

33. *See* Roger Brownsword, *Regulating the Life Sciences, Pluralism, and the Limits of Deliberative Democracy*, 22 SING. ACAD. L.J. 801, 803 (2010).

34. *See id.* at 829 ("[W]e cannot regulate in a way that is compatible with all views but a regulatory position needs to be taken; there will be an opportunity to revisit the issue; but, in the interim, we ask regulatees to respect the position that has been taken."); Roger Brownsword & Jonothan J Earnshaw, *The Ethics of Screening for Abdominal Aortic Aneurysm in Men*, 36 J. MED. ETHICS 827 (2010) (emphasizing that the decision to introduce, or not to introduce, a publicly-funded screening program should be reviewable).

regulatory purposes and agreeing the public rules and standards. In other words, before we set aside any concerns about the amplification of prudential signals, we need to check not only whether there is an impact on regulatees at the point of compliance but also on their ability to participate as members of the political (and aspirant moral) community. However, to do this, members must have the capacity to engage in moral discourse and debate—which is to say, there must be no impairment of their moral development.

Taking stock, we can say that the project of moral community (whatever its particular articulation) presupposes that its members will participate in debating the community's best understanding of its moral commitments, in setting public standards that are compatible with those commitments, and in responding to those standards as regulatees who strive to do the right thing for the right reason.[35] Unless the amplification of prudential signals has no effect on any part of the project, regulators (as stewards for moral community) should proceed with care.

C. THE MORAL MARGIN AND MARGINAL CONSIDERATIONS

If regulators are to act as stewards for moral community, they might interpret this responsibility in a weak or strong sense. In a weak sense, the responsibility is to ensure that the moral life of the community is not altogether extinguished; in a strong sense, the role of regulators is to ensure that there is, at worst, no reduction in the moral life of the community and, at best, some promotion of moral community. In the weak context, regulators would not be concerned that the amplification of prudential signals encroached on and reduced the space for moral reason, provided that there was (in the spirit of the Lockean proviso) still sufficient and plenty[36] of opportunity for the moral life. In contrast, in the strong context, such encroachment and reduction would be unacceptable. Whilst the former evokes a community that is trying to preserve something of its moral project, the latter fits with a community that sees itself on a trajectory toward the completion of its moral project. In the light of previous comments about the

35. There are also questions for any aspirant moral community about its relationship with and responsibilities towards other communities. *See, e.g.*, NUFFIELD COUNCIL ON BIOETHICS, GENETICALLY MODIFIED CROPS: ETHICAL AND SOCIAL ISSUES 67 (1999); GOV'T OFFICE FOR SCI., FORESIGHT, THE FUTURE OF FOOD AND FARMING: FINAL PROJECT REPORT 9–10 (2011), *available at* http://www.bis.gov.uk/assets/bispartners/foresight/docs/food-and-farming/11-546-future-of-food-and-farming-report.pdf.

36. This is an allusion to the famous proviso entered by John Locke in his *Second Treatise on Government*. JOHN LOCKE, SECOND TREATISE ON GOVERNMENT ch. V, § 27 (London: Dent, reprinted 1975) (1690). Stated shortly, Locke allows for the appropriation (and enclosure) of land by improvement provided that there is still "enough, and as good, left in common for others." *Id.*

inclusiveness of moral community, the community as a whole should debate whether it defines itself as undertaking the weak or the strong version of the moral project. To this extent, it might be appropriate to characterize the community as being "communitarian" in the sense that its members identify with the kind of moral project that they have committed to undertake.

Whether the community's aspiration is to retain some part of its moral life or to push forward towards a more complete moral life, there needs to be some kind of regulatory margin that serves as a benchmark for decisions involving the use of technologies that amplify prudential signals. On the weak interpretation of stewardship, the margin represents a minimal zone for moral life to be protected at all costs; on the strong interpretation, the margin will mark the present level of moral life. The function of this margin would be twofold: first, it would serve as *ex ante* guidance for regulators (the marginal question would be one that they should ask themselves). Second, it would serve as a focus for *ex post* review.

What might be the relevant marginal considerations? Despite having used locutions such as the "level of moral life," and the "reduction" of moral community, these are not quantifiable matters. There is no moral barometer of this kind. To be sure, there might be some snapshots of the way in which the amplification of prudential signals impedes or interferes with the opportunities for moral action. But, in general, it is hard to conceive of the existence of reliable and regular quantitative measures that would be workable for either regulators or reviewers. Instead, regulators might be guided by two critical considerations. One consideration is whether there is any possibility that the amplification of prudential signals might interfere with regulatees' development of moral reason and the capacity to participate in the life of the community as moral agents. No doubt, the foreground regulatory environment for children and young persons is that found in the family, at school, and in the neighborhood. We should not assume that the larger public regulatory environment aligns with these most proximate environments. Nevertheless, regulators need to be sensitive to the possibility that the amplification of prudential signals in the background environment might carry over to the foreground. A second consideration concerns the importance of the moral interest served by prudential amplification. For example, if amplified prudential signals serve to protect essential infrastructural conditions for the community or to prevent life-threatening harm, this might be seen overall as an acceptable measure—and, of course, it would be much easier to justify such measures where the weak interpretation of moral stewardship is invoked. Clearly, there is a considerable jurisprudence waiting to be developed here, but it will not get underway unless there is an appropriate doctrinal and institutional opening.

D. *MARPER* RE-INTERPRETED IN TERMS OF THE MORAL MARGIN

In this Part of the Article, I earth some of the foregoing argument and analysis in the leading European case on the compatibility of DNA databases with basic human rights, particularly with the right to privacy. While the case is a rich resource for legal and moral argument about the scope and weight of privacy (against the competing objectives of crime control), it does *not* speak at all to the concerns that this Article raises about the complexion of the regulatory environment. And, this is precisely the point: if we wish to raise such concerns, even in a court that has conspicuous moral aspirations, we do not have the doctrinal means to do so. Moreover, in the absence of a doctrine such as the regulatory margin, the real danger is not just that concerns about the complexion of the regulatory environment might be seen but not heard, but that they are not even seen at all.

There is a considerable distance between the kind of review a community might undertake relative to a regulatory margin of the kind just sketched and what happens in current reviewing practice. Or, at any rate, there is considerable distance in those kinds of practices where legal proceedings test the compatibility of regulatory technologies relative to fundamental human rights commitments. In the jurisprudence of the European Court of Human Rights, the leading case of this kind is *S v. United Kingdom*.[37] It is instructive to see how the Court presented the issue in *Marper* and how it might have done so if the question had concerned compatibility, not with human rights, but with the regulatory margin.

Stated shortly, the question in *Marper* was whether the legislation in England and Wales that permitted the taking and retaining of DNA samples from persons who were arrested, and the making and retaining of DNA profiles, was compatible with the right to private and family life that is protected by Article 8(1) of the European Convention on Human Rights.[38] The legislation authorized the taking of a sample from almost all persons who were arrested and, even more controversially, it permitted the retention of both samples and profiles regardless of whether the person who had been arrested was charged, brought to court, or convicted.[39] Very quickly, the collection of samples and profiles grew to become the largest per capita national DNA database. On the positive side, some headline-catching stories highlighted the relevance of DNA evidence in both exculpating innocent

37. *Marper*, app. nos. 30562/04 & 30566/04, (2009) 48 E.H.R.R. 50 (Eur. Ct. H.R. Dec. 4, 2008), 2008 WL 5044408, at *1169 (combining the applications of S, a British minor, and Michael Marper, a British national).

38. *Id.* at *1187.

39. *Id.* at *1176–80.

persons and leading the police to some serious offenders (often where the offence was old and the case was cold).[40] However, on the negative side, it could be objected that the database contained many samples and profiles from persons who had not actually been convicted of any criminal offense—such persons might have been arrested but they surely were to be treated as innocent.[41]

In the ensuing litigation the applicants complained that the authorizing legislation was not compatible with the Article 8(1) privacy right. The defense was that, even if the privacy right were engaged (which was not conceded), the regulatory objectives (with regard to deterring and detecting crime) were overriding public interest reasons within the meaning of Article 8(2) of the Convention.[42] In the domestic courts, the complainants received little encouragement.[43] For, while it was rather grudgingly accepted that the privacy right was engaged,[44] there was no hesitation in finding that the Article 8(2) reasons were compelling. By contrast, at Strasbourg, the Grand Chamber (the full court), found that privacy was not only clearly engaged, but that the extent of the infringement was disproportionate to the criminal justice objectives. Concluding, the court found that

> the blanket and indiscriminate nature of the powers of retention of the fingerprints, cellular samples and DNA profiles of persons suspected but not convicted of offences, as applied in the case of the present applicants, fails to strike a fair balance between the competing public and private interests and that the respondent

40. *See, e.g.*, *id.* at *1174 ("Lord Steyn noted that the value of retained fingerprints and samples taken from suspects was considerable. He gave the example of a case in 1999, in which DNA information from the perpetrator of a crime was matched with that of 'I' in a search of the national database."); *see also* Andrew Norfolk, *Shoe Rapist Is Trapped by Sister's DNA 20 Years After Serial Attacks*, TIMES (London), July 18, 2006, at 3.

41. *See, e.g.*, NUFFIELD COUNCIL ON BIOETHICS, THE FORENSIC USE OF BIOINFORMATION: ETHICAL ISSUES 18 (2007).

42. *Marper*, 2008 WL 5044408, at *1193–96.

43. R v. Chief Constable of S. Yorkshire Police (*ex parte* LS & Marper), [2004] UKHL 39 (appeal taken from Eng.), *available at* http://www.publications.parliament.uk/pa/ld200304/ldjudgmt/jd040722/york-1.htm.

44. On this point, Baronness Hale was alone in finding that privacy was clearly engaged:

> It could be said that the samples are not "information" . . . . But the only reason that they are taken or kept is for the information which they contain. They are not kept for their intrinsic value as mouth swabs, hairs or whatever. They are kept because they contain the individual's unique genetic code within them. They are kept as information about that person and nothing else. Fingerprints and profiles are undoubtedly information. The same privacy principles should apply to all three.

Id. at [70].

> State has overstepped any acceptable margin of appreciation in this regard. Accordingly, the retention at issue constitutes a disproportionate interference with the applicants' right to respect for private life and cannot be regarded as necessary in a democratic society.[45]

To arrive at this judgment, the court was significantly influenced by the fact that the U.K. was an outlier relative to the position taken by other Contracting States. Thus,

> most of the contracting states allow these materials [i.e., DNA samples] to be taken in criminal proceedings only from individuals suspected of having committed offences of a certain minimum gravity. In the great majority of the contracting states with functioning DNA databases, samples and DNA profiles derived from those samples are required to be removed or destroyed either immediately or within a certain limited time after acquittal or discharge.[46]

Whereas, then, the domestic courts judged that the legal powers were not disproportionate given the weak (as they saw it) engagement of privacy and the strong claims of criminal justice, the Grand Chamber judged that the legal powers were disproportionate given the clear engagement of privacy and the much less sweeping powers adopted by other members of the Strasbourg human rights club.[47]

Whichever view we find convincing, imagine that the complaint had been framed differently. Imagine a complaint that the use of DNA evidence (in conjunction with a raft of other modern technologies) amplified prudential signals at the cost of moral community. However, given that the Convention does not invite or recognise such a strategic complaint, how could the point be put to the legal test? Seemingly, the success of such a claim would require a (moral) regulatory margin relative to which the objection could be assessed. That is, if other Contracting States limited the use of DNA profiling to the most serious criminal offenses, they may appear respectful of privacy. But, arguably, they also would be more sensitive to the need to maintain a moral margin that protects moral signals from being overwhelmed by prudential

45. *Marper*, 2008 WL 5044408, at *1202.

46. *Id.* at *1199.

47. *See id.*; *see also* Wood v. Comm'r of Police, [2009] EWCA (Civ) 414, [2009] All E.R. 4 [951] (Eng.) (holding that on the particular facts, the police action in taking and retaining photographs of the entirely innocent complainant was a disproportionate infringement of his Article 8(1) privacy rights). Looking at the bigger picture, Lord Collins remarked that it was "plain that the last word has yet to be said on the implications for civil liberties of the taking and retention of images in the modern surveillance society." *Id.* at [100].

ones.[48] Such a reframing would transform the terms of the complaint from an unreasonable infringement on privacy to an encroachment on the moral margin. The outcome of the case would be the same but the reasoning would be quite different.

There is, of course, a great deal more one could say about the issues raised by *Marper*, not to mention the more general issues raised by personal genetic profiling. On the one hand, moral communities with commitments to human rights will be heartened by the increased concern for privacy now being shown by the U.K. coalition government as well as by the domestic courts.[49] On the other hand, modern technologies (from social networking platforms to brain imaging) constantly push back the boundary at which our expectation of privacy seems reasonable.[50] However, while these are precisely the kinds of issues that need to be addressed and debated in communities that have moral aspirations, the principal questions in this Article relate to the changing complexion of the regulatory environment, and particularly to the significance of those changes for the employment of moral and prudential (normative) reason.

---

48. This might chime in with the court's sentiment in *Marper* that those who are in the vanguard of using technological instruments to prevent and detect crime bear a special responsibility. The *Marper* Court said:

> The Court observes that the protection afforded by art.8 of the Convention would be unacceptably weakened if the use of modern scientific techniques in the criminal-justice system were allowed at any cost and without carefully balancing the potential benefits of the extensive use of such techniques against important private-life interests. In the Court's view, the strong consensus existing among the contracting states in this respect is of considerable importance and narrows the margin of appreciation left to the respondent State in the assessment of the permissible limits of the interference with private life in this sphere. The Court considers that any state claiming a pioneer role in the development of new technologies bears special responsibility for striking the right balance in this regard.

Marper, 2008 WL 5044408, at *1199–1200.

49. For the coalition government's new approach to DNA profiling, see the Protection of Freedoms Bill, 2010-12, H.C. Bill [189] (Gr. Brit.); and for a marked tilt towards privacy in the courts, see *Wood*, [2009] EWCA (Civ) 414, at [12] (concerning the taking and retention of photographs for forensic purposes).

50. *See, e.g.*, Bert-Jaap Koops & Ronald Leenes, *'Code' and the Slow Erosion of Privacy*, 12 MICH. TELECOMM. & TECH. L. REV. 115, 118–19 (2005).

IV. THE SECOND MOVEMENT: WHEN NORMATIVE SIGNALS FADE

This Part turns to the second shift in the complexion of the regulatory environment. This is the change that occurs when regulators rely on non-normative strategies, so that regulatees are presented, not with rules and regulations, with oughts and ought nots, but with the brute fact that some things can be done and others cannot. Whether we view such a regulatory environment from a moral or a prudential perspective, this is a very different place to be. Whether we reason morally or prudentially, the impact of non-normative regulation is that we lose some degree of choice and, with that, some degree of control.

When regulators rely on technological instruments, they can go beyond the amplification of prudential signals to design in a desired pattern of conduct or to design out conduct that is not desired. Sometimes, the technology will not replace normative signals and the regulatory environment, although employing technologies, will still speak to the moral and prudential interests of regulatees. However, at other times the technology might go a step further and replace any normative signals with an entirely non-normative register. The impact of non-normative regulation is that regulatees lose some degree of control and choice, specifically in relation to their prudential and moral reasoning and action.

Such techno-regulatory strategies might focus on products, places, or persons.[51] For instance, regulators might specify certain safety, privacy-enhancing, or copyright-protecting features to be designed into products. Or they might specify certain architectural features to improve safety (as in the layout of roads) or to facilitate transparency (think about the Bundestag building in Berlin) or adversarial political debate (think about the layout of the House of Commons at Westminster). With regard to persons in some future world, regulators might specify that only those human embryos with acceptable genetic profiles should be implanted for reproductive purposes.

These various possibilities for regulation through technology prompt a number of initial questions. For example, does it matter whether the locus for techno-regulatory interventions is in products, in places, or in persons? Are some locations more suitable than others if the technology is to be customized for individual use? Non-normative regulation can lie anywhere on the spectrum between what is not reasonably practicable and what is impossible. Does it make any difference whether the technology operates at

51. *See* Roger Brownsword, *Code, Control, and Choice: Why East Is East and West Is West*, 25 LEGAL STUD. 1, 12 (2005).

the not reasonably practicable or the impossible end of the spectrum? In normative regulatory settings, the focus of the intervention can be at the point of standard setting or monitoring and detection or correction. Does this carry over to non-normative settings? If so, does it matter at which point or points regulators rely upon the technology?

If the variables in the non-normative forms of regulation are significant, that significance stems from the particular way that those variables impact prudential and moral action. Accordingly, we can start by reviewing the significance of various types of non-normative regulatory interventions with regard to prudential reason and the pursuit of self-interest. Then we can return to the implications of those interventions for the project of building moral community.

A. PRUDENTIAL INTERESTS AS A STARTING POINT

To reason prudentially is to make a judgment about one's own interests. Sometimes those judgments will prioritize short-term costs and benefits. At other times, we will act more strategically, taking the longer view (and, as some would have it, acting on an enlightened understanding of where our interest lies). When others make judgments as to what is in our best interests, they try to simulate the prudential judgment that we would make. However, this is not only an imperfect process, but it also deprives us of making our own prudential judgments. Where our powers of prudential reasoning have not yet developed or where they have waned, there might be no alternative other than to have others make prudential judgments on our behalf. However, this is merely a second best; and, in most communities, members will want to make their own prudential judgments because they believe that if anyone is to judge what is in their best interest, it should be themselves. At all events, for the purposes of this part of the discussion, we will presuppose that regulatees value making their own prudential judgments and that they do not normally welcome such judgments being made for them by others.

Where the regulatory environment employs non-normative technologies, it is less clear that regulation is a three-phase process of direction, detection, and correction.[52] For, in such an environment, there is no phase of normative standard-setting, no need to monitor for compliance with a normative signal, and no need to correct for failure to comply with a normative signal. All three phases are collapsed. Nevertheless, as we shall see, there can and should be a preliminary phase in which normative discourse survives to determine which regulatory purposes should be pursued and how they

52. This so-called "cybernetic" approach is pervasive in the regulatory literature. *See, e.g.*, MORGAN & YEUNG, *supra* note 8, at 3, 73, 103.

should be pursued. Worryingly, the survival of normative discourse might be restricted to a regulatory elite. However, in any community that has democratic commitments, the survival of the discourse needs to be community-wide, encompassing and enfranchising both regulators and regulatees.

We can assess the impact of non-normative technologies on prudential reason in four steps: first, by considering how such technologies might operate where an agent self-regulates (and prudentially elects a certain level of technological regulation); second, by reviewing the imposition of non-normative regulatory environments in products and places; third, by thinking about the significance of in-person technological management; and, finally, by revisiting the idea of a regulatory margin.

1. The Impact of Non-normative Technologies on Self-Regulation

As consumers, we constantly express our prudential preferences. As consumer products become more technologically sophisticated, we have the opportunity to express our prudential preferences relative to ever more features of product design. This Section considers the significance of consumers selecting products that incorporate elements of non-normative management of the user's conduct.

Where products are mass-produced and where their technological features are cheap and cheerful, this might not be what, other things being equal, the particular user would choose; individual tastes and preferences vary enormously. In principle, however, we can imagine that products might be designed in ways that allowed for some tailoring to user preferences. Indeed, at the more expensive end of the product market, we would expect an increase in customization opportunities, such that products become more aligned with the preferences of their users. The more this happens, the greater the options available to users and the greater scope there is for fine-grained prudential choice.

Now, the product features in question might be expressly presented as being of a regulatory nature. That is to say, once incorporated, these features operate in a way that constrains the options available to the user and, on occasion, they might impede the user's particular occurrent prudential preference. Imagine then that a class of products (motor cars, for example) is marketed with the following three design options each reflecting a different level of technological control over the user's self-interested decision-making:

> Level 0 (that is, no) technological assistance or constraint: the user is on his or her own in driving the car.
>
> Level 1 technological assistance or constraint: this is what Mireille Hildebrandt terms a "regulative" technology;[53] it is an amber light alert: it is a motor car fitted with sensors that detect the presence of alcohol or drugs and that cautions against driving under the influence; it is the fridge that warns about food coming up to its eat-by date; it is the energy-smart home that advises the occupier about the levels of fuel consumption; it is the digital voting assistant that advises the user about the voting option that is consistent with the user's standard preferences; and so on. The technological signal is normative.
>
> Level 2 technological assistance or constraint: this is what Mireille Hildebrandt terms a "constitutive" technology.[54] It is red light control: it is the car that is immobilized when its sensors detect the presence of drink or drugs; it is the fridge that destroys the food that has passed its eat-by date; it is the energy-smart home that shuts down the power; it is the gastric band technology that makes it impossible to eat more than a certain amount; and so on. The technological signal is non-normative.

For the individual who elects level 0, there is no technological impingement on prudential reason and action. With level 1 features, the technology simply advises the user, just as a friend might ask, "Do you really think that doing this is in your interest?" The technology is a partner in prudential decision-making, and the user remains in control in the sense that the advice can be ignored. The option that might give some pause is level 2. This is where the technology takes on a non-normative character. For the individual who elects level 2, the power of prudential decision is transferred to the technology. However, the election itself is a strategic prudential decision. For example, an individual might be prone to acting on short-term considerations which lead to actions he subsequently regrets. To minimize this risk, the individual elects level 2 technological features.[55] While this does involve a transfer of control, the transfer is selected and accepted because the

53. Hildebrandt, *supra* note 19, at 172.

54. *Id.*

55. For discussion of how different conceptions of autonomy are implicated in these technological designs, see Roger Brownsword, *Autonomy, Delegation, and Responsibility: Agents in Autonomic Computing Environments*, *in* THE PHILOSOPHY OF LAW MEETS THE PHILOSOPHY OF TECHNOLOGY: AUTONOMIC COMPUTING AND TRANSFORMATIONS OF HUMAN AGENCY 64 (Mireille Hildebrandt & Antoinette Rouvroy eds., 2011).

individual makes a background prudential judgment that, all things considered, this is the way to advance his self-interest. Moreover, provided that the decision is reversible (either by replacing the product or, in some sophisticated designs, by virtue of an override feature), there is no loss of prudential independence.

Perhaps we should not be too sanguine about consumer choices being enhanced by products that offer a range of options of the kind just outlined. After all, consumer preferences can be manipulated and, in markets that use sophisticated profiling and advertising technologies, individuals might have less control over their prudential judgments than they assume. Still, in principle, prudential self-regulatory election of non-normative technological management is much less problematic than the imposition of such management systems by others.

2. The Effect of Non-normative Regulation Imposed by Others

Putting self-regulation to one side, there is the quite different scenario in which non-normative technological features are imposed on regulatees. Here, the question becomes whether such imposition has a cost to the prudential independence that is valued by the community. Prima facie, loss occurs because individuals are no longer making their own self-regulating prudential decisions. However, provided that there is an opportunity to apply prudential reason in public debates in which individuals vote their (collective) preferences, then we are not losing prudential independence from public life—even though the outcome of such debates might be regulatory environments that are non-normative. Arguably, in other words, the loss is not as serious as it first appears.

As we have said, in the marketplace, mass-produced goods might not be designed as one would choose. If this means that the better-off have a better chance of realizing their preferences, there is an obvious concern about the fairness and equity of access to various technological options. But the concern is not about the loss of prudential reason from the life of the community. In the same way, if one market player is in a position to impose a technological restriction on the other (as with DRM technologies and Monsanto's supposed terminator gene in seeds), there is an imbalance of contractual power, with producers using the technology to advance their commercial self-interest against the preferences of purchasers. These facts of market life might give rise to some concern. However, the concern is not about the loss of opportunities for prudential reason so much as the legitimacy of this kind of transactional power play.

Away from the market, what should we make of public impositions of non-normative technologies? For example, what should we make of the non-

normative regulatory environments that are characteristic of many aspects of public transport systems? Suppose officials proposed that, instead of conventional driver-controlled motorcars, there should be a fully technologically-managed road transport system. Even with the best deliberative, democratic, and participatory processes, there is no guarantee that the outcome will align with each participant's judgment of his own personal prudential interest. Once the new transport system is in operation, there is some loss of opportunity for prudential action—drivers will no longer ponder the best route for getting from A to B. However, there has been no loss of prudential reason in the *debates* about the adoption of the system. Indeed, there might be further prudential arguments about possible modifications to the system, or even its abandonment and the restoration of the old system.

However, again, we should not be too sanguine about this. In practice, how often do public debates occur about the adoption of managed environments—rather, how often are there simply incremental changes that just seem to happen? And, even if there are public debates, how often are they framed in terms of a shift from a normative to a non-normative regulatory environment? The current enthusiasm for creating regulatory environments incorporating defaults that "nudge"[56] regulatees towards a particular action indicates how well-intentioned, paternalistic thinking can reshape the regulatory environment. Of course, the beauty of the nudge is that it is analogous to a level 1 technological constraint. The signal is normative, as the regulatee is still in control with the option of opting out. And it seems to be possible to assuage the concerns of those who are worried about the loss of individual autonomy. However, assuming that the purpose of the nudge is legitimate and generally beneficial, there are two aspects of the strategy that invite careful scrutiny. One aspect is whether there has been a public debate about the introduction of the nudge. Have the relevant regulatees signed up for this level 1 steer over level 0 (or possibly level 2)? The other scrutiny-inviting aspect is for the relative ease with which the nudge could become something stronger—so to speak for push to become shove. The nudge will only be effective so long as it produces the pattern of behavior that is desired by the regulators. In some cases, this might require only the gentlest of nudges. However, once the nudge becomes stronger, it starts to approach the boundary that divides normative signals from non-normative signals. The latter, it should be recalled, start at a point at which the regulatee reasons that it is not reasonably practicable to do

56. *See* RICHARD H. THALER & CASS R. SUNSTEIN, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS 6 (2008).

anything other than to go with the flow. Once this happens, even if regulatees have signed up for level 1 nudging, they are unwittingly (and without their endorsement) operating in a level 2, non-normative, regulatory environment.

For a community that values prudential reason and the possibility of acting on one's own prudential judgments, the public imposition of non-normative regulatory environments needs to be preceded by an inclusive public debate that flags to participants the replacement of normative with non-normative signals. Moreover, such communities must take care that what starts out as a level 1 technological regulation does not morph into a level 2 regulation without the community having the opportunity to express its preferences on the matter.

3. Technology Embedded in the Body

Non-normative regulatory technologies that are embedded in persons, rather than in their surrounding environments, raise an interesting problem regarding prudential independence. While there appears to be no loss of the community's aggregate capacity for prudential reason, individuals who are coded for a particular kind of prudence do not seem to enjoy the independence that is integral to valuing prudential decision-making.

Suppose that an individual suffers from depression. He has two options: he can either take a course of drugs or work out regularly at the local fitness club. Whichever option he takes, the biochemistry is identical: serotonin levels are raised and the depression lifts. It is a simple choice, but which option, prudentially, is preferred? Some (perhaps many)[57] will prefer the latter because they are suspicious of drugs for mental health, or they think that their recovery will be more authentic if unaided by drugs, or the like. Others elect to take the drugs. So far as the prudential life of the community is concerned, there seems to be nothing exceptional about any of this. Whichever option is taken, it is the result of the individual's prudential preference, their own independent judgment as to what is in their self-interest.

Suppose that the individual takes the drugs, recovers from depression, but now finds that staying on the drugs enhances their mood. Clearly, there are many questions about the ethics of drugs being used for enhancement

57. *See* ACAD. OF MED. SCI. (UK), BRAIN SCIENCE, ADDICTION, AND DRUGS 28, 54 (2008).

rather than for therapy,[58] but none of this raises concerns about the prudential life of the community. Similarly, whatever doubts we might have about young clubbers choosing to be RFID chipped in order to get fast-track entry or bar service, it cannot be because we sense the loss of prudential reason. Their prudential judgments might be different than our own, but that is quite another matter. Unless the in-person technology is irreversible, it seems that the shape of self-regulatory use of such regulatory technology is much the same as it is with products and places.

In some cases, the initial decision to use an in-person regulatory technology is more suspect than in other cases. For example, we might debate whether an offender's election to be chipped or tagged (for parole or early release or to avoid a custodial disposition) is in any relevant sense "unfree" or forced. Is this really an unforced choice (in the sense required by an appeal to the offender's "election")? Or we might argue about the morality of such measures.[59] However, there is no loss of prudential community here. Similarly, we might debate the merits of tagging children or elderly people for their own health and safety. To the extent that paternalistic reasoning—taking no account of the capacity of these persons for making their own prudential decisions—backs such technological interventions, there is a problem. Although, it should be said that this is a problem that is by no means limited to, or driven by, regulatory technologies. So long as our question is about the preservation of prudential self-determination, the imposition of regulatory technology embedded in the body has not yet hit a nerve.

What would hit such a nerve? Imagine, in the way that Bruce Ackerman once did,[60] that there are master geneticists who can code persons in a way that they have particular talents and, concomitantly, associated preferences. No doubt, for each of us, the way in which we perceive our self-interest, as well as our tendency towards short-term or longer-run calculation, owes something to our genetic inheritance. However, our perceptions have not been designed into us in the self-conscious way that would happen if we employed the services of the master geneticist.[61] So long as this coding is a self-elected, somatic fix, it fits the self-regulatory pattern. However, where

58. *See, e.g.*, JOHN HARRIS, ENHANCING EVOLUTION 7 (2007); MICHAEL SANDEL, THE CASE AGAINST PERFECTION (2007); Roger Brownsword, *Regulating Human Enhancement: Things Can Only Get Better?*, 1 LAW INNOVATION & TECH. 125 (2009).

59. *See, e.g.*, Jeroen van den Hoven, *Nanotechnology and Privacy: Instructive Case of RFID*, *in* NANOETHICS 253 (Fritz Allhoff et al. eds., 2007); Rosenberg, *supra* note 9.

60. BRUCE A. ACKERMAN, SOCIAL JUSTICE IN THE LIBERAL STATE 114–21 (1980).

61. *Id.* at 121–23 (describing parental election of genetic traits in their offspring).

others (the State, our parents, or others) specify the coding for us, this raises a host of *moral* concerns. For liberals, it violates the person's right to an open future,[62] and for dignitarians, it wrongly treats persons as commodities.[63]

But, does it also impinge upon prudential community? In a sense, there is no loss of the community's aggregate capacity for prudential reason, but the individuals who are coded for a particular kind of prudence do not enjoy the independence that is integral when valuing prudential decision-making. Put bluntly, when a person judges that *x* is in her self-interest, she wants that to be her judgment and not a judgment that has been designed into her by others. When her interests are at issue, she wants to be speaking and deciding for herself.

If there is anything exceptional about *imposed* in-person regulatory technologies, it is that they might not be transparent or reversible. In both respects, there seems to be a significant diminution in prudential community.

4. A Regulatory Margin?

We have said already that there needs to be a regulatory margin to facilitate deliberation about, and review of, changes to the complexion of the regulatory environment. Previously, the function of the margin was to provide an opening for considering the amplification of prudential signals (at the expense of moral signals). Now, the margin considers the turning down of such prudential signals in favor of non-normative signals.

Just as before, the marginal responsibilities of regulators might be weak or strong. A regulator's responsibility might be simply to preserve some room for prudential calculation (which one would expect to be relatively undemanding). Or it might be to hold the line and possibly even to promote prudential calculation. For example, this might be a community now facing techno-regulation that has only recently shaken off a culture of paternalism.

Whether weak or strong, a prudential regulatory margin would highlight various considerations. First, provided that individuals self-consciously adopt regulatory technologies because they reason prudentially that this kind of management advances their self-interest, and provided that these decisions are reversible, there seems to be little cause for concerns related to the prudential margin. Second, when public bodies impose non-normative

62. *See, e.g.*, Dena S. Davis, *Genetic Dilemmas and the Child's Right to an Open Future*, 28 RUTGERS L.J. 549, 562–67 (1997) (arguing that a reflexive application of autonomy values will set limits to parents' reproductive autonomy).

63. DERYCK BEYLEVELD & ROGER BROWNSWORD, HUMAN DIGNITY IN BIOETHICS AND BIOLAW 29–47 (2001) (mapping the new dignitarian ethic that holds, *inter alia*, that it is wrong to commercialize or to commodify the human body).

regulatory technologies, an opportunity must exist for prudential and inclusive deliberation before such measures are adopted. The fact that the character of the regulatory environment will change should be highlighted for discussion. If the decision is irreversible, regulators must take a great deal of care before proceeding and ensure that imposed technological management is in line with general prudential preferences.[64]

Finally, any attempt to design-out a person's capacity for prudential reason, or to design-in a particular kind of prudential pathway for a person, should be prohibited.

B. THE IMPACT OF THE SHIFT TO NON-NORMATIVE SIGNALS ON MORAL COMMUNITY

If the amplification of prudential signals can be a problem for moral community, then we might expect a shift from normative to non-normative regulatory signals to accentuate problems regarding actors' agency. As with our discussion of the impact of non-normative management on prudential reason, we can start with self-regulatory choices and then turn to the imposition of techno-regulation.

1. Non-normative Management, Self-Regulation, and Moral Community

This Section discusses three possible concerns that the introduction of non-normative technologies may have on the aspirations of moral community in the particular context of self-regulation.

Let us suppose that we are dealing with an aspirant moral agent—a person who wants to do the right thing relative to the legitimate interests of others. When this person is offered a choice of product design, such as a car, he will be thinking about how the technological management secures his own safety but also about how this safeguards the legitimate interests of others. So, for example, while a purely prudential car-buyer might elect level 1

64. Compare this point to Danielle Keats Citron's argument where she recommended that

> [a]gencies should explore ways to allow the public to participate in the building of automated decision systems . . . .
>
> In the same vein, agencies could establish information technology review boards that would provide opportunities for stakeholders and the public at large to comment on a system's design and testing. Although finding the ideal makeup and duties of such boards would require some experimentation, they would secure opportunities for interested groups to comment on the construction of automated systems that would have an enormous impact on their communities once operational.

Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1312 (2008).

technology that reminds the driver about his or her own safety, a moral car-buyer might elect similar technology that expresses the caution in moral rather than prudential terms, reminding the driver about the safety of other road-users. Indeed, being aware of their own shortcomings, moral agents might choose something stronger than an advisory message, including level 2 technology.

Thus, for moral reasons, the agent has elected technological non-normative management that guarantees that the legitimate interests of other road-users will be respected. Although this seems to align with the aspirations of moral community, this election raises three concerns, respectively: the authenticity of the agent's moral performance, the possibility of expressing human dignity, and the constraints on dealing with moral emergencies.

First, as the motorist proceeds, with the level 2 technological management ensuring that there is no harm to other road-users, one might say that this is an inauthentic moral performance because it is the on-board technology, not the agent, that does all the work. Clearly, this latter point has to be conceded. However, moral reason lies at the root of the technology that has been selected and, arguably, this is good enough.

Second, when the car is in motion and observing the interests of other road-users, the driver cannot proceed otherwise (assuming no facility for overriding the technological controls). There is no possibility that the driver can express his human dignity by choosing not to do the wrong thing. The driver of a car with level 2 technological management never confronts the choice between doing the right thing or the wrong thing. Again, though, the driver is where he is only because the earlier design choice was made. If that choice was made freely, then that seems to be the moment at which he expressed his human dignity. As such, while level 2 technological management may preclude the particular-occurrent expression of human dignity, there is the possibility that humans continue to express their dignity in the prior choice of such technological control.

Thirdly, en route from A to B, the smart car might encounter an emergency in which, without the technological controls, the driver would have deviated to assist another (as in the stock example of a motorist who exceeds the speed limit in order to get a pregnant woman to hospital).[65] No doubt, the really smart car will have an override that allows the moral agent to do the right thing in such an emergency. Failing this, when the moral

65. For discussion of such a case, see Karen Yeung, *Can We Employ Design-Based Regulation While Avoiding* Brave New World, 3 LAW INNOVATION & TECH. 1, 6 (2011).

agent elects level 2 technological management, he must calculate the potential moral cost of subjecting his conduct to the governance of the technology. Still, this does not necessarily signify a loss of moral community.

What if, though, the product was not a car with a specific range of managed compliance? What if, instead, the question concerned the self-administration of a broad sweep moral drug, operating either by tuning up the moral (normative) signals or by repressing the prudential will to defect? Intuitively, we might find this problematic. Recall Mustapha Mond's conversation with the Savage in Aldous Huxley's *Brave New World*,[66] where Mond points out that, in place of all the effort associated with hard moral training, anyone can be moral by swallowing a small amount of soma. As Mond puts it, "Anybody can be virtuous now. You can carry at least half your morality about in a bottle. Christianity without tears—that's what soma is."[67] If the drug simply serves to amplify the moral signals, it might be seen as problematic, but not because of any non-normative characteristics. If the drug functions by repressing any harmful desires, the agent finds it easy to respect others—and, aside from its broad sweep, this seems to be akin to the car with level 2 technological management. For the individual agent who elects to take this shortcut to moral performance, there might be some costs (such as the loss of authenticity[68] or dignity). However, unless the project of moral community requires that moral action be unaided—or, perhaps, unless it becomes so easy for agents to do the right thing that they lose the sense that they face a choice between right and wrong—there is no real problem.

2. Democratic Imposed Regulation

We turn now to consider the effects of techno-regulation where it has been imposed under democratic conditions—that is to say, where the imposition has resulted from a full and inclusive public debate involving regulatees. Moreover, for present purposes, let us suppose that there is general agreement that the moral interests to be protected are important, that techno-regulation will be effective in protecting these interests and that, all things considered, the adoption of a managed environment is the right regulatory strategy. Once this non-normative regulatory environment is in place, regulatees lose the opportunity to do wrong by violating the protected interests of others—which, of course, is precisely the point of making this particular regulatory move. However, it also means that regulatees cannot

66. ALDOUS HUXLEY, BRAVE NEW WORLD 244 (HarperPerennial 1989) (1932).

67. *Id.*

68. Authenticity is by no means a straightforward idea. *See, e.g.*, NEIL LEVY, NEUROETHICS 74–81, 88–94 (2007).

demonstrate in such an environment that they do the right thing for the right reason. How serious a price is this to pay? How serious is it for moral community that agents, in techno-regulated environments, think only about what is practicable or possible rather than what morally is required?

Earlier in the paper, we identified four ideal-typical agents as follows:

> Type 1 agents who act only and always on *moral* reasons;
>
> Type 2 agents who act only and always on *prudential* reasons;
>
> Type 3 agents who act on a mix of *moral and prudential* reasons; and
>
> Type 4 agents who are *erratic*, sometimes acting on moral reasons, sometimes on prudential reasons, and sometimes on mixed reasons.

How does the techno-regulated environment, so designed for moral reasons, impact on each of these agents? While Type 1 agents can remind themselves that they do what they do for moral reasons, they cannot openly demonstrate that this is the case—which a moral community might or might not judge to be problematic. For Type 2 agents (who are pathological in a moral community), there is no loss. These agents, who always act on prudential reasons, are steered by the technological management system towards a moral course of action. To be sure, they lose the opportunity to do the right thing for the right reason; but, if they are never going to do the right thing anyway, this seems to be no loss—and, of course, there is an offsetting moral gain.

What about Type 3 and Type 4 agents? Some of these all-too-human agents will be quite badly conflicted, experiencing weakness of the will as prudential gains trump moral arguments, as well as exhibiting a tendency to rationalize self-serving acts as actually being in line with moral requirements. In short, for such agents, the prudential parts of their practical reason can often defeat their moral aspirations. If a managed regulatory environment prevents this from happening—that is, keeping regulatees on the right moral tracks—this seems to be a positive for moral community. The fact that agents cannot get off once they are set on the right track does not seem too serious a price to pay. Granted, there is no possibility of demonstrating that one is freely doing the right thing for the right reason, but provided that the right reasons were present when the management system was initiated, this seems good enough. Moreover, for some of these Type 3 and Type 4 agents, the problem was always that, when presented with the opportunity, they did not do the right thing.

Having said this, there might be a dual concern that, where techno-regulation is widely employed, Type 3 and Type 4 agents (1) rarely encounter situations where their moral resolve is put to the test and (2) begin to lose a sense of responsibility for their acts. If the former means that the capacity of

these agents for moral reflection and judgment is impaired, this becomes a serious matter for moral community. For, as we saw in our earlier discussion concerning the amplification of prudential signals in Part III, *supra*, moral communities need to keep debating their commitments. In such a community, it is fine to be a passive techno-managed regulatee, but active moral citizenship is also required. As for the latter, David Smith has remarked: "If people are denied any autonomy, then they perceive that the moral responsibility lies entirely with the system, and they no longer retain any obligations themselves."[69] The extent of any such "demoralizing" effects would need to be carefully monitored, for they are clearly corrosive of moral community.

3. In-Person Moral Coding

We said earlier that the coding of persons for prudential preferences could be problematic both for prudential and moral community. Imagine, now, the coding of persons for moral action. In an aspirant moral community, this gives rise to a clutch of concerns, three of which we can highlight.

First, there is the question of whether the coding is an act of self-regulation. If it is, then what is the difference between this and taking a daily dose of soma or whatever that keeps the agent on the moral tracks? Provided that the coding is reversible, then the cases might be comparable, and it is for each agent to make a choice about whether, all things considered, this kind of fix is the best way to lead a moral life. If, however, the coding is imposed, we might want to distinguish between coding before or at birth (which might be seen in a negative or a positive light) and the enforced coding of mature agents who have perhaps shown themselves to be otherwise incapable of respecting the moral interests of others. We might also want to differentiate between coding that amplifies moral signals (or strengthens moral resolve) and that which simply suppresses harmful or dangerous instincts. Whereas, in the former case, we seem to be designing for the moral life, in the latter it seems to be an exercise in risk management. Clearly, there is much devil in the details of such fixes.

Second, as we saw in Part III, *supra*, a moral community will be greatly concerned that technologies are not employed in ways that interfere with the development of a capacity for moral reason and an agent's appreciation of

69. David J. Smith, *Changing Situations and Changing People*, *in* ETHICAL AND SOCIAL PERSPECTIVES ON SITUATIONAL CRIME PREVENTION 147, 170 (Andrew von Hirsch et al. eds., 2000).

morality as a normative code. Famously, in its report *Beyond Therapy*,[70] the President's Council on Bioethics expressed just this concern in relation to the administration of methylphenidate (Ritalin) and amphetamine (Adderall) to children whose conduct is outside the range of acceptability. Thus:

> Behavior-modifying agents circumvent that process [i.e., the process of self-control and progressive moral education], and act directly on the brain to affect the child's behavior without the intervening learning process. If what matters is only the child's outward behavior, then this is simply a more effective and efficient means of achieving the desired result. But because moral education is typically more about the shaping of the agent's character than about the outward act, the process of learning to behave appropriately matters most of all. If the development of character depends on effort to choose and act appropriately, often in the face of resisting desires and impulses, then the more direct pharmacological approach bypasses a crucial element . . . . By treating the restlessness of youth as a medical, rather than a moral, challenge, those resorting to behavior-modifying drugs might not only deprive [the] child of an essential part of this education. They might also encourage him to change his self-understanding as governed largely by chemical impulses and not by moral decisions grounded in some sense of what is right and appropriate.[71]

Accordingly, if we rely on biotechnological or neurotechnological interventions to respond to (or manage) our social problems, there is a danger that, as the President's Council puts it, "we may weaken our sense of responsibility and agency."[72]

Third, once one makes a coding intervention, is that intervention capable of responding to changes in the community's interpretation of their moral commitments and the way in which fundamental principles should be applied? If the coding simply represses anti-social instincts, or if it strengthens the signal to do the right thing, it might continue to be functional even as the substance of morality changes. However, so long as the moral project is understood as an ongoing one, the community will want to take a hard look at in-person measures lest they should inappropriately freeze morals.

---

70. PRESIDENT'S COUNCIL ON BIOETHICS, BEYOND THERAPY: BIOTECHNOLOGY AND THE PURSUIT OF HAPPINESS (2003), *available at* http://bioethics.georgetown.edu/pcbe/reports/beyondtherapy/beyond_therapy_final_webcorrected.pdf.

71. *Id.* at 91–92.

72. *Id.* at 92.

4. The Moral Margin

In Sections III.C and III.D, *supra*, we sketched the idea of a moral margin in the context of the amplification of prudential signals. This sketch continues to apply where the questions for moral community arise not from the amplification of prudential signals but from non-normative regulatory approaches. So, it would continue to be important to determine whether the community's vision of its project implies a weak or a strong stewardship responsibility for regulators. Also, it would continue to be essential to prevent technological interference with the development of the capacity for moral reason and an appreciation of the normative character of morality. The added protection of important moral interests would continue to be material.

Let me offer a few comments on a couple of questions that have previously seemed difficult to resolve.[73] The first is whether, when a techno-regulatory intervention precludes certain kinds of harmful acts, it matters if those acts are intentional or unintentional. The second is whether it matters that harmful acts are prevented by disabling an aggressor or by designing in protection for the victim.

The first puzzle arises where products (such as surgical instruments)[74] or complexes of products (such as transport systems) are designed for safety. Primarily, the purpose of such safety measures will be to safeguard users or passengers—for example, by phasing out trains with slam door carriages[75] or by making it impossible for trains to pass through signals on red. Given that such measures are designed to make routine activities (such as the journey to work) less risky, it is reasonable to assume that most interested parties judge them to be in their prudential interests. And, if public engagement has indeed shown this to be the case, then all is well and good. However, the effect of these measures is not only to replace prudential norms with non-normative design but also to impact on the opportunity to display a moral performance. For example, commuters opening railway train doors might want to show that they do so with due regard for the safety of fellow passengers and persons standing on station platforms. Likewise, train drivers might want to show that they exercise due care by stopping at red signals. Once the train is designed for safety, these displays of due care and concern for others cannot

73. *See* Roger Brownsword, *So What Does the World Need Now? Reflections on Regulating Technologies*, *in* REGULATING TECHNOLOGIES, *supra* note 10, at 24; Yeung, *supra* note 65.

74. *See* Yeung, *supra* note 65; Karen Yeung & Mary Dixon-Woods, *Design-Based Regulation and Patient Safety: A Regulatory Studies Perspective*, 71 SOC. SCI. & MED. 540 (2010).

75. *See* Jonathan Wolff, *Five Types of Risky Situations*, 2 LAW INNOVATION & TECH. 151 (2010).

be made in this way. Assuming that the community values such displays of moral virtue, do regulators have a short answer to these "objections"?

One thought is that regulators might be able to say that, where their primary purpose is the safety of passengers, they do not have to answer for any secondary effects—that they are shielded by a doctrine akin to that of double effect. Surely, though, this will not do. Otherwise, this would involve accepting that, because Robert Moses's bridges were built with safety in mind, there is no need for regulators to answer for their secondary (and racially discriminatory) effects.[76] This is quite contrary to one of the main points in this Article, namely that regulators need to be much more sensitive to the impact of relying on architecture, product design, and the like as features of the regulatory repertoire.

The other thought is that there is no real loss of moral community when such safety features are introduced because, insofar as the intervention targets acts that are harmful to others, its focus is on unintentional rather than intentionally harmful acts. If the technology only prevented non-negligent unintentionally harmful acts, there might be something in this thought. However, technology also blocks negligent acts as well as intentionally harmful acts. Now, as we have indicated already, in a moral community, it is important not only to eschew intentionally violating the protected interests of others but also to respect such interests by taking reasonable care not to cause harm to others. To be sure, a dog might know the difference between being kicked intentionally and unintentionally. But a smart dog will also distinguish between an owner that takes reasonable care not to kick it and one that takes no such care. At all events, for the sake of argument, let us assume that it is conceded that regulators do not have to answer for any impingement on unintentional acts (even negligent acts). Here, the crucial point is that regulators must not interfere with opportunities for intentional wrongdoing. On the face of it, such a norm is strange because under it regulators, whatever other good they may do by using non-normative controls, must not deprive those agents who might intentionally harm others of the opportunity to do so. The deprivation of opportunity to harm others, in turn deprives agents of the opportunity to demonstrate that they are freely doing the right thing. Hence, train drivers must not be prevented from passing through signals when they are on red, lest this prevents the driver from showing that he does the right thing by stopping on red. This, as previously noted, seems a strange view. Indeed, it is tempting to

76. On value-sensitive design, see Noëmi Manders-Huits & Jeroen van den Hoven, *The Need for Value-Sensitive Design of Communication Infrastructures*, *in* EVALUATING NEW TECHNOLOGIES 51, 54 (Paul Sollie & Marcus Düwell eds., 2009).

say that no moral community could reasonably attach such importance to preserving the opportunity to do wrong in order to demonstrate that one does right. Having said that, a moral community might perfectly reasonably attach importance to the existence of some such opportunities and the question then would be whether train drivers or their passengers need this particular opportunity more than they need the design-in safety features—a question for the regulatory margin.

A second puzzle arises from the possibility that regulators might be able to prevent A from causing harm to the protected moral interests of B either by disabling A or by shielding B. Let us suppose that the strategies are equally effective. Nevertheless, if one strategy is, for moral reasons, better than the other, this might be an issue for review within the terms of the regulatory margin. Is there any moral reason to prefer one strategy to the other? Initially, this seemed to be a distinction without any morally significant difference.[77] However, on second thoughts, it might be preferable to shield B rather than to disable A, because this would at least leave open the possibility for A to attempt to deviate and to be aware that such deviation was contrary to the regulatory code.[78] If we place the puzzle in the larger context of the preservation or promotion of moral community, it is surely desirable to retain the relevant moral signals in the interaction between A and B. One way of achieving this might be by coding A so that moral signals are amplified to the point that A is disabled from harming B. Alternatively, there might be scope for traditional moral reasoning with A, knowing that, even if the reasoning fails to restrain A, B cannot be harmed. This leaves the matter unresolved. However, in the absence of a particular context and without knowing the range of the design options, it is difficult to take this any further. All that we can say is that this would be a question to be addressed within the terms of the regulatory margin.

Taking stock, this Part of the Article has reviewed the implications of the adoption of non-normative regulatory strategies. In particular, this Part has focused on the implications for the prudential life of a community where, put simply, agents value the opportunity to make their own decisions about what is in their own best interests. And the discussion has revisited those communities that have moral aspirations to assess the implications of non-

77. *See* Roger Brownsword, *Neither East nor West; Is Mid-West Best?*, 3 SCRIPTED 15 (2006).

78. *See* Brownsword, *supra* note 73, at 42–43 ("[A] community of rights might reason that there is a significant difference between design-out and design-in because, in the former case, agents are only dimly aware (if at all) that they are doing right rather than wrong, while in the latter case agents will be aware they are deviating.").

normative management for their project. For simple prudentialists and for moralists alike, there is much to ponder where the regulatory environment assumes a non-normative complexion. And, for lawyers, there is an overwhelming question to answer. Quite simply, what happens to law when the regulatory environment is dominated by technologies that steer regulatees via non-normative signals? It is to this lawyers' question that this Article now turns.

V. SUSTAINING LEGALITY

This final Part returns to where this Article started, with questions that relate to the ideals of legality and the rule of law. Put starkly, where non-normative instruments dominate the regulatory environment, we seem to be subject to the rule of technology rather than the rule of law. If we value the rule of law, we need to be able to rescue and recycle it even in non-normative regulatory environments. My argument is that we can do this provided that we anchor ourselves to a conceptual understanding of law and legality that captures those aspects of moral community that we are most anxious to preserve.

How much of law survives in regulatory environments that have transitioned to techno-management? To be sure, there might still be some laws in the background, but all the foreground work is done by techno-regulation. If the regulatory environment retains some normative signals, they are so weak as to be irrelevant. This, however, is not the real issue. What really matters is whether the processes that lead to the particular techno-regulatory features are compatible with the ideal of legality.

When Lon Fuller proposed that his eight desiderata (or principles) of legal ordering should be understood as the "inner morality of law," his legal positivist critics saw this as a fundamental error.[79] H.L.A. Hart, for example, ridiculed the idea that the promulgation of clear prospective rules and their congruent administration could be characterized as moral requirements because, quite simply, they were compatible with the pursuit of evil purposes.[80] At most, the legal positivists argued, the Fullerian principles were

79. FULLER, *supra* note 3, at 42–43.

80. H.L.A. Hart, *Review of* The Morality of Law *by Lon L. Fuller*, 78 HARV. L. REV. 1281 (1965) ("He takes me seriously to task for having said that respect for the principles of legality is unfortunately 'compatible with very great iniquity'; but I cannot find any cogent argument in support of his claim that these principles are not neutral as between good and evil substantive aims. Indeed, his chief argument to this effect appears to me to be patently fallacious.").

guidelines for *effective* ordering of social life.[81] Understandably, Fuller was puzzled by such criticism. In response, he might have said that the desiderata were moral requirements *independently* of the underlying morality of the regulatory purposes (just as contract lawyers might argue that good faith and fair dealing are moral requirements even though the transaction might be unconscionable or illegal or contrary to public morals, and the like). Or, Fuller might have said that compliance with the procedural principles was necessary although not sufficient for fully moral performance. Or, he might have stuck with his first instinct that the critics' line of argument was "so bizarre, and even perverse, as not to deserve an answer."[82] However, Fuller did not rely on such short retorts. Instead, he went right back to what he took to be his own starting point and, as it now seemed, the somewhat different starting point of the legal positivists.[83]

For both sides, it was agreed that law, in a pre-theoretical sense, refers to the enterprise of subjecting human conduct to the governance of rules. However, Fuller traces his differences with his critics to two key assumptions made by the legal positivists, namely:

> The *first* of these is a belief that the existence or non-existence of law is, from a moral point of view, a matter of indifference. The *second* is an assumption . . . that law should be viewed not as the product of an interplay of purposive orientations between the citizen and his government but as a one-way projection of authority, originating with government and imposing itself upon the citizen.[84]

The second of these assumptions is elaborated in a contrast that Fuller draws between a legal form of order and simple managerial direction. He sketches the distinction between the two forms of order in the following terms:

> The directives issued in a managerial context are applied by the subordinate in order to serve a purpose set by his superior. The law-abiding citizen, on the other hand, does not apply legal rules to serve specific ends set by the lawgiver, but rather follows them in the conduct of his own affairs, the interests he is presumed to serve in following legal rules being those of society generally. The directives of a managerial system regulate primarily the relations between the subordinate and his superior and only collaterally the relations of the subordinate with third persons. The rules of the legal system, on the other hand, normally serve the primary

81. *Id.*
82. FULLER, *supra* note 3, at 201.
83. *Id.* at 190–91.
84. *Id.* at 204.

> purpose of setting the citizen's relations with other citizens and only in a collateral manner his relations with the seat of authority from which the rules proceed. (Though we sometimes think of the criminal law as defining the citizen's duties towards his government, its primary function is to provide a sound and stable framework for the interactions of citizens with one another.)[85]

As Fuller concedes, these remarks need "much expansion and qualification."[86] He tries to give more substance to them by characterizing the relationship, in a legal order, between government and citizens in terms of "reciprocity" and "intendment."[87] Perhaps, Fuller's most evocative observation is that "the functioning of a legal system depends upon a cooperative effort—an effective and responsible interaction—between lawgiver and subject."[88]

No doubt, these seminal Fullerian ideas are open to many interpretations. However, for our purposes, it is the association of legal ordering with a two-way reciprocal process that is most fruitful. For, in the larger context of the regulatory environment, it implies that the legal approach—an approach to be valued—is one that embeds participation, transparency, due process, and the like. Hence, if we take our lead from Fuller, we will surely reason that, as the translation is made from a normative to a non-normative regulatory environment, we certainly need to hold on to the idea that what we value is a reciprocal enterprise, not just a case of management by some regulatory elite.

Accordingly, while various kinds of self-regulation that adopt measures of technological control might be fine, even empowering, the imposed public ordering of the community needs to respect the values of legality. This means that a comprehensively transparent and democratic relationship between regulators and regulatees must exist.

How far should that relationship extend? If we try to tease out an answer to this question by pouring over Fuller's text, we will surely think that regulators should engage with the prudential preferences of their regulatees. However, we might be less sure about how far Fuller sees legal order as a community's best expression of its moral commitments. Let me cut through this by saying that, for those who take (as I do) a morally-driven view of law, then it is not just the prudential preferences of regulatees that matter. There is more to law than assisting regulatees to know where they stand so that they can maximize their self-interested preferences. A moral community is an

85. *Id.* at 207–08.
86. *Id.* at 208.
87. *Id.* at 209.
88. *Id.* at 219.

interpretive community and the regulatory environment at any one time should reflect the community's best understanding of its moral project. It is critical for such an aspirant moral community that there be no technological interference with the moral development of agents; that technological interventions should be reviewable and reversible; and that there be, at minimum, a clear and protected margin for moral action.

It follows that one of the challenges for legal forms of ordering in the twenty-first century is to construct regulatory environments that enable moral community to flourish, even though the normativity of the foreground signals might have given way to techno-regulation. Provided that the character and content of the regulatory environment flows from a reciprocal engagement with regulatees, and provided that the background discourse continues to be informed by prudential and moral reason, the things that we value about law will not have been lost. As Fuller rightly says, whether or not we have a regulatory environment of this kind is far from being a matter of moral indifference.[89] To this we might add: if we are indifferent to the kinds of questions raised in this Article, the regulatory environment that we have will be, at best, no more than we deserve.

VI. CONCLUSION

In this Article, there are three "take home messages." The first is that to appreciate the potential impact of emerging technologies on our regulatory environments and on our cultural and social lives, we need to understand the significance of the regulatory registers. With this understanding, we can identify two key movements: the amplification of prudential signals with the use of that register and the shift away from normative signals as technological management takes over. The second is that, in the transition from legal normativity to techno-regulation, we do not have to lose the spirit of legality. Even though normative signals might fade from the foreground, we can (and should) ensure that the relationship between regulators and regulatees is reciprocal. Regulatory environments might become techno-managed but they should not be backed by managerialism. The third is that the relationship between regulators and regulatees can only become fully reciprocal if the complexion of the regulatory environment becomes a matter for public debate and review. Responsible and responsive regulators necessarily engage with their regulatees in setting policies that are in line with general preferences as well as being compatible with the community's moral commitments. However, this is not sufficient. Regulatees also need to be

89. *Id.* at 181–82.

engaged in determining the instrumental complexion of the regulatory environment or to what extent it relies on normative and non-normative signals. Thus the idea of a regulatory margin creates some opening for this kind of engagement. If we are not sensitive to the importance of these features of the regulatory environment, technological management will be adopted where it seems to be effective and the drift away from normativity will simply happen. Recognizing the need to debate these important questions will give us the chance to exert some collective control over our legal, moral, and prudential futures.

FROM PREEMPTION TO CIRCUMVENTION: IF TECHNOLOGY REGULATES, WHY DO WE NEED REGULATION (AND VICE VERSA)?

Helen Nissenbaum†

TABLE OF CONTENTS

I. INTRODUCTION

Thank you so much for the kind introduction; however, I am now absolutely terrified: how am I possibly going to live up to it? Learning about the type of person David Nelson was makes me realize what a tremendous

† Professor of Media, Culture, and Communication & Computer Science, and Senior Fellow of the Information Law Institute, New York University. I gratefully acknowledge support from the National Science Foundation, Grant Nos. 1058333 and 083124. Thanks to Luke Stark for superb assistance throughout the preparation of this manuscript, to Berkeley's Center for Law and Technology for the opportunity to think through the issues, and to BTLJ editors for truly admirable work toward a polished final text.

This Article is adapted from the David Nelson Memorial Keynote given on March 3, 2011, at the symposium *Technology: Transforming the Regulatory Endeavor*, sponsored by the Berkeley Center for Law and Technology.

honor it is to present this memorial lecture.[1] It is also an honor and a pleasure to present to so many of my friends and colleagues whose work I follow closely, learn from, and respect enormously.

The title, I confess, is slightly misleading as the tap on my shoulder to send it came before I had written the talk. When I began work on the text, I realized that filling out the argument implied in the title would better fit a book length project than anything I could manage in a single lecture. Luckily, excellent presentations made earlier today by other speakers, covering the "vice versa" in the title—that is, issues that arise when technologies function in the service of existing law or regulation and particularly when they do so preemptively—address gaps I shall be leaving in my talk in focusing on flows of influence in the other direction. In particular, my attention will mostly be drawn to the role of law and regulation in circumstances where regulation by technology seems already to be in place, or, put another way, where regulation is already encoded in architecture.

My final confession is that I am trained as a philosopher, not as a lawyer. While I am committed to exploring the role of explicit regulation and law in relation to technology, I prefer to frame this exploration not merely in terms of artifacts encoding, enforcing, or preempting law, but as embodying values—specifically, political and ethical values. In so doing, I have tried to connect a set of questions that has been discussed in the field of information law for just over a decade[2] to questions discussed in the philosophical and social study of technologies for approximately half a century.[3] The more general question that has puzzled me since I began working in this area is this: if technology embodies values, and if technology is capable of regulation, what role is left for law and regulation? In this presentation, I focus on one specific aspect of this question about the role of technology, beginning with some background.

1. David Nelson started his practice in Silicon Valley in the 1960s with the firm that became Morrison & Foerster LLP. He was renowned not only for his knowledge of and interest in high tech law, but also for his remarkable memory for popular culture trivia. J.L. Pimsleur, *David E. Nelson*, SFGATE.COM (Feb. 4, 1999), http://articles.sfgate.com/1999-02-04/news/17678718_1_morrison-foerster-mr-nelson-douglas-s-nelson.

2. *See generally* LAWRENCE LESSIG, CODE: VERSION 2.0 (2006); Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 TEX. L. REV. 553 (1998).

3. *See, e.g.*, Lewis Mumford, *Authoritarian and Democratic Technics*, 5 TECH. & CULTURE 1 (1964).

II. POLITICAL TECHNOLOGY

Consider this quote by Langdon Winner from his most famous article, "Do Artifacts Have Politics?"[4] As you know, Winner, of course, answers, "Yes, they do!" and writes:

> In that sense, technological innovations are similar to legislative acts of political foundings that establish a framework for public order that will endure over many generations. . . . The issues that divide or unite people in society are settled not only in the institutions and practices of politics proper, but also, and less obviously, in tangible arrangements of steel and concrete, wires and semiconductors, nuts and bolts[5]

—and, I want to add, lines of code.

When I first read Winner's article many years ago, I immediately wanted to be a soldier in this intellectual struggle, and much of my early work looked at specific information systems and devices in order to point out their politics and why these politics were problematic if we did not pay sufficient attention to them.[6] Revealing the politics in search engines[7] and bias in computer systems (in collaboration with Lucas Introna and Batya Friedman, respectively) are two of my published contributions to this line of work.[8] But Winner's exhortations went beyond merely arguing that artifacts "have" politics—that is, the capacity to settle political issues or inherently favor certain structures of power and authority in society; they also call on creators of technical systems and devices to pay heed, early on in development, to moral and political factors:

> By far the greatest latitude of choice exists the very first time a particular instrument, system, or technique is introduced. Because choices tend to become strongly fixed in material equipment, economic investment, and social habit, the original flexibility vanishes for all practical purposes after the initial commitments are made. . . . For that reason the same careful attention one would give to the rules, roles, and relationships of politics must also be

4. Langdon Winner, *Do Artifacts Have Politics?*, *in* THE WHALE AND THE REACTOR: A SEARCH FOR LIMITS IN AN AGE OF HIGH TECHNOLOGY 19 (1986).

5. *Id.* at 29.

6. *Cf.* Finn Brunton & Helen Nissenbaum, *Vernacular Resistance to Data Collection and Analysis: A Political Theory of Obfuscation*, FIRST MONDAY, May 2011, http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3493/2955.

7. Lucas D. Introna & Helen Nissenbaum, *Shaping the Web: Why the Politics of Search Engines Matters*, 16 INFO. SOC'Y 169 (2000).

8. Batya Friedman & Helen Nissenbaum, *Bias in Computer Systems*, 14 ACM TRANSACTIONS ON INFO. SYS. 330 (1996).

> given to such things as the building of highways, the creation of television networks, and the tailoring of seemingly insignificant features on new machines.[9]

My own chance to respond to Winner's exhortation to design with politics in mind came in 2006 when I collaborated with Daniel Howe in designing and developing TrackMeNot ("TMN"),[10] and again in 2009 with colleagues at Carnegie Mellon and Stanford on Adnostic.[11] In creating these small systems, we were also testing out Values-at-Play ("VAP"), a framework developed with Mary Flanagan, a games designer, and Daniel Howe.[12] Inspired by the practical turn in Winner's work and others in a similar vein,[13] VAP was conceived as a systematic approach to guide an analysis of values embodied in technology, as well as the practical foundation for a heuristic aimed at designers wanting to include values among the standards they considered as they created new systems.

In what follows, a brief description of TMN, Adnostic, and VAP will serve as a departure point for a more general discussion of the relationship between law and technology: how they affect one other, how they support, or how they obstruct. The goal of this discussion is ultimately to shed light on our question, "if we have technology, why do we need law?"—regarding the limits to regulation by technological design. I am still an ardent admirer of Winner, but—and I am sure he would agree with this sentiment—it is not quite as straightforward as simply plugging values into a technology and then believing that you have immediately had some positive and protracted impact on society.

---

9. Winner, *supra* note 4, at 29.

10. Daniel C. Howe & Helen Nissenbaum, *TrackMeNot: Resisting Surveillance in Web Search*, *in* LESSONS FROM THE IDENTITY TRAIL: ANONYMITY, PRIVACY, AND IDENTITY IN A NETWORKED SOCIETY 418 (Ian Kerr et al. eds., 2009).

11. *See generally* VINCENT TOUBIANA, ARVIND NARAYANAN, DAN BONEH, HELEN NISSENBAUM & SOLON BAROCAS, ADNOSTIC: PRIVACY PRESERVING TARGETED ADVERTISING (2010), *available at* http://crypto.stanford.edu/adnostic/adnostic-ndss.pdf.

12. Mary Flanagan, Daniel C. Howe & Helen Nissenbaum, *Embodying Values in Technology: Theory and Practice*, *in* INFORMATION TECHNOLOGY AND MORAL PHILOSOPHY 322 (Jeroen van den Hoven & John Weckert eds., 2008); Mary Flanagan, Daniel C. Howe & Helen Nissenbaum, *Values at Play: Design Tradeoffs in Socially-Oriented Game Design*, *in* PROCEEDINGS OF ACM CHI 2005 CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS 751 (SIGCHI & Ass'n for Computing Mach. eds., 2005).

13. *See, e.g.*, James H. Moor, *What Is Computer Ethics?*, 16 METAPHILOSOPHY 266 (1985).

III. TRACKMENOT, ADNOSTIC, VALUES-AT-PLAY

So what is TrackMeNot? Initially created as a lightweight Firefox extension, available for free download, it now also functions in Chrome.[14] TMN automatically generates fake search queries and sends them to Google, Bing, Baidu, and Yahoo!. Although it does not do anything to protect against online identification, the idea is to guard users against profiling based on logs of search queries accumulated by search engines; the fake queries sent by TMN perform an obfuscatory function. Further discussion of how TMN works can be found on the website and in various publications.[15]

Among those who have chosen to use TMN—at a modest estimate approximately 800,000 total downloads (not people!)—there has been tremendous support and enthusiasm. However, there has also been criticism ranging from charges that it does not work because search engines will easily be able to find and delete TMN-generated queries to charges that it is immoral because it involves dishonesty and wasted resources. As designers committed to the practical importance of moral as well as engineering standards, we cannot focus only on the "Does it work?" question and ignore the charges of immorality; accordingly, we have been attending to both.[16]

VAP, developed with Mary Flanagan and Daniel Howe, provides not only an analytic framework for revealing moral and political values in the design of technologies but the foundations for a heuristic to which designers, who are committed to taking values into account in practice, may refer. It posits two key activities: (1) Discovery and (2) Translation, which are performed not in a single rigid order but in iterative cycles. In Discovery, designers seek to identify values relevant to given projects, sometimes revealed in efforts to locate the sources of these values. In some instances, the sources include the very functional definition of a project; in others, the sources of values emerge when designers discover that seemingly technical decisions have implications for values promoted or blocked by a system in question. Having enumerated values by attending the possible sources, a second component of Discovery is an activity we have labeled *operationalization*. Operationalization typically involves developing concrete

14. To download TrackMeNot, see TRACKMENOT, http://cs.nyu.edu/trackmenot/ (last visited Sept. 24, 2011).

15. Brunton & Nissenbaum, *supra* note 6; Howe & Nissenbaum, *supra* note 10; Vincent Toubiana, Lakshminarayanan Subramanian & Helen Nissenbaum, TrackMeNot: Enhancing the Privacy of Web Search (Mar. 22, 2011) (unpublished manuscript), *available at* http://arxiv.org/abs/1109.4677; TRACKMENOT, *supra* note 14.

16. Brunton & Nissenbaum, *supra* note 6; Toubiana, Subramanian & Nissenbaum, *supra* note 15.

definitions of relevant values for the context of a given design project. It forms a crucial bridge between highly abstract concepts—such as privacy, security, and autonomy—and Translation, the second key activity. Translation, in turn, involves three components: (a) implementing values in design features and architecture, (b) resolving the inevitable conflicts of values that arise as design proceeds, and (c) verifying that one's efforts have been sound by both engaging with designers' reflective capacities as well as ascertaining users' responses.

In the case of TMN, one source of values was, clearly, the definition of the project—that is, protecting privacy in web search. We understood privacy as contextual integrity, or as flow of personal information that is consistent with context-specific informational norms, and we operationalized this as preventing access by search engines to the accurate record of your web searches. When we realized that our implementation did not protect against identification, we realized that we were not able to pursue all our functional and value ends simultaneously and decided to prioritize the effort to obfuscate users' profiles and maintain simplicity and ease-of-use above the effort to anonymize. When users complained that TMN would occasionally send politically or sexually charged search terms, we responded by offering users the ability to select the RSS feeds that would seed TMN's searches, but we decided against direct censorship. When critics pointed out TMN's vulnerabilities to side channel and other attacks, we defended against these attacks, all the while mindful of usability goals, and continuously engaged in iterative cycles of Discovery, operationalization, implementation, conflict resolution, verification, and back again to Discovery.[17]

Adnostic, developed for the purpose of protecting privacy in the face of online behavioral advertising, also constitutes a case in which the functional definition was among the sources of values embodied in it. Responding to concerns over dominant models of behavioral advertising, in which users are tracked across websites in order to target ads to them based on their behavioral profiles, we took seriously its defenders, who claimed that better targeting of ads lured more advertising revenue—the web's lifeblood—ultimately supporting innovation and free content. But, we took issue with the belief that pervasive tracking was necessarily tied to this promise. So, we set about designing a system—Adnostic, a Firefox extension—that internalizes tracking to the browser and is based on the profile it builds from users' browsing habits, which it never shares with third parties; it selects ads

17. Howe & Nissenbaum, *supra* note 10; Toubiana, Subramanian & Nissenbaum, *supra* note 15.

to present to users. For Adnostic to function properly, ad servers would need to collaborate, serving browsers not one, but many, ads from which Adnostic selects the most appropriate.

As with TMN, Adnostic raises questions on various fronts. For example, some might consider its touch too soft for it merely asks websites it visits not to allow third-party tracking and does not attempt to block this entirely. Others might argue that in enabling targeting at all, it buys into a paradigm of user profiling that true protection of privacy should not support at all. These criticisms amount to disagreements over the ways we have operationalized the relevant values, ways we have chosen to implement them, and ways we have chosen to resolve values in conflict. I will pursue this discussion no further, except to say that you can learn more about Adnostic and download the code from the Adnostic website, but, unfortunately, without ad networks willing to step forward to present ads in accordance with Adnostic's requirements, unlike TMN, the system is not able to function "in the wild."[18]

IV. LAW AND TECHNOLOGY; TECHNOLOGY AND LAW

From these brief accounts of TrackMeNot and Adnostic, let us return to the question at the heart of this talk. As I, and others, have observed, law and technology both have the power to organize and impose order on society. Panel presentations earlier today richly observed ways in which they both systematically afford certain behaviors, activities, and practices, and systematically impede others; both have capacities to enable, constrain, allow, and prevent. Both have—to invoke a term Bruno Latour uses—prescriptive capacities.[19] Of course, the prescriptive capacity of law and public policy has been studied for centuries; technology's prescriptive character, however, has emerged as a subject of explicit philosophical, social, and legal study only recently and even so, on a relatively small scale. For some of us, it is the multifaceted relationship between these two prescriptive systems—law and technology—that has been most compelling.

Addressing this relationship were some of the comments we heard earlier today about techno-law, as well as a body of literature that is concerned with whether built objects, including technology, that preempt or enforce law are problematic.[20] In artifacts like subway turnstiles, ten-foot-high barbed wire

18. To download Adnostic, see ADNOSTIC, http://crypto.stanford.edu/adnostic/ (last visited Sept. 24, 2011).

19. Bruno Latour, *Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts*, *in* SHAPING TECHNOLOGY/BUILDING SOCIETY 225 (Wiebe Bijker & John Law eds., 1992).

20. *Cf.* Ian Kerr, *Digital Locks and the Automation of Virtue*, *in* "RADICAL EXTREMISM" TO "BALANCED COPYRIGHT": CANADIAN COPYRIGHT AND THE DIGITAL AGENDA 247

fences at border crossings, devices for monitoring criminals on parole, or full-body scans at airports, technology is put in the service of law in these ways. Tying these cases to the title of my talk, they bear on the question of why we need technical pre- and proscriptions when we already have legal ones. There is much to discuss: observing the different ways technology can support law, evaluating which forms of technical enforcement or preemption are acceptable, and finally, whether it is better to achieve desired behaviors by forcing, or merely by "nudging."[21] As mentioned in my introductory remarks, however, I will not be pursuing this line of thinking further, beyond acknowledging it as a reference point in the larger landscape and, of course, recognizing its importance.

Unlike these scenarios, in which law exists and technology's prescriptive role needs to be explained, the scenarios I want to discuss in the time remaining are those in which technology embodies values or regulates behavior, and we are left to wonder what, if anything, is the role left for law and public policy. In the words of the talk's title, where it seems possible to rule effectively by technology alone, is not law simply redundant? Even if this potential exists in only a handful of cases, one might pose the question: why not dispense with legal regulation entirely when code purportedly regulates on its own?

V. REGULATION BY DESIGN AND ITS LIMITS

Law may be needed in cases where regulation by technology contradicts societal values, a concern Winner raises when arguing that artifacts, by themselves, have the capacity to settle political controversies.[22] Bypassing political channels of collective decision-making, he calls into question the moral legitimacy of technology-based prescriptions (or regulation), particularly when they run afoul of values to which a society explicitly subscribes.[23] In the paradigmatic case of Robert Moses's infamous overpasses, regulators ought to have been awake to the fact that they reinforced socio-economic and racial prejudices, thereby obstructing our

(Michael Geist ed., 2010); Danny Rosenthal, *Assessing Digital Preemption (and the Future of Law Enforcement?)*, 14 NEW CRIM. L. REV. 576 (2011).

21. *Cf.* RICHARD H. THALER & CASS R. SUNSTEIN, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS (2008); *see also generally* Roger Brownsword, *Lost in Translation: Legality, Regulatory Margins, and Technological Management*, 26 BERKELEY TECH. L.J. 1321 (2011) (discussing techno-regulation, regulatory registers, the amplification of prudential signals, and the shift away from normative signals in a technologically managed environment).

22. Winner, *supra* note 4, at 22–29.

23. *Id.* at 38.

explicit commitments to equality of opportunity.[24] Another case, discussed by Rachel Weber, was the ultimately successful effort to redesign the cockpits of fighter planes in the U.S. Air Force so that they fit the generally smaller dimensions of women's bodies.[25]

Yet in countless instances where there may be no obvious contradiction between legal commitments and built systems, technology emerges as a mode of governance outside of government. Technology mediates and gives texture to certain kinds of private relationships; it weighs in on the side of one vested interest over others. What is law doing in these cases; what might it do? In addressing these questions, I acknowledge that I am working in a tiny corner of the sweeping philosophical terrain that covers fundamental questions such as "What is the function of law?", "What is the function of technology?", "Who makes the decisions?", and so on. My aim, here, is to tug on special instances of these larger questions, particularly as they apply to recent controversies involving information technologies and digital media, including some of the questions we have asked at this workshop.

Taking a few steps back, I offer this thought: life confronts us with many obstacles. For example: steep inclines; cancer; gravity; tall trees with the juicy fruit at the very top; and the case Winner made famous, of Robert Moses's nine-foot-high overpasses making it difficult for commuters on the twelve-foot-high buses to reach many of Long Island's beaches.[26] Locks are obstacles, as are seatbelts, body scanners, machine safety overlocks, one-way streets, and queues in banks and supermarkets. Among these obstacles are mechanisms that get in our way in systematic, politically relevant ways—for example, discriminating against some people over others. Not all are problematic: locks, for example, are very nice for the security of owners, but they badly discriminate against thieves. A question we may think to pose is this: among the obstacles we confront each day of our lives, there are some that we seem to accept as impenetrable barriers around which we adjust the patterns of our activities—we may even welcome, support, or invite them—and there are others that we conceive as challenges that must be overcome. We say, "No, we're not going to be stopped from flying—we are going to build airplanes" or "Cancer is a terrible thing—we have to find a cure for it."

24. *See id.* at 22.

25. Rachel N. Weber, *Manufacturing Gender in Military Cockpit Design*, *in* THE SOCIAL SHAPING OF TECHNOLOGY 372 (Donald MacKenzie & Judy Wajcman eds., 1999).

26. Winner, *supra* note 4, at 22–25.

VI. PFAFFENBERGER'S TECHNOLOGICAL DRAMAS

To explain how this distinction relates to our discussion of technology and law, I would like to introduce a second theorist to you, Bryan Pfaffenberger, an anthropologist by training, working in the field of Science and Technology Studies ("STS"). His seminal article, "Technological Dramas," offers an approach to understanding how technologies evolve in response to social and political factors.[27] Pfaffenberger asks the same question as Winner—"Do artifacts have politics?"—but in contrast to Winner, answers a resounding "No." Now I do not entirely agree with Pfaffenberger because I do believe that values may be embedded in technical systems due to specific material characteristics. But Pfaffenberger's challenge to Winner, at the very least, forces us to acknowledge that the processes by which technology comes to embody values, or comes to have the power to regulate, are complex and, at times, even indeterminate. Let us consider the argument a bit more closely.

As I see it, there are two main contentions forming Pfaffenberger's argument. One establishes that as technologies evolve—whether a simple pen or a complex network, such as the Internet—there is always flexibility in its design: there are many inflection points and a variety of paths a designer could have chosen, arriving not at something that looks like *this* but *that*.[28] With information technology, because systems can be developed and changed quickly, one may actually observe this evolution over a relatively short period of time, and experience shows how much flexibility there is in the design of digital artifacts. The second contention highlights technology's interpretative flexibility: when you introduce a (technological) system into a society, it does not arrive fully thought out. Without interpretation it may not even be immediately understood.[29] As symbolic actors, we need to—and these are the particular terms Pfaffenberger uses—"regulate" technology discursively, drawing on some of our familiar cultural mythologies and secular rituals to establish what it is and what it can do. This discursive "regularization" is one of the processes that establish the political aims of a technology.[30] In Pfaffenberger's words, "artifacts [] are projected into a

27. Bryan Pfaffenberger, *Technological Dramas*, 17 SCI. TECH. & HUMAN VALUES 282 (1992).

28. *Id.* at 283; Trevor J. Pinch & Wiebe E. Bijker, *The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other*, *in* THE SOCIAL CONSTRUCTION OF TECHNOLOGICAL SYSTEMS 17 (Wiebe E. Bijker et al. eds., 1989).

29. Pfaffenberger, *supra* note 27, at 285.

30. *Id.* at 291.

spatially defined, discursively regulated social context, which is crucial to actualizing the technology's constructed cultural and political aims."[31]

According to Pfaffenberger's theory of technological dramas, these aims are not built in to a technical artifact itself; instead, they are shaped by the accompanying interpretations given to it by the technical artifact's "design constituency"—the group of people introducing an artifact into society. In the process of regularization, the design constituency promulgates a particular culturally and politically laden interpretation by structuring the discursive regulation.[32] For the design constituency, this process involves not only explaining how a technology links into and satisfies certain cultural or symbolic needs that people have, but also actively intervening to adjust the legal and legislative landscape to make it hospitable to the system in question.[33]

You may well wonder why the design constituency must seek these adjustments in the landscape. For me, the answer to this question constitutes Pfaffenberger's core insight: the reason they do it is to define away alternatives, to define away the design and interpretative flexibility; the design constituency wants us to see the technology in a certain way and only in that way.[34] And it is through this process, according to Pfaffenberger, that technology has politics. Armed with these concepts, we may return to the terms of a dichotomy I introduced earlier—namely, how it comes to be that certain artifacts are made to seem not merely like challenges that must be overcome, but like impenetrable barriers that must be accepted as they are. Where the regularizing activities of design constituencies (or others) are successful, they manage to define away the alternatives. When it happens that a design constituency manages to regularize a particular meaning for a given technical system, Pfaffenberger would say that the constituency has asserted "logonomic control" over it.[35] Finally, it is through the exercise of logonomic control and not by dint of material features that artifacts become political because, by then, the political aims will be experienced as inevitable, irresistible.[36]

There is more to technological drama than regularization; it is merely a first move, the first act, if you will. It is hardly ever all there is. Pfaffenberger admits that even for a determined design constituency, it is difficult to unite

31. *Id.*
32. *Id.*
33. *Id.* at 294.
34. *Id.* at 295.
35. *Id.* at 296.
36. *Id.*

whole-heartedly behind one unambiguous account of a technology's symbolic meaning—there is always going to be some symbolic ambiguity, and out of this ambiguity the opponents of a particular system and the values expressed through it can stir discontent. Because cultures are so often a riot of beliefs and mythologies, opponents of a design constituency's interpretations of a particular system may find conflicting cultural norms amidst ambiguous root paradigms from which to construct competing interpretations. In this second act, called "adjustment," the politicization of a technology shifts course. An example Pfaffenberger uses is an entryway bench provided by the upper classes in Victorian Britain for servants, who were generally disenfranchised through uncomfortable living conditions and humiliated by deliberately poorly designed furniture. The servants "adjusted" the meaning of these benches by drawing on a competing cultural norm around compassion and the common good, which facilitated solidarity among themselves and pity for employers whom they saw as needlessly cruel and insecure.[37]

A third act in Pfaffenberger's technological dramas, "reconstitution," involves not merely reinterpreting a given system but also redesigning it materially and, potentially, beginning a drama all over again.[38] This redesigned system, which Pfaffenberger calls a "counterartifact," is given its cultural meaning and values by those who before were part of the impact constituency, that is, those negatively affected by the original system. The new drama involves an accompanying societal discourse aimed at earning acceptance for the counterartifact and its associated mythology. Accordingly, "[r]egularization can indeed become a tool of reconstitution; it can be used to enforce change as well as continuity."[39]

Now if you, like me (and Langdon Winner), stubbornly resist the idea that the material character of an artifact does not matter at all to its political character and that political values can be fully shaped by adjusting symbolic context alone, then you will welcome reconstitution as evidence that not even Pfaffenberger holds that interpretative flexibility knows no bounds. Sometimes design alterations are necessary to complement the political aims of supporting mythologies.[40] How much of the politics inheres in material design and how much in discourse and interpretation is the heart of much philosophizing about technology and is too big an issue to resolve here. My

37. *Id.* at 301.
38. *Id.* at 304.
39. *Id.*
40. *Id.*

aim, thus far, has been to present one plausible account of these intertwined influences.

There is still one more act in Pfaffenberger's technological drama, you could say, the final act: closure, or normalization (or "designification").[41] It is the phase when competing discourses are no longer present in the public's attention, and the politics of a technical system recede from consciousness. Whatever public controversy has played out in the public sphere has gone silent for the time being. (It can, of course, flare up again with new dramas following.) During this phase, an artifact, like a mountain or a tree, has become something "natural" or, in Pfaffenberger's words, "the drama [] drop[s] out of the technology."[42] This, in some sense, is a dangerous phase, when people are inclined to accept that technology is neutral because it is when people forget that there are values or politics involved in technology at all. It is not that there is no longer politics in these normalized systems, but merely that we are no longer attuned to it; we forget that political values are still there. To counteract such losses in astuteness, Pfaffenberger recommends STS as the "political philosophy of our time" so that we may continue to advance our understanding of the ways technology is linked to fundamental principles and values of ethics and politics.[43]

VII. IF TECHNOLOGY REGULATES, WHY DO WE NEED REGULATION?

Having armed ourselves with insights from Pfaffenberger's technological dramas, we return to consider the role of regulation in relation to technologies that, so to speak, regulate. As illustrations, I will refer to a couple of cases that are quite well known, cookies and the Digital Millennium Copyright Act ("DMCA"), the latter far better known to many of you in this audience than to me! Specific inflection points in each of these cases, understood in Pfaffenberger's terms, reveal fascinating variations in the relationship between these two regulators.

A. DMCA AND ANTI-CIRCUMVENTION

As we know, the DMCA, which came into law in 1998, followed fraught multi-year, multi-national deliberations. In large part, its passage was driven by a combination of radical changes due to digital information technologies in the creation, use, and distribution of creative content, and an ensuing

41. *Id.* at 308–09.
42. *Id.* at 308.
43. *Id.* at 309.

panic in the content industry due to mortal threats these changes seemed to pose to established business models. Within the big picture, the inflection point of greatest interest to me here is the DMCA's anti-circumvention clause, urged by the content industry as a necessary antidote to vulnerabilities in technology-based copyright protection measures. I have yet to hear an expert in the field who is willing to say that any given technological protection measure ("TPM") or digital rights management ("DRM") system cannot be broken.

In Pfaffenberger's terms, the drama begins with the industry attempting to regulate, that is, prevent unauthorized copying of content by means of TPMs. The industry, or in this case the design constituency, introduces these systems, buttressing them with a rich discourse about intellectual property—supporting creativity, protecting their copyrights, preventing piracy, and so forth. The trouble is, this is not enough because TPMs are not tamper-proof and a large segment of the impact-constituency is not swayed by the mythos. In other words, there is interpretative flexibility and a struggle over which version will prevail. Many proponents of peer-to-peer file sharing systems, such as Gnutella and BitTorrent, and circumvention software, such as DeCSS, reject the discourse of the likes of the content industry; they see TPMs as obstacles, but obstacles that must be challenged, overcome.[44] The content industry, however, wants us to see TPMs not as challenges to be overcome but as impenetrable barriers. Since the technology is not itself impenetrable, the industry must find other ways to, in Pfaffenberger's terms, define away technical alternatives; this they have achieved through the DMCA's anti-circumvention clause.

I find the concept of a "handoff" useful. To begin, it was the law of copyright that constrained people's behaviors. When digital technologies radically loosened the hold of copyright, defenders of intellectual property rights turned to technical means—that is, they handed off power to technology. When technology proved imperfect, there was another handoff, again, back to law and regulation, not directly to constrain behavior as with the law of copyright itself, but to shape how people saw, understood, and interpreted prevailing TPMs.

As we know, the drama is not yet over, even though, in my view, systems such as iTunes are close to normalized. There are continued efforts to *adjust*

44. *Cf.* Edward W. Felten, *DRM and Public Policy*, 48 COMM. ACM 112 (2005); Cory Doctorow, *Pushing the Impossible*, GUARDIAN (Sept. 4, 2007, 2:10 PM), http://www.guardian.co.uk/technology/2007/sep/04/lightspeed; Fred von Lohmann & Wendy Seltzer, *Death by DMCA*, IEEE SPECTRUM (June 2006), http://spectrum.ieee.org/computing/software/death-by-dmca.

our interpretation of TPMs and associated anti-circumvention measures. Some, such as the Electronic Frontier Foundation, portray them not as defensible protectors of property rights but as violators of other rights, compiling stories of economic calamity, injustice for people trying to use digital content legally, and problems with market competition.[45] Ed Felten, another in this vein, defied the music industry's ban on the publication of his method for breaking the industry's digital watermarking technologies of the day. The industry cited anti-circumvention, Felten cited academic freedom and the "freedom to tinker."[46] And the cycle continues!

B. COOKIES

The development of HTTP cookies is another case that can be illuminated by the notions of technological dramas. In contrast with our previous case, it illustrates the perils of believing one can leave all regulation to technology alone. According to the standard account, which we will accept here, Lou Montulli created web cookies in order to facilitate shopping carts, and shopping, on the web.[47] Because the web is stateless, without cookies, each time you would visit a particular website—in fact, each time you would send a command to this website—these actions would be treated as distinct, unconnected events. The exchange of cookies, simple data-objects, back and forth between browser and website enables a continuity in the relationship. In 1995 the cookie was integrated into the Mosaic and Microsoft Internet Explorer browsers, and in 1997 it was introduced as a standard in Request for Comments ("RFC") 2109 to the Network Working Group ("NWG").[48]

In RFC 2109, the original design specification for the cookie ensured that cookies generated by a particular website could only be retrieved by that website. Thus, when a person revisits a given website, that website could retrieve only the cookie that it had placed in the person's browser. Montulli and his colleague David Kristol made their intention clear when they wrote in the RFC, "The intent is to restrict cookies to one, or a closely related set of hosts. . . . We consider it acceptable for hosts host1.foo.com and host2.foo.com to share cookies, but not a.com and b.com."[49] So the intent of

45. ELEC. FRONTIER FOUND., UNINTENDED CONSEQUENCES: TEN YEARS UNDER THE DMCA (2008), *available at* http://www.eff.org/files/DMCAUnintended10.pdf.

46. *Id.* at 2.

47. *Cookies*, STUDIO 360 (Dec. 17, 2010), http://www.studio360.org/2010/dec/17/cookies/.

48. David Kristol & Lou Montulli, *HTTP State Management Mechanism*, INTERNET ENG'G TASK FORCE (Network Working Grp., Request for Comments No. 2109, Feb. 1997), *available at* http://tools.ietf.org/pdf/rfc2109.pdf.

49. *Id.* at 16.

the original designers of the cookie was quite clearly to build the value of user privacy into cookies from the beginning of their use online. But by 1997, the advertising industry had already figured out ways to circumvent Montulli and Kristol's restriction on cookie exchange. In that year, articles in the business press were trumpeting DoubleClick's workaround, which involved dropping their own cookies into people's browsers by attaching them to ad images incorporated into the websites that people visited.[50] Thus was born the so-called "third-party" cookie that was able to follow people around from site to site. In head-to-head competition with Montulli and Kristol, the advertising industry-backed competing standard RFC 2965, which allowed placement of third-party cookies, was ultimately victorious.[51] The rest, as they say, is history.

In hindsight and in light of Pfaffenberger's dramas, I am inclined to say that while Montulli and Kristol's RFC 2109 embodied the value of privacy into web cookie design, there was no accompanying attempt to "regularize" it—that is, to discursively regulate the social context into which this design proposal was being projected. I would like to think that had we—"we" being those of us who would have liked this standard to win out—done the discursive groundwork, things might have turned out differently. In contrast with the DMCA case, there was no design constituency equivalent to the content industry undertaking the challenge of defining away alternatives. There was no defender of the original standards and practices to champion a web cookie anti-circumvention clause before the online advertising industry, with an urgent and vested interest, hijacked momentum and succeeded in regularizing a different technical standard. Things could have turned out differently; had the standard defined in RFC 2109 won out, been "regularized," and been buttressed with supporting anti-circumvention regulation, the prohibition on third-party cookies would have seemed an impenetrable barrier. As it happened, however, alternatives were not successfully defined away, and the standard was treated as a mere challenge to overcome.

At this time, with interest in a Do Not Track option for the web percolating in Congress,[52] the Federal Trade Commission,[53] and the

50. Judith Messina, *New Media's Hot Play*, 13 CRAIN'S N.Y. BUS., no. 24, June 16, 1997, at 1.

51. David Kristol & Lou Montulli, *HTTP State Management Mechanism*, INTERNET ENG'G TASK FORCE (Network Working Grp., Request for Comments No. 2965, Oct. 2000), *available at* http://tools.ietf.org/pdf/rfc2965.pdf.

52. Do-Not-Track Online Act of 2011, S. 913, 112th Cong. (2011), http://www.govtrack.us/congress/bill.xpd?bill=s112-913.

Department of Commerce,[54] there is reason to hope that political controversies that were settled with the selection of a cookie standard may be revisited. If so, the privacy community, now far more extensive and better organized, will have an opportunity to articulate adjustment strategies (perhaps including regulation) and develop counter-technologies that were missing in the late 1990s. There is evidence that both are occurring; the struggle is on.

VIII. TRACKMENOT, ADNOSTIC, AND THE POLITICS OF PRIVACY

We have come full circle back to TrackMeNot and Adnostic, systems that embody the value of privacy. As you may recall, these systems were inspired by the aim of saying and doing something political through technology, particularly in domains where there seemed to be foot-dragging and resistance from policy makers both in the private and public sectors. If Winner is correct and we can say that these systems have politics, or in other words, can assert that they regulate by affording and constraining behaviors in politically relevant ways, have we done enough? What role is left for regulation, per the question of my title? I hope to have shown that indeed, beyond design, there is work to do to regulate discursive conditions so they are hospitable to one's political ends. Law and regulation are an important part of both practically and symbolically preparing the environment so that a technical system projected into it can do its work—and sometimes that work is political. We cannot take it for granted that with clever enough design, TMN, Adnostic, or anything else will enter the scene and protect privacy.

Drawing wisdom from Pfaffenberger's dramas, cultural mythologies and secular rituals are rich sources for the creation of discourses that can serve to regularize a system. Experience with TMN is consistent with Pfaffenberger's claims that for systems to have politics—that is, to function politically in society—more than design is needed. In the case of counterartifacts, in particular, this will mean looking for cultural beliefs that contradict those buttressing entrenched systems. If search engines and ad networks draw on cultural mythologies of individualism to support their particular brand of personalization, opponents can tap into the same pool of cultural

53. FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE (2011), *available at* http://www.ftc.gov/os/2010/12/101201privacyreport.pdf.

54. INTERNET POL'Y TASK FORCE, U.S. DEP'T OF COMMERCE, COMMERCIAL DATA PRIVACY AND INNOVATION IN THE INTERNET ECONOMY: A DYNAMIC POLICY FRAMEWORK (2010), *available at* http://www.ntia.doc.gov/files/ntia/publications/iptf_privacy_greenpaper_12162010.pdf.

mythologies but focus on stereotyping and unfair discrimination, forms of illegitimate personalization that impinge on individual freedoms. There are other powerful cultural beliefs surrounding a right to freedom in educational, intellectual, informational, and political pursuits that, too, could lend support to adjustment and reconstitution of entrenched discourses of personalization. Mobilizing this symbolic discourse and placing web search within it could contribute to an environment that is more welcoming to the political aims of systems such as TMN.

In the effort to gain a toehold for Adnostic, technical functionality is not the greatest barrier. We have found ourselves up against a cultural mythology of innovation, incredibly powerful in the context of the Internet and web. In projecting a new system into this context, one should avoid at all cost being viewed as going against it. So powerful is this context that it succeeds against just about any effort to develop regulation restricting what established actors can do either in the systems they develop or the policies guiding the application of these systems. In seeking to protect privacy, any inclination that agencies such as the Federal Trade Commission and the Department of Commerce may have to develop regulation must clear the great hurdle of innovation. It seems hardly to matter that the clarion call of innovation seems oddly stacked in favor of industry-based innovation and is so vaguely defined that it amounts more to a free-for-all than a serious guidepost.

With Adnostic, we have tried to pick holes in the logic of the online advertising industry. They say that personalized ads support robust commerce online as well as both technical and business innovation. Our answer is that Adnostic still offers personalization (i.e., targeting) of ads but without third-party tracking and the vast and sinister apparatus of surveillance that goes along with it. Thus far, logic has not worked in our favor, and we have not yet interested regulatory bodies in putting pressure on ad networks to enable individuals to choose Adnostic as a medium for ad presentation. We may need to go looking for the ambiguities in root paradigms, to pit the halo of innovation that industry seems to have cornered to something equally powerful. Perhaps it will be to encourage a norm of equal opportunity to innovate, or to promote the idea that your web browser is *your* agent, not an infiltrator on your system working in someone else's service.

IX. A ROLE FOR LAW AND REGULATION

So that is the discursive realm of cultural myth and secular ritual. What might be a role for law in contributing to a hospitable environment into which systems for protecting privacy are projected? Clearly, laws that would directly regulate what can and cannot be done with data in specific areas,

such as web searches or online tracking, would set parameters in ways that are favorable to such efforts. Or, even more ambitiously, we may try to resuscitate a general privacy law applicable to private as well as public sectors. If successful, the efficacy of systems designed to protect privacy would be enhanced (just as laws against burglary enhance the efficacy of locks). My colleague Ira Rubinstein believes we might actually pass a law, but other colleagues who work in this area of regulation do not see much cause for optimism.[55]

Short of these, what else might contribute to a hospitable environment? I have often thought that legal recognition of something like "respect for expressive choice" holds promise. It might work as follows: the law would recognize an expressed desire not to be tracked for those who have installed TMN on their browsers. Within the technical community, where the favored means of protection is strong encryption, there is skepticism over whether TMN "works," by which they mean whether it could withstand a concerted attack by an entity determined to filter out the fake, TMN-generated queries. Giving legal recognition to those who have installed TMN as an expressed desire not to be profiled is a bit like an anti-circumvention clause for privacy in web search. In such cases, the law plugs holes that technology leaves open; it, to quote Pfaffenberger, defines away alternatives.[56] Recognizing expressive choice is also consistent with one element of the expectation of privacy test because adoption of a system such as TMN clearly indicates an actual expectation of privacy on the part of a user (the other element being that this expectation is reasonable).

A second role for law that falls short of privacy regulation but may create a more hospitable environment for the private development and uptake of privacy protective systems is to stipulate limits on the Terms of Service, perhaps by affording greater latitude to individuals to negotiate terms that allow them greater control over their relationship with online entities. At present, the pressure on users to agree to lengthy or unclear Terms of Service documents can be a one-click "flank attack" on rights we ought to retain in relation to websites we visit. In the early days of TMN, there was considerable debate over whether it was illegal. I have to admit, I was actually quite afraid for awhile—likely without cause—because TMN did violate some of the search engines' Terms of Service agreements, which forbid users from initiating automatically-generated searches. This net, which was probably cast to prevent denial-of-service attacks, also captured TMN.

55. Ira S. Rubinstein, *Privacy and Regulatory Innovation: Moving Beyond Voluntary Codes*, I/S: J.L. & POL'Y INFO. SOC'Y (forthcoming winter 2011).

56. Pfaffenberger, *supra* note 27, at 295.

Although no angry search companies have come after us, we should not have to depend on another party's largesse to develop or use privacy enhancing technologies.

In preparing for this talk, the distinction between obstacles that are viewed as challenges and obstacles that are viewed as impenetrable barriers was an exciting discovery for me as I sought to understand why technology cannot regulate all by itself, even if we allow that it can regulate at all. (Or, in other words, why politics is needed even if we allow that technology itself has politics.) However well-designed, well-executed, and well-fortified our praiseworthy, value-enhancing systems are, incipient weaknesses are inevitable and pose a threat to their programmed action. In these cases, an important role for regulation is to remove the temptation to exploit these weaknesses. As such, regulation contributes to a view of these systems as impenetrable barriers rather than challenges to work around.

I will stop here and leave you with that thought.

Thank you.

SEEING THE FORESTS AND THE TREES: TECHNOLOGICAL AND REGULATORY IMPEDIMENTS FOR GLOBAL CARBON MONITORING

Molly K. Macauley[†] *& Nathan Richardson*[††]

TABLE OF CONTENTS

I. INTRODUCTION

Forests play a major role in Earth's carbon cycle and have the potential to play an equally significant role in any national or global policy to reduce net carbon emissions and the risks of climate change. The chief vehicle advanced for incorporating forests into carbon mitigation policy is the use of international carbon offsets, by which reductions in net carbon emissions in relatively low-cost regions can be used in lieu of similar reductions in higher

† Senior Fellow and Vice President for Research, Resources for the Future, Washington, D.C.

†† Resident Scholar, Resources for the Future, Washington, D.C.

The financial support of the U.S. Geological Survey and the Alfred Sloan Foundation for portions of the research in this Article is greatly appreciated. Responsibility for errors and opinions rests exclusively with the authors.

cost regions. The costs to the U.S. economy of reducing greenhouse gas ("GHG") emissions would be reduced by half or, in some scenarios, by more than half by managing forests to exploit their natural storage of carbon.[1]

Although international forest carbon offsets could be financially beneficial, two problems continue to limit their use. The first is a lack of data. Perhaps surprisingly in an age when detailed neighborhood maps are available at a touch on smartphones, there is no map with much detail about the world's forests. Accurate measures of extant forest acreage and the capacity to monitor changes in acreage are necessary for carbon markets, just as countable units are required for market exchange of any commodity.

The second problem is legal. It appears likely that the U.S. Environmental Protection Agency's ("EPA") tools under the existing Clean Air Act ("CAA")[2] will be the primary means of limiting carbon emissions from most sectors of the economy in the near future. Under the CAA, the EPA has proposed and finalized rules for reporting GHG emissions and announced steps to regulate emissions from mobile sources (cars and trucks) and new or modified stationary sources (power plants and industrial facilities). The agency is next expected to regulate existing stationary sources (like fossil-fuel power plants and petroleum refineries). Such measures could achieve meaningful reductions in U.S. carbon emissions.[3] Unfortunately, the CAA, as it stands, is likely incompatible with the use of international offsets.[4]

---

1. U.S. ENERGY INFO. ADMIN. (EIA), SR/OIAF/2009-05, ENERGY MARKETS AND ECONOMIC IMPACTS OF H.R. 2454, THE AMERICAN CLEAN ENERGY AND SECURITY ACT OF 2009, at xiv (2009), *available at* http://www.eia.doe.gov/oiaf/servicerpt/hr2454/pdf/sroiaf(2009)05.pdf; Cong. Budget Office (CBO), *The Use of Offsets To Reduce Greenhouse Gas Emissions* 7 (Aug. 3, 2009), http://www.cbo.gov/ftpdocs/104xx/doc10497/08-03-Offsets.pdf; U.S. Envtl. Prot. Agency (EPA), Appendix to EPA Analysis of the American Clean Energy and Security Act of 2009 H.R. 2454 in the 111th Congress 64–65 (June 23, 2009), *available at* http://www.epa.gov/climatechange/economics/pdfs/HR2454_Analysis_Appendix.pdf.

2. Clean Air Act, ch. 360, 69 Stat. 322 (1955) (codified as amended at 42 U.S.C. §§ 7401–7671q (2006)).

3. *See* Nathan Richardson et al., *Greenhouse Gas Regulation Under the Clean Air Act: Structure, Effects, and Implications of a Knowable Pathway*, 41 ENVTL. L. REP. 10098, 10099–115 (2011) (stating that the EPA can achieve emissions reductions via performance standards, and detailing specifics for the coal energy sector); *see also* Dallas Burtraw et al., *Greenhouse Gas Regulation Under the Clean Air Act: A Guide for Economists*, 5 REV. ENVTL. ECON. & POL'Y 293, 299–301 (2011), *available at* http://reep.oxfordjournals.org/content/5/2/293.full.pdf (summarizing evidence for emissions reductions available via EPA regulation from various sectors, and suggesting that a 10% overall reduction in US emissions is plausible).

4. *See generally* Nathan Richardson, *Playing Without Aces: Offsets and the Limits of Flexibility Under Clean Air Act Climate Policy* (Resources for the Future, Discussion Paper No. 11-49, 2011), *available at* http://www.rff.org/RFF/Documents/RFF-DP-11-49.pdf; *see also* Nathan

Superficially, this incompatibility appears to reduce the role that forests can play in U.S. climate policy. As a result, the United States is poised to undertake an unnecessarily expensive approach to GHG management, one that overlooks a relatively low-cost means of sequestering carbon.

Part II of this Article describes the science of forests in the global carbon cycle and the economic benefits of sequestration in offsetting greenhouse gas emissions. Part III describes the problem of forest measurement and the technical means necessary for physical inventory of forests. Part IV discusses forest carbon offsets under the CAA. Part V deals with implications for innovation and international diplomacy with respect to forest carbon offsets and suggests areas for future research.

II. THE SCIENCE AND ECONOMICS OF FOREST CARBON SEQUESTRATION

The global carbon cycle is made up of atmospheric, oceanic, and terrestrial processes that circulate and filter carbon, methane, fluorinated gases, and other natural and anthropogenic emissions.[5] Forests are a quantitatively significant link in two of these processes. Forests store carbon by taking in carbon dioxide from the atmosphere during respiration; trees draw the carbon atoms into the plant cell and release oxygen back into the atmosphere.[6] By contrast, when forests are removed (for purposes of agricultural production, development, or other uses) or damaged (by wildfire, pests, drought, or other occurrences), carbon is released (though some portion remains stored in lumber, furniture, and other timber products). Trees are particularly efficient at storing, or sequestering, carbon. Estimates of carbon emissions from forest removal range from seven percent to thirty percent of all GHG emissions.[7] Because changes in forests have the potential to significantly impact those emissions, maintaining intact forests and adding

Richardson, *International Greenhouse Gas Offsets Under the Clean Air Act*, 40 ENVTL. L. REP. 10887 (2010).

5. *See generally* SUSAN SOLOMON ET AL., CONTRIBUTION OF WORKING GROUP I TO THE 4TH ASSESSMENT REPORT OF THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (2007).

6. *See id.* at 1–40.

7. R.A. Houghton & S.J. Goetz, *New Satellites Help Quantify Carbon Sources and Sinks*, 89 EOS TRANS. 417 (2008); *see also* G.R. van der Werf et al., Commentary, *CO_2 Emissions from Forest Loss*, 2 NATURE GEOSCIENCE 737 (2009), *available at* http://www.biology.duke.edu/jackson/ng09.pdf (stating that carbon emissions from deforestation and forest degradation account for about 20% of global anthropogenic CO_2 emissions).

to forest stocks through reforestation and afforestation confers a benefit in the form of carbon sequestration.[8]

The U.S. Energy Information Administration ("EIA") has estimated the cost to the U.S. economy of managing greenhouse gases through actions such as reducing power plant and vehicle emissions, changing forest management, and other steps. The estimates are presented in various economic terms, including in absolute dollar amounts and as a percentage of the nation's gross domestic product ("GDP"). The annual value of forest carbon offsets in reducing expenditures that would otherwise need to be made (say, instead of reducing power plant emissions) could reach $60 billion annually by 2030.[9] Without the use of forest offsets, U.S. emissions reductions will have to largely come from the domestic electricity and transportation sectors. Expressed another way, in terms of the value of the reduction in U.S. GDP, failing to make use of forest offsets would increase the loss in GDP from 0.2 percent to 0.3 percent over the period from 2012 to 2030.[10]

One of the biggest hurdles to using forest carbon offsets—and reaping their financial benefits—is that they may be difficult to measure accurately. The EIA emphasizes that its estimates are based on the assumption that physical forest carbon sequestration capacity can be accurately measured over time and around the world. The agency further emphasizes, and other experts agree, that this assumption may be too strong, underscoring concerns about whether adequate measurement and monitoring is in fact attainable.[11]

---

8. Other concurrent benefits of forests include their role in watershed protection and habitat and biodiversity conservation. Note that these attributes are not necessarily correlated with forest carbon storage; in other words, carbon rich forests are not necessarily species rich. *See* Erin Myers Madeira & Juha Siikamäki, *Progress and Challenges for Forests in Climate Policy—Seeing REDD*, *in* CLIPORE, ANNUAL REPORT 2009, at 17 (2010), *available at* http://www.clipore.se/download/18.2a759bb41277b00e3c380001166/Annual+Report+2009-6.pdf.

9. EIA, *supra* note 1, at xi (multiplying domestic and international offset quantities by estimated domestic and international offset prices for 2030).

10. The costs reported here are discounted because they accrue over time. *See* EIA, *supra* note 1, at xiv, 40.

11. EPA, *supra* note 1, at 53; ROSS W. GORTE & JONATHAN L. RAMSEUR, CONG. RESEARCH SERV., RL34560, FOREST CARBON MARKETS: POTENTIAL AND DRAWBACKS 15–18 (2010), *available at* http://www.cnie.org/NLE/CRSreports/10Jun/RL34560.pdf; PERVAZE A. SHEIKH & ROSS W. GORTE, CONG. RESEARCH SERV., R40990, INTERNATIONAL FORESTRY ISSUES IN CLIMATE CHANGE BILLS: COMPARISON OF PROVISIONS OF S. 1733 AND H.R. 2454, at 13 (2009), *available at* http://www.nationalaglawcenter.org/assets/crs/R40990.pdf; CBO, *supra* note 1, at 4–6. The U.S. Governmental Accountability Office has also written about this problem. *See* U.S. GOV'T ACCOUNTABILITY OFFICE (GAO), GAO-08-1048, CARBON OFFSETS: THE U.S. VOLUNTARY MARKET IS GROWING BUT QUALITY

At present, no global measurements of forests meet the desired accuracy; in fact, current measures fall so far short that few meet existing federal guidelines for voluntary carbon management (undertaken by industries wishing to reduce their carbon footprint) or voluntary carbon exchanges (the precursors to actual carbon markets).[12] The U.S. Government Accountability Office notes that "ensuring the credibility of carbon offsets poses challenges because of the inherent uncertainty in measuring emissions reductions or sequestration relative to a projected business-as-usual scenario."[13] Legislation passed by the U.S. House of Representatives (H.R. 2454) and proposed in the U.S. Senate (S. 1733) limited the total number of offset credits and discounted the purchase of international offsets by requiring companies to buy 1.25 international offsets for one domestic offset credit. As further evidence of the concern about capacity to measure and track offsets, both the House and Senate provisions established an Offset Integrity Advisory Board.[14] As described in the Senate provisions, the Board would establish "methodologies to address the issues of additionality, activity baselines, quantification methods, leakage, uncertainty, permanence, and environmental integrity."[15]

Several measurement issues are of particular concern: the requirement for better baseline estimates and the need to monitor changes in the baseline over time. Both measures help scientists ascertain whether, in fact, atmospheric concentrations of GHGs appear to be stabilizing. Both measures would also be required to satisfy regulators and other parties using forests to offset greenhouse gas emissions from other sources. The two concerns with respect to these measurements are known as leakage and permanence, and both problems could prevent a forest carbon market from functioning effectively to offset GHG emissions. Leakage refers to reduced deforestation in one area that drives deforestation to another area. Forecasts of how much forested area worldwide may be protected for sequestration—which are needed to ascertain if enough carbon sequestration is taking place—may be incorrect if it is assumed that no leakage occurs. Murray et al.

ASSURANCE POSES CHALLENGES FOR MARKET PARTICIPANTS 7–9 (2008), *available at* http://www.gao.gov/new.items/d081048.pdf.

12. *See* MOLLY MACAULEY ET AL., RESOURCES FOR THE FUTURE, FOREST MEASUREMENT AND MONITORING: TECHNICAL CAPACITY AND 'HOW GOOD IS GOOD ENOUGH?' 17–20 (2009), *available at* http://www.rff.org/rff/documents/rff-rpt-technical%20capacity_macauley%20et%20al.pdf (discussing the lack of capacity to meet existing guidelines and the standards set by voluntary markets).

13. GAO, *supra* note 11, at 37.

14. H.R. 2454, 111th Cong. § 731 (2009); S. 1733, 111th Cong. § 731 (2009).

15. S. 1733, § 731.

estimate the potential for leakage at ten to ninety percent in the United States.[16] Without adequate monitoring of forests in all countries throughout the world, leakage could undermine efforts to stabilize GHG emissions.[17]

Additionally, changes in forests from logging, conversion to agriculture, or disturbances such as wildfires and drought affect the long-term physical capacity of forests to store carbon. Monitoring these changes is another component of forest carbon sequestration as an element of greenhouse gas management. Some forest carbon management proposals assume that forests would be rented to account for the possibility of their impermanence.[18] In short, measuring forest sequestration is hard—with leakage and sequestration making it even harder.

III. THE TECHNOLOGY OF MEASURING AND MONITORING FORESTS

This Part briefly describes why improving the technology for measuring forests requires the deployment of new technology. Although necessary, institutional and financial constraints limit deployment of new technology. These limits have also led to inaccuracies in the forest measurement data that are now available.

Forest measurement requires a number of steps. The procedure is allometric, and is based on forested land area, the growing volume (height) of the trees, and their biomass, which is determined largely by the tree species and overall tree health. High quality data on these variables have been available only in a small number of places that predominantly fall into two categories: forests that are managed for commercial timber; or small areas where field work has taken place, usually for research on the use of forests for fuelwood, watershed management, or poverty alleviation in developing countries.[19] The only worldwide inventory of forests consists of voluntary,

16. *See* Brian C. Murray et al., *Estimating Leakage from Forest Carbon Sequestration Programs*, 80 LAND ECON. 109, 109 (2004).

17. *See id.*

18. *See* Man-Keun Kim et al., *Permanence Discounting for Land-Based Carbon Sequestration*, 64 ECOLOGICAL ECON. 763 (2008).

19. *See, e.g.*, Sandra Brown & Barbara Braatz, *Methods for Estimating CO_2 Emissions from Deforestation and Forest Degradation*, *in* GOFC-GOLD, SOURCEBOOK: A SOURCEBOOK OF METHODS AND PROCEDURES FOR MONITORING AND REPORTING ANTHROPOGENIC GREENHOUSE GAS EMISSIONS AND REMOVALS CAUSED BY DEFORESTATION, GAINS AND LOSSES OF CARBON STOCKS IN FORESTS REMAINING FORESTS, AND FORESTATION, ch. 2.4, at 2-72 (Report No. COP16 ver. 1, 2010), *available at* http://www.gofc-gold.uni-jena.de/redd/sourcebook/Sourcebook_Version_Nov_2010_cop16-1.pdf [hereinafter GOFC SOURCEBOOK]

self-reported data sent roughly every five years by countries to the United Nations Food and Agricultural Organization ("FAO") for the FAO's Forest Resource Assessment. Inventory practices for the FAO reports vary widely among countries. Many countries differ in their definitions of "forested land." Some extrapolate forested acreage from a sample of field measures and others use state-of-the-art instruments on airplanes to map the height of trees and forested acreage, thus obtaining high quality measures of carbon sequestration. Developing countries, including many that are thought to be rich in forest carbon, often lack measurement capacity altogether. Some countries with large acreages of boreal forests, such as Canada and Russia, use altogether different measurement techniques for estimating forested lands than many other countries, making comparisons even more difficult.[20] The FAO itself acknowledges these problems,[21] which result in wide discrepancies in reported measures compared with actual field data.[22]

Improving forest measurement on a global scale is technically feasible. Specialized remote sensing instruments carried on aircraft or satellites could provide highly accurate, well-calibrated, and spatially consistent measures.[23] At present, aircraft instruments are deployed in only a few places and would require massive deployment for global coverage. Furthermore, the unique vantage point of space satellites, coupled with the fact that they generally cover the same location on the Earth's surface every few days, makes them an especially good choice to meet the requirement of such coverage. Satellites can readily collect measurements globally over time; the repeated coverage of the same location every few days would allow for monitoring of forests, including monitoring of leakage, permanence, and degradation. When

(discussing measurements taken to understand the supply of fuelwood provided by a forest in a developing country).

20. MACAULEY ET AL., supra note 12, at 8.

21. *See* EMILY MATTHEWS & ALAN GRAINGER, UNITED NATIONS FOOD & AGRIC. ORG. (FAO), EVALUATION OF FAO'S GLOBAL FOREST RESOURCES ASSESSMENT FROM THE USER PERSPECTIVE (2002), *available at* http://www.fao.org/docrep/005/y4001e/Y4001E07.htm.

22. *See, e.g.*, Lloyd C. Irland, *Assessing Sustainability for Global Forests: A Proposed Pathway To Fill Critical Data Gaps*, 129 EUR. J. FOREST RES. 777 (2009); *see also* Paul Waggoner, *Forest Inventories: Discrepancies and Uncertainties* (Resources for the Future, Discussion Paper No. 09-29, 2009), http://www.rff.org/documents/RFF-DP-09-29.pdf.

23. *See* MATTHEW FAGAN & RUTH DEFRIES, RESOURCES FOR THE FUTURE, MEASUREMENT AND MONITORING OF THE WORLD'S FORESTS: A REVIEW AND SUMMARY OF TECHNICAL CAPABILITY, 2009–2015 (2009), *available at* http://www.rff.org/rff/documents/rff-rpt-measurement%20and%20monitoring.pdf; GOFC SOURCEBOOK, *supra* note 19; Molly Macauley & Roger Sedjo, *Forests in Climate Policy: Technical, Institutional and Economic Issues in Measurement and Monitoring*, 16 MITIGATION & ADAPTATION STRATEGIES FOR GLOBAL CHANGE 499 (2011).

collected routinely over many years, these data provide a time series to inform baselines and track changes of forests.

Several impediments prevent deployment of aircraft and satellite technologies despite their technical feasibility. First, in the absence of a market for carbon, there is no mechanism to privately finance aircraft or satellite data collection. Without a climate policy that allows inclusion of forest offsets, there is no incentive for public or private financing. Second, although some space satellite systems, deployed mostly by national space agencies, serve to measure air and water quality, land use, urbanization, and other terrestrial processes, none are optimized to measure forested acreage, tree height, or other parameters from which to estimate carbon sequestration. Third, few countries consider forest inventories to be high priority activities, and national space agencies pursue goals other than forest resource management or climate policy. Some countries with sophisticated satellite and other inventory methods use these data on behalf of timber industries and are unlikely to make these data public. For instance, the legislation that authorized the establishment of the National Aeronautics and Space Administration does not include language about natural resources (although it does include language about the "expansion of human knowledge of phenomena in the atmosphere and space").[24] The international organization to which many countries belong for purposes of organizing coordination among Earth-observing satellites, the Group on Earth Observations, has recognized the need to give priority to forest measures and collaborate to overcome the problem that these measures presently receive little priority.[25]

Forest management jurisdiction is another complication. Forests within a country are nationally sovereign resources but their carbon sequestration capacity is a global public good. Nations may undersupply information to global authorities about the extent and health of forests because forest resources, much like deposits of oil, copper, and other resources, are seen as nationally sovereign resources. A nation may fear that sharing data about the quantity, quality, and geographic extent of these resources may reveal information that is commercially important or information that reveals how

24. National Aeronautics and Space Act of 1958, Pub. L. No. 85-568, § 102(c)(1), 72 Stat. 426, 427 (codified at 42 U.S.C. § 2451(c)(1) (1958)), *repealed by* Pub. L. No. 111-314, § 6, 124 Stat. 3328, 3444 (2010).

25. *See Professor José Achache, Director, Group on Earth Observations*, RESEARCHMEDIALTD (Aug. 27, 2011, 1:50 PM), http://www.research-europe.com/index.php/2011/08/professor-jose-achache-director-group-on-earth-observations/.

poorly the nation is managing its resources, such as data pertaining to forest degradation or illegal logging.

A United Nations resolution allows satellite observations of countries (unlike airborne imaging, which requires overflight permission).[26] Thus, collecting forest measures from space might be a solution if financing were available to pay for the systems. The question then becomes one of securing financing. Policies that value forest carbon explicitly would encourage national decision makers to change their priorities in favor of better information about forests. If national decision makers favored forest carbon, forest carbon would finally attain monetary value, which would in turn provide an incentive to finance improved measurement.

IV. OFFSETS AND POLICY OPTIONS: THE CLEAN AIR ACT

Offset policies beyond voluntary actions are only possible if carbon emissions are controlled in some way in the first place. If there is no reason to reduce carbon emissions, there is correspondingly no reason to look for ways to do so cheaply (such as reducing deforestation) or to finance the necessary technological improvements for adequate measurement and monitoring of forests. The most obvious way to control carbon emissions and create an opportunity for offsetting is to price those emissions. In a cap-and-trade system for example, offsets can be used to generate additional allowances. Moreover, if carbon is taxed, offsets can generate tax credits and revenue to underwrite measurement and monitoring.

But a price on carbon is not necessary for offsets—purely regulatory policies are equally compatible, at least in principle. Regulators could require specific emissions reductions, perhaps via a performance standard, but accept offsets in lieu of the required reductions. Offsets are similarly compatible with various hybrid policies. In short, any policy that controls or limits carbon emissions is, in principle, compatible with offsets and sufficient to generate at least some incentive to use them and to fund the required monitoring technology. The degree to which offsets are appealing, however, depends on their relative cost compared to that of the primary emissions reductions required by the policy.[27] Moreover, compatibility in principle does not necessarily mean compatibility in practice. Political and legal limitations

26. G.A. Res. 41/65, ¶ 2, U.N. Doc. A/RES/41/65 (Dec. 3, 1986).

27. If a policy targets only relatively cheap emissions reductions (low-hanging fruit), forest offsets may not be a very attractive alternative. But if caps or regulatory requirements are stringent, requiring deep and costly emissions cuts, offsets will appear very attractive (and regulators' decisions on whether to include them will have a large effect on the program's overall costs).

may prevent offsets from being included in a given policy or limit their scope.

The EPA has the authority to regulate carbon emissions under the Clean Air Act. In *Massachusetts v. EPA*,[28] the Supreme Court held that GHGs are "air pollutants" under the CAA and directed the EPA to investigate whether regulating them was warranted.[29] The EPA issued an "endangerment finding" in late 2009.[30] In it, the agency stated its view that GHG emissions do endanger public health and welfare—a finding that, under the CAA, both enables and compels actual regulation.[31]

The 2009 endangerment finding set the EPA on a path to widespread regulation of U.S. carbon emissions. But is this pathway compatible with offsets? As discussed above, it is compatible in *principle*, but it may not be legally compatible. The EPA may lack the authority to implement offsets via CAA carbon regulation, particularly for international and forest offsets.[32] In addition, even if the agency has the authority to implement offsets, there are a variety of reasons to be skeptical about whether the agency will actually exercise that authority. Limitations on the EPA's authority under the CAA appear to be a significant barrier—though not all analysts agree.[33] Moreover,

28. 549 U.S. 497 (2007).

29. *Id.* at 528. The EPA had argued that it could not regulate GHGs via the CAA since they were not pollutants of the type Congress intended the agency to regulate with its CAA powers. *Id.* The Court rejected this argument, but in doing so it did not compel the EPA to regulate carbon. *Id.* at 534–35. Rather it removed the EPA's justification for refusing to do so, leaving the agency with no choice other than to investigate whether regulating GHGs was necessary based on the statute's requirements. *Id.* Essentially, the holding required the agency, if it continued to refuse to regulate GHGs, to articulate a *scientific* rather than a purely *policy* reason for doing so.

30. Endangerment and Cause or Contribute Findings for Greenhouse Gases Under Section 202(a) of the Clean Air Act; Final Rule, 74 Fed. Reg. 66,496 (EPA Dec. 15, 2009).

31. *Id.*

32. For a more complete review of the EPA's chosen regulatory pathways, see, e.g., NATHAN RICHARDSON, RESOURCES FOR THE FUTURE, ISSUE BRIEF NO. 11-02, EPA GREENHOUSE GAS PERFORMANCE STANDARDS: WHAT THE SETTLEMENT AGREEMENT MEANS (2011); Burtraw et al., *supra* note 3; Richardson, *supra* note 4.

33. *See* Coal. for Emission Reduction Policy, Memorandum, *Comments of the Coalition for Emission Reduction Policy on EPA's Forthcoming Proposal To Establish New Source Performance Standards for GHG Emissions from Electric Generating Units and Refineries* 6 (Mar. 18, 2011), *available at* http://www.uscerp.org/assets/attachments/CERP%20Mar%2018%202011%20Comments%20on%20NSPS%20Rulemakings.pdf (claiming that offsets are "adequately demonstrated," found in other CAA programs, and therefore are compatible with CAA § 111 regulation); *see also* INIMAI M. CHETTIAR & JASON SCHWARTZ, N.Y. UNIV. SCH. OF LAW, THE ROAD AHEAD: EPA'S OPTIONS AND OBLIGATIONS FOR REGULATING GREENHOUSE GASES 88 (2009), *available at* http://policyintegrity.org/files/publications/TheRoadAhead.pdf (arguing that legislative history suggests Congressional intent to allow EPA to consider third-party emissions reductions under § 111 regulation).

institutional and political preferences within the EPA may be just as significant as agency general counsel's evaluation of the legal arguments.[34]

A. OFFSETS UNDER THE CAA: THE PAST

Offsets are not unknown in CAA regulation, having been formally included since the 1977 amendments to the statute—albeit in limited fashion. But these existing CAA offsets, known as emissions reduction credits ("ERCs"), are inapplicable to carbon regulation and are a poor model for both legal and practical reasons.

ERCs, as specified in section 173(c) of the CAA, work in the following manner. Under the CAA, restrictions are placed on the construction of new emitting facilities in areas that violate national air quality standards set by the EPA under section 110 of the Act (called "nonattainment areas").[35] In order for a permit to be issued for construction of a new facility in a nonattainment area, the firm seeking the permit must do two things: first, it must install tight emissions controls (lowest achievable emission rate, or "LAER");[36] second, it must offset the residual emissions from the project.[37] This offsetting is achieved via ERCs. Firms that reduce emissions obtain credits, which can be sold to other firms seeking permits for new projects or used by the reducing firm for its own projects. The ERC program is specified in section 173 of the Act itself; it is not based on the EPA's interpretation of general pollution-control powers under the CAA.[38]

The ERC program, although widely used, is narrow in scope. It is only relevant in nonattainment areas. Furthermore, even in nonattainment areas, offsets are not a general emissions-control tool because offsets only become

34. There has been relatively little discussion of offsets under CAA GHG regulation in the literature. Richardson discussed prospects for *international* offsets alone under the CAA and concluded at the time that such offsets probably could not be included in CAA regulation, though the analysis was necessarily somewhat superficial as the EPA had not yet chosen a regulatory pathway for existing stationary sources. *See* Richardson, *supra* note 4. A recent World Resources Institute and Columbia Law School working paper also briefly addressed offsets under the CAA. The paper concluded that "[i]t is unlikely . . . that offsets could be used to meet the minimum reductions required by EPA's guidelines issued under section 111(d)." Franz T. Litz et al., *What's Ahead for Power Plants and Industry? Using the Clean Air Act To Reduce Greenhouse Gas Emissions, Building on Existing Regional Programs* 20 (Columbia Law Sch. Ctr. for Climate Change Law & World Res. Inst., Working Paper, 2011), *available at* http://pdf.wri.org/working_papers/whats_ahead_for_power_plants_and_industry.pdf. This conclusion is based largely on the simple fact that § 111 includes no explicit mention of offsets, though analysis of this issue was not the focus of the working paper.

35. 42 U.S.C. § 7503 (2006).

36. § 7503(a)(2).

37. § 7503(a)(1).

38. § 7503(c).

relevant when preconstruction permits are needed. Even then, the ERC offsets required for the permit must generally be created within the same nonattainment area.[39] Because of these restrictions, ERC offsets are best viewed as a safety valve that prevents strict regulations on nonattainment areas from completely shutting down economic growth, rather than a general tool for reducing compliance costs.

Despite their long history within CAA regulation, ERCs have little relevance for offsets under CAA programs not related to the section 110 air quality standards, such as those for GHGs. Superficially, the existence of ERCs indicates that offsets are not incompatible with the CAA, at least in principle. By including ERCs in the CAA, Congress demonstrated an awareness of the benefits of offsets and similarly demonstrated it could craft language that specifically includes offsets within the EPA's authority. Congress's failure to include offsets elsewhere in the CAA could thus be interpreted to indicate that Congress did not intend to grant such authority anywhere else. This *expressio unius* argument should not be taken too far. The CAA is a flexible statute, with many different programs aimed at different pollutants from different sources.[40] The EPA has a long history of interpreting these programs relatively independently, and an *expressio unius* argument that depends on Congressional consistency throughout the statute is thus relatively weak. It is difficult to argue that the scope of authority delegated to the EPA should be exactly the same for each of the CAA programs, despite their wide variation in aims and structure.

Even if the *expressio unius* argument fails, however, ERCs are not a useful model for GHG offsets under the CAA for two reasons. First, it is unlikely that national air quality standards will be set for GHGs.[41] Second, the limited

39. *See id.* There is one exception to this rule: offsets can come from another nonattainment area if the other nonattainment area has an equal or higher nonattainment classification than the area in which the source is located and emissions from the other area affect compliance in the area where the permit is being sought.

40. For example, CAA § 112 targets hazardous pollutants from a wide range of sources with strict emissions limits; CAA Title II targets mobile sources with fleetwide and fuel standards; and CAA Title IV implements a national cap-and-trade program for sulfur dioxide emissions. *Id.* §§ 7412, 7521–7554, 7651(b). Each uses different tools to address different pollution problems from different classes of sources. *See id.*

41. *See* Robin Bravender, *EPA Chief Signals Opposition to CAA Curbs on GHGs*, GREENWIRE (Dec. 8, 2009), http://eenews.net/Greenwire/2009/12/08/archive/4?terms=naaqs+petition (quoting Administrator Lisa Jackson saying, "I have never believed and this agency has never believed that setting a national ambient air quality standard for greenhouse gases was advisable"). But it is possible that courts might *force* EPA to set a GHG NAAQS. *See, e.g.*, Ctr. for Bio. Diversity & 350.org, *Petition To Establish National Pollution Limits for Greenhouse Gases Pursuant to the Clean Air Act* 15 (Dec. 2, 2009), *available at* http://www.biologicaldiversity.org/programs/climate_law_institute/global_warming_litigation/clean_air

geographic scope of ERCs precludes the use of the best and cheapest class of offsets—those from forests abroad. The EPA will have to look elsewhere to find a legal basis for including offsets in its GHG regulation.

B. OFFSETS UNDER THE CAA: THE FUTURE

If ERCs are not a useful model for carbon offsets under the CAA, are there other avenues available to the EPA? Perhaps, but it appears unlikely. To understand why, it is first necessary to examine the EPA's plans for carbon—and their statutory basis—in more detail.

The CAA grants authority to the EPA under various regulatory programs that apply to different kinds of pollution from different sources, including mobile sources as well as both new and existing stationary sources. As a result, the EPA's regulatory approach to GHGs is fragmented among different programs and proceeds sequentially as agency actions trigger links between them.

Although the EPA could, in principle, attempt to include offsets in any of its three GHG regulatory programs, the analysis that follows focuses on the EPA's performance standards for existing stationary sources for two reasons. First, the mobile source and stationary source permitting programs are relatively mature at this point and do not include offsetting. Existing source regulation is the only opportunity to include them without redesigning an existing program. Second, regulation of existing stationary sources is the more natural venue for offsets because it covers the sources of the majority of U.S. emissions and will probably generate the largest compliance costs (and therefore the largest opportunity for offsetting).

The EPA's choice of regulatory program for existing stationary sources appears to be performance standards. Under section 111 of the CAA, the EPA can set performance standards for new sources and, via the states, also set standards for existing sources. Under these programs, the agency identifies the "best system of emission reduction"[42] for categories of sources (such as fossil fuel steam power plants and petroleum refineries). These sources are then required to meet the level of emissions set by the standard,[43]

_act/pdfs/Petition_GHG_pollution_cap_12-2-2009.pdf (containing a petition filed by environmental groups claiming that the EPA is required by the CAA to issue a GHG NAAQS); Nathan Richardson, *Greenhouse Gas Regulation Under the Clean Air Act: Does* Chevron *Set the EPA Free?*, 29 STAN. ENVTL. L.J. 283, 308–15 (2010) (arguing that *Chevron* doctrines are likely insufficient to insulate EPA against the statutory arguments that it is required to issue a GHG NAAQS).

42. 42 U.S.C. § 7411(a)(1).

43. § 7411(e).

though they are not required to use any specific technology to do so.[44] The standards for new sources under section 111(b) are termed new source performance standards ("NSPS") and the standards for existing sources under section 111(d) are termed existing source performance standards ("ESPS"). As a result of a settlement agreement, the EPA has announced plans to issue proposed NSPS and ESPS in two categories by the end of 2011: fossil fuel power plants[45] and refineries.[46]

It appears very likely that emissions trading[47] in some form can be incorporated into the GHG NSPS/ESPS program, as discussed in detail in some of our earlier work.[48] To summarize, it appears possible for the EPA to define emissions trading as part of section 111 "standards of performance," which must be based on the "best system of emission reduction," even though the statute does not explicitly permit doing so.[49]

Prospects for including offsets under section 111 performance standards are not as favorable, however. International offsets in particular appear to be legally problematic. As one of us pointed out in a recent paper, there is no precedent for international offsets under the CAA, and it is difficult to interpret section 111 so as to allow their inclusion in performance standards.[50] Nevertheless, it may not be impossible. If, under section 111 performance standards, the "best system of emission reduction" can be interpreted so as to include emissions trading, it might be possible to

---

44. § 7411(b)(5).

45. *See* Settlement Agreement, New York v. EPA (*Boiler GHG*), No. 06-1322 (D.C. Cir. Dec. 21, 2010), *available at* http://www.epa.gov/airquality/pdfs/boilerghgsettlement.pdf. Note the specific source category covered by the settlement is category Da, which includes only fossil fuel powered steam boiler EGUs. Gas-fired turbines and a few other types of fossil power plants are not included in the source category specified in this agreement. *See also* RICHARDSON, *supra* note 32.

46. *See* Settlement Agreement, Am. Petroleum Inst. v. EPA (*Refinery GHG*), No. 08-1277 (D.C. Cir. Dec. 21, 2010), *available at* http://www.epa.gov/airquality/pdfs/refineryghg settlement.pdf.

47. The line between what is considered emissions trading and what is considered use of offsets is not always clear. For purposes of this Article, we consider exchange of emissions credits in some form between different regulated sources to be trading; offsets, in contrast, are the exchange of emissions or carbon credits between a regulated source and an unregulated source, like a forest landowner.

48. *See, e.g.*, Burtraw et al., *supra* note 3, at 297–99; Richardson et al., *supra* note 3, at 10105–06, 10108–11.

49. *See* Richardson et al., *supra* note 3, at 10105–06. The EPA articulated this argument in its 2005 Clean Air Mercury Rule. *See* Standards of Performance for New and Existing Stationary Sources: Electric Utility Steam Generating Units; Final Rule, 70 Fed. Reg. 28,606, 28,616 (May 18, 2005) (stating that "[i]n the final rule, EPA interprets the term 'standard of performance,' as applied to existing sources, to include a cap-and-trade program").

50. *See* Richardson, *supra* note 4.

interpret section 111 to include offsets as well.[51] But pressing the statutory language to support both trading and offsets is a heavy burden for it to bear. Also, offsets—unlike emissions trading—allow emissions reductions to come from outside the regulated sector. This is in tension with the section 111 sectoral regulatory approach, in which standards are set for EPA-defined "source categories." Although the EPA could in principle draw broad source categories and allow trading or offsetting within them, it could never include international sources, since emissions sources outside the United States are almost certainly outside the reach of section 111.[52] Finally, many offsets, including most types of forest offsets, do not result in emissions reductions, but rather in putative reductions in atmospheric carbon concentrations. Addition or preservation of carbon sinks like forests increases the rate at which carbon is pulled from the atmosphere but does not change the amount of emissions generated from any source, source category, the United States, or human activities in total. In this sense, offsets are not a "system of emission reduction" at all and therefore may be fundamentally incompatible with performance standards as defined in the CAA.

Prospects for purely domestic offset programs are perhaps not as grim as those for international offsets, because not all of the above arguments apply in the domestic context. However, the most fundamental arguments against offsets still apply. In any case, domestic offsets alone would likely have a much smaller impact on the cost of emissions reductions, since the most cost-effective sources of offsets are believed to be in developing countries with large forested areas.[53]

The EPA may also lack the powers and institutional capacity to negotiate, implement, and enforce the agreements necessary to support an international offset regime. This is particularly true if agency resources are threatened by congressional budget-cutting. Moreover, to the extent that bold legal arguments would be necessary to include offsets in CAA regulation, the

51. Or it might not: many types of offsets, most notably forestry offsets, do not reduce GHG emissions, but rather GHG *concentrations*, via sequestration of atmospheric carbon. They therefore are not (arguably) a "system of emissions reduction" at all.

52. While the definition of "stationary source" at 42 U.S.C. § 7411(a)(3) (2006) does not explicitly exclude sources outside the United States (it makes no mention of sources' location), nothing in § 111 appears to counter the presumption against extraterritoriality.

53. *See, e.g.*, ADRIAN DEVENY ET AL., RESOURCES FOR THE FUTURE, FOREST CARBON INDEX: THE GEOGRAPHY OF FORESTS IN CLIMATE SOLUTIONS 40–42 (2009), http://www.forestcarbonindex.org/RFF-Rpt-FCI_small.pdf (measuring likely availability of forest carbon offsets and concluding that 18 of the top 20 sources are developing countries).

agency's troubles in the Court of Appeals for the D.C. Circuit in recent years[54] may temper its enthusiasm for legal risk.

C. THE STATES: OFFSETS' LAST, BEST HOPE?

If the EPA is unable to include offsets in federal-level GHG regulation, might states be able to include offsets instead? Yes, but not without complications. In recent years, states have taken the lead on U.S. climate policy, filling the gap left by federal inaction. A group of Northeastern states under the Regional Greenhouse Gas Initiative ("RGGI") have led the way with an electricity-sector GHG trading program.[55] California is also nearing implementation of its own emissions control program under AB32, which is planned to include an expansive emissions trading system.[56] Though these programs are necessarily smaller in scope than a federal program, they are also likely to be more stringent. State programs are therefore capable of generating small but substantial markets for offsets as well as incentives to assess and monitor offset availability and quality—but only if barriers to inclusion of offsets can be overcome.

Since state legislatures (and, via powers delegated from those legislatures, state environmental regulators) control the design of state-level climate polices, those policies can, in principle, include almost any particular tool. For example, a state could enact a cap-and-trade system (as California and the RGGI have), a renewable portfolio standard (as many states have done), a carbon tax, or other mechanisms, including offsets. Nothing prevents California, for example, from allowing land-use changes of some type in the state to generate offsets for use in the AB32 cap-and-trade program. Assuming that interstate trading programs like the RGGI are generally legal, offsets from any of the participating states could similarly be included. For that matter, a program could include out-of-state offsets even if those states are not part of the program.

But legal problems with offsets may arise in two scenarios: first, international offsets may present constitutional problems that could limit or prevent their adoption; second, the limitations on federal-level offsets

54. *See* North Carolina v. EPA, 531 F.3d 896, 901–29 (D.C. Cir. 2008) (striking down the EPA's Clean Air Interstate Rule); *see also* New Jersey v. EPA, 517 F.3d 574, 583 (D.C. Cir. 2008) (striking down the EPA's Clean Air Mercury Rule, on grounds unrelated to the agency's interpretation of CAA § 111 allowing a trading program).

55. *See Memorandum of Understanding*, REGIONAL GREENHOUSE GAS INITIATIVE (Dec. 20, 2005), *available at* http://rggi.org/docs/mou_final_12_20_05.pdf.

56. *See* CAL. AIR RES. BD., CLIMATE CHANGE SCOPING PLAN (2009), *available at* http://www.arb.ca.gov/cc/scopingplan/document/adopted_scoping_plan.pdf.

discussed above could cause compatibility issues with state-level programs, perhaps even causing states to forgo them.

1. States and International Offsets

The power of the states to set their own climate policies is, as noted above, limited only by the Constitution. Although international offsets appear to be the source of the lowest-cost emissions reductions, and have the potential to significantly reduce the costs associated with state offset programs, state-level international offsets are legally risky at best. A variety of constitutional objections to state-level international offsets can be raised, but the common thread among the various arguments is that the power to regulate and conduct foreign affairs is traditionally reserved for the federal government. These arguments have been discussed by some scholars, though they have not by any means been resolved. As Douglas Kysar and Bernadette Meyler put it in a 2008 paper:

> [A]nalyses must necessarily depend on assuming debatable positions within notoriously underdetermined areas of constitutional law, including various restrictions on state foreign affairs activities that emanate from the Treaty Clause, the Compact Clause, the Foreign Commerce Clause, and the foreign affairs preemption doctrine. Although unsatisfying, the safest conclusion to draw in this context is that the recent foreign affairs activities of state and local governments exist in a constitutional fog, similar in many respects to the dim doctrinal haze that covers the interbranch distribution of foreign affairs authority at the federal level.[57]

Full analysis of these constitutional issues is beyond the scope of this paper. In any case the implications for state-level international offsets remain ambiguous. These constitutional issues are at least a legal risk the architects of state programs must consider. Analysts studying the issue disagree over how significant that risk is.[58]

Although academic discussions have focused on linkage of state trading programs with parallel foreign programs (such as the EU's Emissions Trading System), the legal issues for state-level offsets are similar. Incorporating international offsets into a domestic trading program requires some form of agreement with the public or private provider of the emissions reduction credits. It also requires ongoing verification, oversight, and enforcement if there is to be any assurance of offset quality. These ongoing

57. *See* Douglas Kysar & Bernadette Meyler, *Symposium: Like a Nation State,* 55 UCLA L. REV. 1621, 1625 (2008).

58. *See id.* at 1624 n.11.

commitments compel the regulating state to enter into an ongoing and likely formalized relationship with a foreign government, raising similar concerns to those relationships created by the agreements necessary to link trading markets. Although it might in principle be possible to conduct some of the necessary negotiation and agreement exclusively with private parties abroad, it seems likely that some form of negotiation with the political powers responsible for enforcement and capable of giving consent to monitoring efforts would be necessary. And even if states maintain relationships purely with foreign private actors, some of the constitutional objections may remain legally significant.

On the other hand, the relevant constitutional doctrines are ambiguous. It is impossible to say that any one of them conclusively or even probably limits the states' ability to include international offsets in its programs.

2. Federal-State Compatibility

Constitutional concerns regarding *domestic* offsets are much less significant, though again the value of purely domestic offsets in both cost and environmental terms is limited. Whether the focus is on domestic offsets or the constitutional issues are simply set aside for the time being, interaction between EPA regulation under the CAA and any state programs may create practical obstacles for state-level offset programs. That is, even if states face no legal restrictions on their ability to incorporate offsets (domestic, international, or both), states could ironically be discouraged from doing so by the presence of the parallel EPA program.

If the EPA cannot include offsets in its federal program, or simply chooses not to, the EPA's program may become partly or wholly incompatible with state programs that do include offsets. If emitters in, say, California, comply with emissions cuts required by the state solely via offset purchases without reductions in emissions from the regulated source category (sector) itself, the emitters would be out of compliance with the federal standard. This is true even if state requirements are stricter than federal requirements in emissions terms. In this scenario, offsets would only be useful for *additional* emissions reductions required by states, increasing the cost of those programs without any emissions benefit.[59] It is possible that the

59. This simple scenario hides much legal complexity. *See* 42 U.S.C. § 7411(d) (2006). ESPS regulation is primarily a state activity; the EPA simply sets initial guidelines and reviews state plans, intervening only if states fail to act. But states do not have complete discretion in writing their § 7411(d) plans. As discussed above, states must set standards of performance within the definition of the CAA, which appears to limit their ability to incorporate offsets. Section 7411(d) does not restrict states' ability to regulate emissions more stringently, but this does not grant states the ability to use tools other than "standards

EPA could take more creative approaches to section 111(d) regulation, such as setting state-level budgets rather than facility-level targets. But it remains unclear whether such approaches would permit any additional flexibility regarding offsets (or trading that includes uncovered sources).

If states with their own climate programs are unable to use offsets for compliance with section 111(d), these states' program choices will be limited and the likelihood of other states joining existing interstate climate agreements may decrease. States with existing programs will be much less likely to include offsets in their programs since they would only be useful for emissions restrictions beyond federal requirements. The administrative, enforcement, and compliance costs of an offset program might not be justified under these conditions. As a result, the cost of cutting emissions would increase (assuming offsets are the cheapest option available to emitters for meeting state requirements). This has obvious effects on the regulating states, but it also makes other states less likely to join interstate programs. Increasing costs for state programs may be only the first part of a double blow—federal climate regulation could also undermine these programs' political momentum.

V. SUMMARY AND FUTURE STEPS

Even with the physical and economic significance of forest carbon offsets in climate policy, and despite the technological capacity to measure and track forests around the world, federal climate regulation under the CAA appears unlikely to allow inclusion of forest offsets. Although it is not possible to rule out GHG offsets under the CAA on legal grounds, the arguments against legality discussed above make it much less likely that the EPA will take the risk of including them. Although the agency has made some bold interpretive moves in the recent past, most notably in its Clean Air Mercury Rule, courts have not generally been receptive to these ambitious statutory interpretations.[60] The agency may therefore have lost some of its appetite for ambitious interpretation of the CAA, particularly in the context of the EPA's already-controversial GHG regulatory programs.

Ultimately, more research is needed to determine the legality and feasibility of carbon offsets in state and federal climate policy. Some legal issues surrounding the plausibility of *domestic* offsets remain unclear; while the opportunities for cost reduction from these offsets are limited, they are not

of performance." This is in contrast to state plans under the CAA NAAQS program, which does grant such broad flexibility so long as environmental targets are met.

60. *See, e.g.*, cases cited *supra* note 54.

necessarily trivial. And if such offsets can be included in EPA regulation, there would be incentives to invest in technology that might later be expanded internationally. Many aspects of the relationship between EPA/CAA section 111 performance standards and state emissions programs also remain unclear. Although this Article has attempted to explore these issues to some degree, more study is needed. The coming months will likely reveal a great deal about the EPA's intentions and states' preferences with regards to the use of carbon offsets, though the final boundaries of legality may not be known until likely ensuing litigation is resolved.

Moreover, implementing offsets will likely be difficult in light of the probable reductions in EPA funding in the coming years. Offset programs are likely to be administratively complex and labor-intensive for the agency, especially relative to more traditional approaches to performance standards under section 111. Although some of this workload could in principle be shifted to states under section 111(d), much could not—especially insofar as international offsets are concerned. And the budgetary situations in the states are hardly more favorable.

With domestic action at the federal and state levels uncertain, opportunities for exploiting forest carbon sequestration are likely to continue to play a role, albeit a limited one, in the United States' international diplomatic actions. For example, the Fiscal Year 2011 U.S. Budget includes $347 million, to be administered by the U.S. Department of the Treasury and the Agency for International Development, for enhancing forest sequestration management in developing countries.[61] This program will draw on the technical capacity to measure and monitor but will fall far short of realizing the economic benefits of including forest management in domestic climate policy. Meanwhile other nations appear to be moving ahead with forest carbon offsetting as a component of domestic policy. For example, Norway is actively engaged in the mapping, monitoring, and financing of

61. *See* Dep't of State, USAID & Dep't of the Treasury, *FY 2011 Budget for International Climate Change Financing*, *available at* http://www.usclimatenetwork.org/resource-database/resource-database/fy-2011-summary-of-core-climate-assistance-budget (last visited Nov. 5, 2011). This action is a step towards fulfilling U.S. promises made at the 15th Conference of the Parties (CoP) to the United Nations Framework Convention on Climate Change (UNFCCC) in December 2009 in Copenhagen. Robert N. Stavins and Robert C. Stowe assess in detail the Copenhagen meetings and these provisions. *See* Robert N. Stavins & Robert C. Stowe, *What Hath Copenhagen Wrought? A Preliminary Assessment*, ENV'T MAG., May/June 2010, *available at* http://www.environmentmagazine.org/Archives/Back%20Issues/May-June%202010/what-wrath-full.html.

forest carbon storage in developing countries and then counting this effort toward Norway's domestic emissions reduction target.[62]

If the U.S. climate policy pathway leaves offset opportunities unexploited, as legal and institutional barriers indicate it will, at least over the short term incentives to innovate and invest in the technology necessary to support global forest offsets will be substantially blunted. Without this technological investment, *future* offset programs will be more difficult and more costly to implement. Countries that do pursue international offsets as part of their emissions reduction policy will have to bear a much larger share of the technological burden. This will likely reduce the quality of offsets that are available (increasing their cost or reducing their real emissions impact), dissuade countries from including offsets in their policies at all, or both. Assuming land use and offset information would be shared (and there is little reason it would not be), monitoring and verification technology is a public good. Although the lack of U.S. participation does not necessarily doom development of this technology, it is a big blow.

Moreover, the likely failure to include offsets and promote forest monitoring technology is an unnecessary artifact of the particular path the United States has chosen for climate policy. If the United States adopted no emissions reduction policy at all, then of course there would be no incentive for U.S. investment in offset technology. But this would be the least of the environmental problems flowing from the choice to do nothing. In fact, the United States does and will continue to have an emissions reduction policy, even over the short term, driven by the EPA and the states. The difficulty of integrating offsets into this regime is among its largest failures. This limitation of Clean Air Act climate policy will increase costs, decrease achievable emissions benefits, and result in a missed opportunity for technological and environmental investments with large present and future benefits.

62. *See* NORWEGIAN MINISTRY OF THE ENV'T & MINISTRY OF FOREIGN AFFAIRS, THE GOVERNMENT OF NORWAY'S INTERNATIONAL CLIMATE AND FOREST INITIATIVE (2010), *available at* http://www.regjeringen.no/upload/MD/Vedlegg/Klima/klima_skogprosjektet/mai2010.pdf.

REGULATING PRIVACY BY DESIGN

Ira S. Rubinstein[†]

TABLE OF CONTENTS

---

† Adjunct Professor of Law and Senior Fellow, Information Law Institute, New York University School of Law. This Article was presented at the NYU Privacy Research Group, the Princeton's Center for Information Technology Policy, and the Privacy Law Scholars Conference and I am grateful for the comments of workshop participants. For detailed comments on an early draft, I am indebted to Kelly Caine, Peter Cullen, Erin Egan, Jacques Lawarrée, Ron Lee, Paul Schwartz, and Tal Zarsky. Thanks are also due to Solon Borcas, Travis Breaux, Anapum Datta, Cathy Dwyer, Kenneth Farrall, Foster Provost, and Adam Shostack for insights on various technology-related issues, and to Jeramie Scott for able research assistance. A grant from The Privacy Projects supported this work.

I. INTRODUCTION

Privacy officials in Europe and the United States are embracing privacy by design as never before. This is the idea that in designing information and communications technologies ("ICT"), building in privacy from the outset achieves better results than bolting it on at the end.[1] The European Union Data Protection Directive has always included provisions requiring data controllers to implement "technical and organizational measures" in the design and operation of ICT.[2] But this has proven insufficient and in their new call for privacy by design, the European Commission ("EC") hopes to see data protection principles taken into account at the outset of designing, producing, or acquiring ICT systems. In particular, they are encouraging both the use of Privacy Enhancing Technologies, or PETs, as well as default settings that favor privacy.[3]

1. *See* Ann Cavoukian, *Privacy by Design*, INFO. & PRIVACY COMM'R, 1 (2009), http://www.ipc.on.ca/images/Resources/privacybydesign.pdf (stating that she "first developed the term 'Privacy by Design' back in the '90s" and that " 'Build in privacy from the outset' has been [her] longstanding mantra, to 'avoid making costly mistakes later on, requiring expensive retrofits' "); *see also Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A Digital Agenda for Europe*, COM (2010) 245 final/2 (Aug. 26, 2010), *available at* http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0245:FIN:EN:PDF [hereinafter EC, *A Digital Agenda for Europe*].

2. Council Directive 95/46 requires data controllers to "implement appropriate technical and organizational measures" for safeguarding personal data. In addition, Recital 46 calls for such measures to be taken, "both at the time of the design of the processing system and at the time of the processing itself." Directive 95/46/EC, 1995 O.J. (L 281) 31 (Nov. 23, 1995), *available at* http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML [hereinafter EU Data Protection Directive].

3. *See* ART. 29 DATA PROTECTION WORKING PARTY, 02356/09/EN, WP 168, THE FUTURE OF PRIVACY (2009), *available at* http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2009/wp168_en.pdf; John J. Borking & Charles D. Raab, *Laws, PETs and Other Technologies for Privacy Protection*, 2001 J. INFO L. & TECH., no. 1, http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2001_1/borking/ (noting that PETs have a very specific meaning, namely, "a coherent system of ICT measures that protects privacy by eliminating or reducing personal data or by preventing unnecessary and/or undesired processing of personal data, all without losing the functionality of the information system"). This same definition is cited in *Commission Communication on Promoting Data Protection by Privacy Enhancing Technologies (PETs)*, COM (2007) 228 final (May 2, 2007). In 1995, in one of the earliest discussions of PETs, privacy commissioners from Ontario, Canada and the Netherlands collaborated on a paper describing the privacy concerns associated with the trail of identifying information created by electronic transactions and a number of techniques that would permit users to engage in

In the United States, a recent staff report of the Federal Trade Commission ("FTC") describes a Proposed Framework with three main components: privacy by design, simplified consumer choice, and increased transparency of data practices.[4] According to the Staff Report, companies engage in privacy by design when they promote consumer privacy throughout their organizations and at every stage of the development of their products and services.[5] More specifically, privacy by design has two main elements: first, incorporating substantive privacy protections into a firm's practices; and second, maintaining comprehensive data management procedures throughout the life cycle of their products and services.[6] The report also briefly mentions the use of PETs such as identity management, data tagging tools, transport encryption, and tools to "check and adjust default settings."[7] In short, regulators on both sides of the Atlantic agree on the need for a new legal framework to protect online privacy in the twenty-first century and that one of its major aspects should be privacy by design.

Although PETs and privacy by design resist precise definition and even overlap as to their usage, the two ideas are not identical. Their differences may be summed up as follows: PETs are applications or tools with discrete goals that address a single dimension of privacy, such as anonymity, confidentiality, or control over personal information. Frequently, PETs are added onto existing systems, sometimes as an afterthought by designers and sometimes by privacy-sensitive end-users.[8] In contrast, privacy by design is not a specific technology or product but a systematic approach to designing

transactions without revealing their identity. *See* 2 INFO. AND PRIVACY COMM'R (Ontario, Canada) & REGISTRATIEKAMER (Netherlands), PRIVACY-ENHANCING TECHNOLOGIES: THE PATH TO ANONYMITY (1995), *available at* http://www.ontla.on.ca/library/repository/mon/10000/184530.pdf [hereinafter 1995 PETS REPORT] (proposing the use of "identity protectors" that "separate one's true identity from the details of one's transactions through the use of 'pseudo-identities' "). Although this 1995 report treats PETs primarily in these terms, the 2007 communication from the EC reflects a much broader view of PETs as encompassing not only identity protection but various encryption tools, cookie managers and other filtering devices, as well as data management protocols such as the Platform for Privacy Preferences ("P3P").

4. BUREAU OF CONSUMER PROTECTION, FED. TRADE COMM'N (FTC), PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE (2010), *available at* http://www.ftc.gov/os/2010/12/101201privacyreport.pdf [hereinafter FTC STAFF REPORT].

5. *Id.* at 41.

6. *Id.* at 44–52.

7. *Id.* at 52 n.131.

8. Indeed, many PETs now take the form of so-called "browser add-ons." *See* Jim Brock, *Are Privacy Add-Ons Effective? Surprising Results from Our Testing*, PRIVACYCHOICE (Nov. 17, 2010), http://blog.privacychoice.org/2010/11/17/are-privacy-add-ons-effective-surprising-results-from-our-testing/ (comparing the effectiveness of a single type of privacy add-on that blocks efforts by data and marketing companies to track online activity).

any technology that embeds privacy into the underlying specifications or architecture.[9] Although PETs and privacy by design are distinguishable, these two phrases have no established usage and regulators and commentators often use them interchangeably.

Despite the endorsement of regulators, PETs have not achieved widespread acceptance in the marketplace, and relatively few firms have embraced privacy by design.[10] Confusion over a variety of key definitions contributes to this slow pace of adoption. For instance, it is not yet clear how the concept of "privacy by design" relates to certain technologies or organizational measures, nor what regulators really have in mind when they urge firms developing products to build in privacy.

Economics also plays an important role in determining the adoption rate of PETs and privacy design practices. On the consumer side, few PETs have proven popular and the demand for products and services with strong privacy safeguards seems quite limited. Reasons include consumers' lack of knowledge concerning the privacy risks associated with web surfing, search, social networks, e-commerce, and other daily internet activities and their limited understanding of how PETs or privacy by design might help reduce these risks. Further, cognitive and behavioral biases may prevent some individuals from acting in accordance with their stated preference for greater privacy. Other consumers just do not care very much about privacy.[11] On the business side, weak consumer demand discourages information technology ("IT") spending. Moreover, given the huge profits many firms derive from online advertising, they are reluctant to voluntarily implement PETs or design practices that would limit their ability to collect, analyze, or share valuable consumer data.[12]

Although the European Commission sponsored a study of the economic costs and benefits of PETs, and the United Kingdom is looking at how to improve the business case for investing in privacy by design, there is scant evidence that privacy technology pays for itself, much less confers a competitive advantage on firms that adopt it. Indeed, the economic or regulatory incentives for adopting privacy by design need more attention in Europe and are largely absent from the FTC report. In the meantime, the regulatory implications of privacy by design are murky at best, not only for firms that might adopt this approach but for free riders as well.

---

9. *See* Cavoukian, *supra* note 1, at 1 (noting that "[embedding privacy] may be achieved by building the principles of Fair Information Practices (FIPs) into the design, operation and management of information processing technologies and systems").

10. *See infra* Sections III.A, III.C.

11. *See infra* Section III.A.

12. *See infra* Section III.C.

This Article seeks to clarify the meaning of privacy by design and to suggest how privacy officials might develop appropriate regulatory incentives that offset the certain economic costs and somewhat uncertain privacy benefits of this new approach. Part II begins by developing a taxonomy of PETs, classifying them as substitutes or complements depending on how they interact with data protection or privacy laws. Substitute PETs aim for zero-collection of personal data[13] and, if successful, make legal protections less important or even superfluous. In contrast, complementary PETs fall into two subcategories: those which are privacy-friendly and those which are privacy-preserving. These are familiar terms within the privacy literature but they have no fixed meaning. As used here, "privacy-friendly" means literally a system or even a feature that welcomes individual control over personal data, mainly through enhanced notice, choice, and access, whereas "privacy-preserving" refers to a much smaller number of systems offering provable guarantees of privacy, mainly through cryptographic protocols or other sophisticated measures.

Part III explores the meaning of privacy by design in the specific context of the FTC's emerging concept of comprehensive information privacy programs ("CIPPs"). It also looks at how privacy by design practices relate to the use of PETs and at the activities of a few industry leaders, who rely on engineering approaches and related tools to implement privacy principles throughout the product development and the data management lifecycles. Building on this analysis, and using targeted advertising as a primary illustration against the backdrop of the FTC analysis, the Article then suggests that economic incentives are inadequate to ensure widespread adoption of PETs or significant investments in the design aspects of CIPPs.

Finally, Part IV considers how regulators might achieve better success in promoting the use of privacy by design by (1) identifying best practices, including prohibited practices, required practices, and recommended practices, which are compiled in the Appendix; and (2) situating these best practices within an innovative regulatory framework that (a) promotes experimentation with new technologies and engineering practices; (b) encourages regulatory agreements through stakeholder representation, face-to-face negotiations, and consensus-based decision making; and (c) supports flexible, incentive-driven safe harbor mechanisms as defined by newly proposed privacy legislation.

---

13. The EU Data Protection Directive defines "personal data" as "any information relating to an identified or identifiable natural person." EU Data Protection Directive, *supra* note 2, art. 2(a). In the United States, the cognate concept is personally identifiable information ("PII").

II. PETS AND PRIVACY BY DESIGN

The FTC's Proposed Framework states that companies should develop and implement CIPPs to ensure proper incorporation of the four substantive principles identified in the report (data security, reasonable collection limitations, sound retention practices, and data accuracy).[14] The two core elements of CIPPs are (1) assigning specific personnel the responsibility of privacy training, and (2) promoting accountability for privacy policies and assessing and mitigating privacy risks. These privacy assessments should occur before a product launches and periodically thereafter to address any changes in data risks or other circumstances. The size and scope of a CIPP should be determined based on the data at stake and the risks of processing such data, with companies that collect vast amounts of consumer data or sensitive data required to devote more resources than those collecting small amounts of non-sensitive data. Finally, the report mentions in passing that the FTC staff supports the use of PETs.[15]

The Staff Report's enticing description of privacy by design has great intuitive appeal. Why is this? The FTC's discussion suggests that privacy by design generally reduces errors and costs,[16] yet this discussion remains short on specifics and never quite explains what privacy by design amounts to. Do companies engage in privacy by design by making more and better use of PETs and, if so, what sorts of PETs are most effective and why? The report recommends, without discussion, the use of several kinds of PETs (identity management, data tagging tools, transport encryption, and tools to check and adjust default settings),[17] but it makes no effort to differentiate them according to relevant criteria. Alternatively, does privacy by design mean that companies should implement specific design practices or compliance measures? Without more detailed guidance, firms will not know what they are supposed to do (or not do), how much they should spend to achieve the desired outcomes, or to what extent this approach will enhance their standing

14. *See* FTC STAFF REPORT, *supra* note 4, at 50.

15. *See id.* at 44–52. For a more detailed FTC statement describing CIPPs in terms of five major elements, see *infra* notes 66–71 and accompanying text.

16. There is evidence that resolving security issues during the design phase is more efficient and less costly than having to deal with it later in the development process. *See* MARK GRAFF & KENNETH VAN WYK, SECURE CODING: PRINCIPLES AND PRACTICES 56 (2003) (citing evidence that the cost of a bug fix at design time is considerably less than the cost of fixing the same bug during implementation or testing, a disparity that only increases if a patch is required). It is beyond the scope of this Article to determine whether privacy design flaws are analogous to security bugs or if it is also cheaper to fix the former at an early as opposed to a later stage.

17. *See* FTC STAFF REPORT, *supra* note 4, at 52.

with regulators. The following discussion lays the groundwork for examining these issues by developing a new taxonomy of PETs, exploring the meaning of privacy by design, and comparing existing private sector approaches to the FTC's analysis in the Staff Report.

A. THE SUCCESSES AND FAILURES OF PETS

PETs have been around for about twenty-five years. Many PETs reflect major advances in cryptographic research, which have also enabled advanced privacy features such as anonymous payment systems, anonymous protection for real-time communications, authentication via anonymous credential schemes, and methods for anonymously retrieving online content.[18] Identity protectors and related PETs were first introduced as a regulatory strategy in the 1995 report on the "path to anonymity."[19] However, as Feigenbaum and her colleagues summed it up a little more than fifteen years later: "Despite the apparent profusion of such technologies, few are in widespread use. Furthermore, even if they were in widespread use, they would not necessarily eliminate" various deployment problems.[20]

Of course, not all PETs rely on anonymity protocols. The term encompasses a range of tools beyond anonymity including those that enhance notice and choice, help automate communication and/or enforcement of privacy policies, or ensure confidentiality via encryption. Arguably, anonymity tools are the most effective PETs precisely because they prevent identification or collection of personal data in the first place, irrespective of legal requirements. As a result, they are sometimes referred to as true or pure PETs.[21] In contrast, other privacy tools permit data collection and analysis but seek to assist knowledgeable and motivated consumers in

18. *See* Joan Feigenbaum et al., *Privacy Engineering for Digital Rights Management Systems*, 2320 LECTURE NOTES COMPUTER SCI., art. 6, 2002, *available at* http://cs-www.cs.yale.edu/homes/jf/FFSS.pdf.

19. *See* 1995 PETs REPORT, *supra* note 3.

20. These include overdependence on abstract models as opposed to "real-world" uses, insecure implementations, ease-of-use issues, and integration of PETs with legacy systems. *See* Feigenbaum et al., *supra* note 18, at 6–10; *see also* Ira Rubinstein et al., *Data Mining and Internet Profiling: Emerging Regulatory and Technological Approaches*, 75 U. CHI. L. REV. 261, 274–77 (2008) (discussing underutilization of anonymity tools due to apathy, consumer ignorance, and difficulty in finding, understanding, and configuring the relevant tools).

21. For an explicitly normative treatment of PETs, see, e.g., Roger Clarke, *Introducing PITs and PETs: Technologies Affecting Privacy*, 7 PRIVACY L. & POL'Y REP., no. 9, Feb. 2001, at 181, *available at* http://www.austlii.edu.au/au/journals/PLPR/2001/12.html (distinguishing PETs from so-called PITs (Privacy-Invasive Technologies), whose primary function is surveillance, and distinguishing "savage" PETs, which set out to deny identity and to provide untraceable anonymity, from "gentle" PETs, which include pseudonymity tools that balance the shielding of identity with accountability).

exercising greater control over what data they share and with whom they share it.

Although the Commission recommends the use of PETs, the Staff Report fails to discuss the different kinds and uses of PETs or their historical successes and failures. There is, in fact, a large literature on PETs including a number of proposed classifications. Most classifications of PETs take a functional approach (i.e., they distinguish PETs based on whether they ensure anonymity, confidentiality, transparency, and so on).[22] However, this is sometimes combined with other factors such as whether end-users deploy the PET on the client side or if firms deploy them on the server side. Other researchers classify PETs based on their underlying conception of privacy (e.g., control, autonomy, seclusion), but this has not proven very useful.[23]

This Article takes a different approach by classifying PETs in terms of how they relate to government regulation. The next Section suggests that *all* PETs fall into one of two very broad categories: substitute PETs (which take the place of privacy regulation by shielding identity and/or preventing the collection of personal data or personally identifiable information ("PII")) or complementary PETs (which support regulatory goals by using technical measures to achieve specific goals). The Article demonstrates that this categorization is far more likely to result in useful guidance to the private sector on their adoption of PETs.[24]

22. *See, e.g.*, COLIN J. BENNETT & CHARLES D. RAAB, THE GOVERNANCE OF PRIVACY: POLICY INSTRUMENTS IN GLOBAL PERSPECTIVE 180–202 (2006); Lorrie Faith Cranor, *The Role of Privacy Enhancing Technologies*, *in* CONSIDERING CONSUMER PRIVACY: A RESOURCE FOR POLICYMAKERS AND PRACTITIONERS 80 (Paula J. Bruening ed., Mar. 2003), *available at* http://old.cdt.org/privacy/ccp/roleoftechnology1.pdf (full volume available at http://old.cdt.org/privacy/ccp/ccp.pdf); Ian Goldberg, *Privacy Enhancing Technologies for the Internet III: Ten Years Later*, *in* DIGITAL PRIVACY: THEORY, TECHNOLOGIES, AND PRACTICES 3 (A. Acquisti et al. eds., 2007); NAT'L RESEARCH COUNCIL, ENGAGING PRIVACY AND INFORMATION TECHNOLOGY IN A DIGITAL AGE 107–16 (2007).

23. *See, e.g.*, Herbert Burkert, *Privacy-Enhancing Technologies: Topology, Critique, Vision*, *in* TECHNOLOGY AND PRIVACY: THE NEW LANDSCAPE (Philip E. Agre & Marc Rotenberg eds., 1998); L.J. Camp & C. Osorio, *Privacy-Enhancing Technologies for Internet Commerce*, *in* TRUST IN THE NETWORK ECONOMY (O. Petrovic et al. eds., 2003); Herman T. Tavani & James H. Moor, *Privacy Protection, Control of Information, and Privacy-Enhancing Technologies*, 31 COMPUTERS & SOC'Y 6 (2001).

24. For a similar distinction, see BENNETT & RAAB, *supra* note 22, at 153, 180 (noting that in Europe, PETs are often regarded as "a useful complement to existing regulatory and self-regulatory approaches" while in the United States they have sometimes been positioned as "an alternative to regulatory intervention").

B. A TAXONOMY OF PETS: SUBSTITUTES VS. COMPLEMENTS

Substitute PETs seek to protect privacy by blocking or minimizing the collection of personal data, thereby making legal protections superfluous. In contrast, complementary PETs permit the collection and use of such data as long as these activities are consistent with privacy laws and related statutory requirements.[25]

The main types of substitute PETs rely on anonymity to shield or reduce user identification and/or on client-centric architectures to prevent or minimize the collection of PII.[26] Their design is motivated by an underlying assumption that commercial IT systems are flawed, while legal rules and sanctions are in most (if not all) cases ineffective. These PETs shift the locus of protection from oversight of firm behavior to prevention or avoidance of the data collection and analysis requiring oversight in the first place. Most of the best known substitute PETs are discrete applications deployed by individual end-users to provide limited functionality (e.g., anonymous browsing or encrypted email).[27] Some substitute PETs also require ongoing maintenance, research, and support from non-profits and volunteers (e.g., the Tor anonymity network), but it is rare to see businesses deploy substitute PETs in their own products or services.

In practice, many substitute PETs are more theoretical than practical. Few are widely deployed,[28] for the reasons discussed above, and the firms that have sought to create a business around such tools have failed, which in turn discourages further investment.[29] This is hardly surprising. Profit-motivated internet firms collect and analyze personal data for multiple purposes—serving targeted ads, personalizing their services, and charging

25. Most U.S. privacy laws focus on the collection and use of PII, while EU privacy law turns on the related concept of personal data. Under both regimes, the collection and use of information other than PII or personal data is unregulated.

26. *See* S. Spiekermann & L. Cranor, *Engineering Privacy*, 35 IEEE TRANSACTIONS ON SOFTWARE ENGINEERING 67 (2009).

27. This is at least partly the result of the inhibitory effect of a regulatory environment driven by concerns over money laundering and other financial crimes, which have undermined government (and hence private-sector) support for anonymous payment systems and other forms of anonymity.

28. For a discussion of the most popular and useful substitute PETs, see Ethan Zuckerman, *How To Blog Anonymously*, *in* HANDBOOK FOR BLOGGERS AND CYBER-DISSIDENTS (Reporters Without Borders ed., 2005), *available at* http://www.rsf.org/IMG/pdf/Bloggers_Handbook2.pdf.

29. *See* Goldberg, *supra* note 22, at 12. Nevertheless, firms persist in trying to distinguish themselves on the basis of privacy. For recent examples of search engines that seek to maximize user privacy, see IXQUICK, http://www.ixquick.com (last visited Nov. 6, 2011); DUCKDUCKGO, http://www.duckduckgo.com (last visited Nov. 6, 2011).

prices that extract as much surplus as possible from any sale (which economists refer to as price discrimination).[30] As a result, they are reluctant to adopt substitute PETs voluntarily, which further erodes any market in such tools.

In sharp contrast, complementary PETs are designed to implement statutory privacy principles or related legal requirements. Thus, businesses are eager to deploy them both to ensure regulatory compliance and/or to give customers a positive impression of their commitment to privacy (here understood in terms of control over personal data). Developers of complementary PETs take it for granted that firms will collect data for various useful (and profitable) purposes. Their goal in developing complementary PETs is not to block or minimize such collection, but to reduce the risk of consumer harms by ensuring that data is collected and processed in compliance with regulatory requirements based on Fair Information Practice Principles ("FIPPs"). Complementary PETs can focus on the front-end user experience (e.g., informed consent mechanisms, access tools, and preference managers), or address privacy issues that arise with back-end infrastructure and data sharing networks (e.g., IBM's Tivoli Privacy Manager, which helps enterprises manage user identities, access rights, and privacy policies across an entire e-business infrastructure, and HP's proposed Policy Compliance Checking System).[31]

Complementary PETs fall into two subcategories: *privacy-friendly* and *privacy-preserving* PETs. Privacy-friendly PETs seek to give people more control over their personal data through improved notice and consent mechanisms, browser management tools, digital dashboards, and so on. In contrast, privacy-preserving PETs in many cases resemble substitute PETs. They rely on sophisticated cryptographic protocols that may lead to

30. *See* LONDON ECON., STUDY ON THE ECONOMIC BENEFITS OF PRIVACY-ENHANCING TECHNOLOGIES (PETS) 46–49 (2010), *available at* http://ec.europa.eu/justice/policies/privacy/docs/studies/final_report_pets_16_07_10_en.pdf [hereinafter LONDON ECON. STUDY] (also noting the use of personal data as a "productive resource" and "tradable commodity"). Some economists argue that price discrimination is the principle motivation for businesses to collect personal data and that privacy erosion is driven to a large extent by the incentives to price discriminate. *See* Andrew Odlyzko, *Privacy, Economics, and Price Discrimination on the Internet*, *in* ICEC2003: PROCEEDINGS OF THE FIFTH INTERNATIONAL CONFERENCE ON ELECTRONIC COMMERCE (N. Sadeh ed., 2003), *available at* http://www.dtc.umn.edu/~odlyzko/doc/privacy.economics.pdf.

31. On the front-end/back-end distinction, see LONDON ECON. STUDY, *supra* note 30, at 13 (noting that Yoram Hacohen, Head of the Information and Technology Authority of Israel, draws a similar distinction between "technologies that are used before any personal data is used ('pre-usage') and technologies that safeguard privacy while personal data is being processed"), and *infra* Section II.C.1.

deployable solutions with strong privacy guarantees but that also satisfy legal requirements. This combination of features permits companies and government agencies to engage in activities that might otherwise be viewed as privacy invasive while preserving privacy in a rigorous manner. Good examples include privacy-preserving data mining[32] and privacy-preserving targeted advertising.[33]

Why are these distinctions important?[34] The answer relates to the incentives for developing and using PETs. Bluntly, the market incentives for substitute PETs are feeble. On the other hand, a much stronger business case exists for complementary PETs because they both support existing compliance obligations and tend to enhance a firm's reputation as a trustworthy company that cares about privacy. Of course, businesses will adopt complementary PETs only if they determine that the direct and opportunity costs of doing so are low enough to justify the investment. Thus, firms are less likely to adopt privacy-preserving PETs because they are both harder to implement and less flexible than privacy-friendly PETs. These observations suggest that regulatory incentives may still be necessary to overcome the reluctance of private firms to increase their investments in PETs, especially in the face of limited consumer demand, competing business needs, and a weak economy.

The distinction between substitute and complementary PETs and the incentives for adopting them are well-illustrated by PETs designed to control the receipt of targeted advertising.[35] This Section concludes with brief

32. *See* Rakesh Agrawal & Ramakrishnan Srikant, *Privacy-Preserving Data Mining*, 29 SIGMOD REC. 439 (2000).

33. *See* VINCENT TOUBIANA ET AL., ADNOSTIC: PRIVACY PRESERVING TARGETED ADVERTISING (2010), *available at* http://crypto.stanford.edu/adnostic/adnostic.pdf.

34. For the sake of completeness, we may also distinguish a third category of PETs consisting in certain hybrid privacy solutions that may exhibit characteristics of privacy by design and utilize one or more kinds of PETs. Examples of such hybrid solutions may be found in Daniel J. Weitzner et al., *Information Accountability*, 51 COMM. ACM 82 (2008) (describing an accountability framework that combines strict legal rules on the permissible uses of data with a technical architecture that supports policy-aware transaction logs, a policy-language framework, and policy-reasoning tools); PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH. (PCAST), REALIZING THE FULL POTENTIAL OF HEALTH INFORMATION TECHNOLOGY TO IMPROVE HEALTHCARE FOR AMERICANS: THE PATH FORWARD 46 (2010), *available at* http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-health-it-report.pdf (recommending a new health IT architecture that offers much stronger privacy and security protections than existing systems by using a universal exchange language and "tagged" data elements, i.e., each unit of data is accompanied by a mandatory "metadata tag" that describes the attributes, provenance, and required privacy and security protections of the data).

35. *See* FTC, SELF-REGULATORY PRINCIPLES FOR ONLINE BEHAVIORAL ADVERTISING 9 n.21 (2009), *available at* http://www.ftc.gov/os/2009/02/P085400behavad

descriptions of the PETs in targeted advertising sorted into the categories of PETs distinguished above:

> **1. Substitute PETs:** Various anonymity tools are available that prevent tracking and targeted advertising by enabling consumers to surf the web anonymously. For example, anonymous proxy servers permit users to surf the web without revealing their IP addresses. The Tor Browsing bundle offers similar functionality using a much stronger cryptographic protocol. Consistent with their business models, however, none of the major search or network advertising firms support the use of such tools in their web services, either by building in such functionality or educating users about where to find and how to use these PETs. It seems unlikely that the FTC could devise attractive enough incentives to overcome the opportunity costs associated with substitute PETs short of threatening highly restrictive regulations for those failing to adopt them.
>
> **2. Complementary Privacy-Friendly PETs:** On the other hand, many of the most popular commercial internet and network advertising firms strongly support tools that enable users to control their online advertising by editing their inferred interest and demographic categories or opting-out of behavioral targeting with respect to participating firms. Examples include ad preference managers, standalone and browser-based cookie managers, additional browser controls that allow users to delete cookies (including Flash cookies), "private browsing" features (which delete cookies each time the user closes the browser or turns off private browsing, effectively hiding his history), new icons that link to additional information and choices about behavioral advertising, and new, browser-based "do not track" tools from all three of the major browser vendors. These PETs are attractive to companies for obvious reasons: they enhance notice and choice in a privacy-friendly manner without disrupting the advertising business model.
>
> **3. Complementary Privacy-Preserving PETs:** Finally, a group of privacy researchers at Stanford and New York University recently developed a privacy-preserving approach to targeted advertising, which they call Adnostic.[36] This proposed system would allow ad networks to engage in behavioral profiling and ad targeting but without having a server track consumers. Rather, all of the tracking and profiling necessary for serving targeted ads takes place on the client side, i.e., in the user's own browser.

report.pdf (defining targeted advertising as "the collection of information about a consumer's online activities in order to deliver advertising targeted to the individual consumer's interests"); *see also* FTC STAFF REPORT, *supra* note 4, at 63–69 (discussing the "do not track" option).

36. TOUBIANA ET AL., *supra* note 33.

> When a site wants to serve an interest-based ad, the user's browser chooses the most relevant ad from a portfolio of ads offered by the ad network service but the browser doesn't reveal this information to the ad service or to any third-party. Adnostic is a promising technology because it offers much greater privacy protections than privacy-friendly PETs while preserving much of the advertising business model.[37] On the other hand, Adnostic imposes new costs and complexity on the online advertising industry and arguably undermines the ability of different ad services to compete based on which of them has the best ad matching algorithms. Adnostic has not found any takers as of this writing and seems unlikely to do so absent much stronger regulatory incentives.

These main characteristics of the three categories of PETs are summarized in Table 1:

Table 1: Main Characteristics of PETs

Type of PET	Purpose	Examples	Incentives To Adopt
Substitute PET	Prevent tracking and profiling	Anonymous proxy servers; Tor Browsing Bundle	Weak due to high opportunity costs
Complementary: Privacy-Friendly	User control of online advertising	Ad-preference and cookie managers; advertising icons; "do not track" tools	Strong: PETs enhance user controls with minimal disruption of advertising business model
Complementary: Privacy-Preserving	Allow tracking and profiling without revealing user's preferences to third parties	Adnostic	Weak: Even though it supports the business model, Adnostic adds complexity and shifts control from advertisers to users

C. ANALYZING PRIVACY BY DESIGN

Privacy by design is an amorphous concept. At the very least, it means implementing FIPPs in the design and operation of products and services that collect, or in any way process, personal data. One way of accomplishing this is by using existing PETs or creating new ones in response to emerging privacy concerns. Alternatively, privacy by design may refer to the adoption of processes, systems, procedures, and policies—any of which may also have

37. *See also* ANN CAVOUKIAN, REDESIGNING IP GEOLOCATION: *PRIVACY BY DESIGN* AND ONLINE TARGETED ADVERTISING (2010), *available at* http://www.ipc.on.ca/images/Resources/pbd-ip-geo.pdf (discussing Bering Media's "doubleblind" privacy architecture).

a technological dimension—and which may be referred to collectively as privacy safeguards. EU privacy officials have long embraced PETs[38] but have begun to embrace a more expansive approach to privacy by design that emphasizes sound design practices as well.[39] In the United States, the FTC gives short shrift to PETs[40] and instead highlights a broad set of safeguards including certain design practices. The following discussion attempts to put some meat on these bones by analyzing the Staff Report in greater detail.

The Staff Report suggests that privacy by design consists of an integrated set of development and management processes and practices.[41] As with PETs, it is necessary to differentiate front-end software development activities from back-end data management practices. Front-end activities are a design process for customer-facing products and services (i.e., those with which customers interact by downloading software, using a web service, and/or sharing personal data or creating user content). Back-end practices consist of data management processes that ensure that information systems (for both internal use and for sharing data with affiliates, partners, and suppliers) comply with privacy laws, company policies (including published privacy policies), and customers' own privacy preferences. Although distinctive, the two lifecycles overlap in that most products and services designed for the Internet combine a front-end component with back-end data handling.[42]

The software development lifecycle seeks to ensure that in designing products and services, software developers take account of both customer privacy expectations and the relevant threat model that needs to be guarded against. This approach empowers users to control their personal data (for example, by improving their understanding of what information will be collected from them, how it will be used and what choices they have as to its

38. *See supra* note 3.

39. *See* EC, *A Digital Agenda for Europe*, *supra* note 1, at 17 n.21 (explaining that the principle of "privacy by design" means "that privacy and data protection are embedded throughout the entire life cycle of technologies, from the early design stage to their deployment, use and ultimate disposal").

40. For example, the FTC staff report on consumer privacy discusses privacy-friendly choice mechanisms for online behavioral advertising (including "do not track") but otherwise barely mentions any substitution or privacy-preserving PETs. FTC STAFF REPORT, *supra* note 4.

41. *Id.* at 41.

42. Cavoukian avoids distinguishing front-end software development activities from back-end data management practices and instead takes a more holistic approach. *See* Cavoukian, *supra* note 1, at 4 (noting that "the PbD concept can be applied at many levels, from specific technologies, to organizational practices, extending to entire information ecosystems and architectures").

transfer, storage, and use). At the same time, it seeks to minimize the risks of privacy incidents (such as surreptitious or unanticipated data collection, unauthorized data use, transfer or exposure, and security breaches). The data management lifecycle, on the other hand, focuses more on how firms should engineer and manage information systems with privacy in mind as firm employees access, use, disclose, and eventually delete customer data.

This front-end/back-end distinction is generally consistent with the chief concerns discussed in subsections V(B)(1) and (2) of the Staff Report. The former advises companies on "incorporating substantive privacy protections into their practices,"[43] while the latter recommends that companies maintain "comprehensive data management procedures."[44] Yet there are notable shortcomings in the Commission's analysis. For instance, there is a lack of detail describing software design guidelines and data management practices. Overall, the Staff Report lacks the more robust discussion of best practices and other actionable steps that companies require to deploy privacy by design effectively. The next two Sections elaborate upon these concerns.

1. Privacy by Design in the Private Sector: Front-End and Back-End Approaches

Several of the older and more well-established multinational IT companies have developed guidelines, policies, tools, and systems for building privacy into software development and data management. For example, Microsoft's Security Development Lifecycle ("SDL") for software development is the best-known example of how privacy can be built into the design process.[45] The SDL aims to integrate privacy and security principles into each of the five stages of the software development lifecycle (requirements, design, implementation, verification, and release).[46] Privacy impact ratings are given to each project and these ratings determine the

43. FTC STAFF REPORT, *supra* note 4, at 44.

44. *Id.* at 49.

45. *See* Steve Lipner & Michael Howard, Microsoft Corp., *The Trustworthy Computing Security Development Lifecycle*, MDSN (Mar. 2005), http://msdn.microsoft.com/en-us/library/ms995349.aspx. Microsoft claims—with some independent support—that when compared to software that has not been subject to the SDL, software that has undergone SDL processes has a significantly reduced rate of external discovery of security vulnerabilities. *See SDL Helps Build More Secure Software*, MICROSOFT, http://www.microsoft.com/security/sdl/learn/measurable.aspx (last visited Nov. 6, 2011). The Department of Homeland Security's Software Assurance program adopts a similar approach, which it refers to as "Build Security In." *See* Nat'l Cyber Sec. Div., Dep't of Homeland Sec. (DHS), *Build Security in Home*, https://buildsecurityin.us-cert.gov/bsi/home.html (last visited Nov. 6, 2011).

46. *See* MICROSOFT CORP., SIMPLIFIED IMPLEMENTATION OF THE MICROSOFT SDL 3, http://www.microsoft.com/download/en/details.aspx?displaylang=en&id=12379 (last updated Nov. 4, 2010).

design specifications needed for compliance.[47] The SDL guidelines are supplemented by Microsoft's "Privacy Guidelines for Developing Software and Services," a fifty-one-page document that lays out basic concepts and definitions based on FIPPs and related U.S. privacy laws; discusses different types of privacy controls and special considerations raised by shared computers, third parties, and other situations; and then enumerates nine specific software product and web site development scenarios.[48] For each scenario, the guidelines identify required and recommended practices relevant to notice and consent, security and data integrity, customer access, use of cookies, and additional controls or requirements.[49]

On the data management side, IBM's Tivoli Privacy Manager is a comprehensive enterprise privacy management system that supports a variety of privacy functionalities.[50] HP is also developing a comprehensive approach to managing the information lifecycle—storage, retrieval, usage, prioritization, update, transformation, and deletion—as well as identity management tasks such as the collection, storage, and processing of identity and profiling information, authentication and authorization, "provisioning" of digital identities (i.e., account registration and related tasks), and user management of personal data and identities. According to researchers in HP's Trusted Systems Lab, this requires both a model of privacy obligations (based on the rights of data subjects, any permission they have granted over the use of their personal data, and various statutory obligations associated

47. *Id.* at 10.

48. *See* MICROSOFT CORP., PRIVACY GUIDELINES FOR DEVELOPING SOFTWARE PRODUCTS AND SERVICES (ver. 3.1, Sept. 2008), http://www.microsoft.com/downloads/en/details.aspx?FamilyID=c48cf80f-6e87-48f5-83ec-a18d1ad2fc1f&DisplayLang=en (describing nine scenarios at length). Although these guidelines mainly treat privacy design issues for front-end products and services, they also address back-end services such as "Server Deployment." This implies that "front-end" and "back-end" are not exclusive categories so much as primary areas of focus. One of the very few comparably detailed sets of privacy guidelines is the European Privacy Seal ("EuroPriSe") for IT products and services, which has developed a fifty-nine-page document with four sets of detailed criteria that firms must satisfy to demonstrate compliance with the EU Data Privacy Directive. *See* EUROPRISE, EUROPRISE CRITERIA (2010), *available at* https://www.european-privacy-seal.eu/criteria/EuroPriSe%20Criteria%20201011.pdf.

49. *See generally* MICROSOFT CORP., *supra* note 48.

50. *See* Paul Ashley & David Moore, *Enforcing Privacy Within an Enterprise Using IBM Tivoli Privacy Manager for E-business*, IBM DEVELOPERWORKS (2002), http://www.ibm.com/developerworks/tivoli/library/t-privacy/index.html (describing functions such as tracking different versions of privacy policies; storing consent of the individual to the privacy policy when PII data is collected; auditing of all submissions and accesses to PII; and authorization of submissions and accesses to PII).

with FIPPs) and a framework for managing these obligations.[51] The resulting "obligation management system" enables enterprises to configure information lifecycle and identity management solutions to deal with the preferences and constraints dictated by privacy obligations and ideally to do so in an automated and integrated fashion.[52]

Although product development and data management emphasize different aspects of privacy by design, the goal of both approaches is roughly the same: to build in privacy protections using a combination of technological and organizational measures that ensure compliance with applicable rules. Over the past decade, computer scientists have begun to develop formal methods for extracting descriptions of rules from the policies and regulations that govern stakeholder actions,[53] formal languages for representing such rules,[54] and methods for enforcing such rules via software systems that perform run-time monitoring and post hoc audits to ensure that disclosure and use of personal information respects these rules.[55] As Breaux and Anton note in a paper using the HIPAA Privacy Rule as a model: "Actions that are permitted by regulations are called rights, whereas actions that are required are called obligations. From stakeholder rights and obligations, we can infer system requirements that implement these rules to comply with regulations."[56] The idea of using formal languages to align

51. *See* MARCO CASASSA MONT, HP LABS., ON PRIVACY-AWARE INFORMATION LIFECYCLE MANAGEMENT IN ENTERPRISES: SETTING THE CONTEXT 5–8 (2006), http://www.hpl.hp.com/techreports/2006/HPL-2006-109.pdf.

52. *Id.* at 7.

53. *See* Travis D. Breaux & Annie I. Anton, *Analyzing Regulatory Rules for Privacy and Security Requirements*, 34 IEEE TRANSACTIONS ON SOFTWARE ENGINEERING 5 (2008).

54. *See* A. Barth et al., *Privacy and Contextual Integrity: Framework and Applications*, *in* PROCEEDINGS OF 2006 IEEE SYMPOSIUM ON SECURITY AND PRIVACY 184 (2006), *available at* http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1624011 (describing a language for representation of rules based on Helen Nissenbaum's theory of contextual integrity and showing how to represent a collection of rules from several federal statutes using this language).

55. *See* D. GARG ET AL., A LOGICAL METHOD FOR POLICY ENFORCEMENT OVER EVOLVING AUDIT LOGS (Carnegie Mellon Univ., Technical Report No. CMU-CyLab-11-002, rev. May 6, 2011), *available at* http://arxiv.org/PS_cache/arxiv/pdf/1102/1102.2521v3.pdf. One challenge in automated enforcement of rules that appear in privacy regulations is that they sometimes include subjective concepts (e.g., related to beliefs of individuals). Such policies cannot be automatically enforced in their entirety, but recent results demonstrate that software systems can in fact support a best-effort enforcement regime by checking all parts of the rules that do not contain subjective concepts and outputting the rest for inspection by human auditors. I am grateful to Anapum Datta for this reference.

56. Breaux & Anton, *supra* note 53 (explaining that the 55-page HIPAA Privacy Rule yielded 300 stakeholder access rules, which in turn were comprised of 1,894 constraints); *see also* TRAVIS D. BREAUX & DAVID G. GORDON, REGULATORY REQUIREMENTS AS OPEN

privacy requirements of software systems with legal regulations no doubt exceeds anything that the FTC has in mind when it recommends that companies incorporate substantive privacy protections into their practices. On the other hand, requirements engineering, formal languages, and related tools and techniques are precisely what software developers need in order to transform privacy by design from a vague admonition (that it is better to build in privacy than to bolt it on later) into a planned and structured design process.

2. Privacy by Design in the Staff Report and FTC Enforcement Actions

In comparison to these front-end and back-end commercial approaches, which are both rich in detail and very comprehensive, or to the emerging discipline of requirements engineering, the discussion of privacy development guidelines in Section V(B)(1) seems incomplete. To begin with, it considers only four substantive privacy protections that firms should incorporate into their practices (security, collection limits, retention practices, and accuracy) but fails to explain why all eight FIPPs are not applicable.[57] Certainly, two of these other principles—purpose specification and use limitation—are highly relevant to building privacy protections into products and services. An equally serious omission of this section (but not of later sections of the report) is the failure to discuss common use scenarios or the rules that should govern them, the severity of threat associated with each of them, and the safeguards needed to address these threats consistent with customer expectations and legal requirements.[58] In Section V(B)(2), the report's guidance consists of two recommendations. First, that firms implement CIPPs, and second, that they assess risks (in a manner akin to

SYSTEMS: STRUCTURES, PATTERNS AND METRICS FOR THE DESIGN OF FORMAL REQUIREMENTS SPECIFICATIONS (Carnegie Mellon Univ., Technical Report No. CMU-ISR-11-100, 2010), http://reports-archive.adm.cs.cmu.edu/anon/isr2011/CMU-ISR-11-100.pdf (describing a formal requirements specification language that allows developers to turn regulations into computational requirements that they can "design and debug" using formal structures, patterns, and metrics, and validating the approach using state data breach notification laws).

57. *See, e.g.*, Hugo Teufel III, Privacy Policy Guidance Memorandum, *The Fair Information Practice Principles: Framework for Privacy Policy at the Department of Homeland Security*, DEP'T HOMELAND SECURITY (Dec. 29, 2008), http://www.dhs.gov/xlibrary/assets/privacy/privacy_policyguide_2008-01.pdf (identifying eight principles including purpose specification and use limitation).

58. In fact, sections V(C) and (D) of the FTC staff report on consumer privacy examine a number of scenarios involving choice, notice, access, and material changes. FTC STAFF REPORT, *supra* note 4, at 58–77. Unfortunately, the report does not incorporate this analysis into the discussion of privacy by design.

Privacy Impact Assessments ("PIAs")) "where appropriate." But these insights are not sufficiently developed to provide much useful guidance.

For example, the report neglects to define when risk assessments are appropriate. This is surprising considering that section 208(b)(1)(A) of the E-Government Act of 2002[59] offers relevant guidelines, requiring federal agencies to perform a privacy assessment prior to developing or procuring IT systems or projects that collect, maintain, or disseminate information in identifiable form from or about members of the public.[60] Although the Staff Report offers a few illustrations of privacy reviews (notably in its discussion of peer-to-peer file sharing) and some prescriptive guidance, it does not go far enough in providing detailed rules or requirements for privacy assessments to help companies determine when to conduct them or whether they have done so in a meaningful way. Of course, PIAs are the most widely used tool for privacy risk assessments, especially in the public sector.[61] Interestingly, the privacy Green Paper recently published by the Department of Commerce ("DOC") also encourages firms to use PIAs to enhance transparency, increase consumer awareness, and identify alternative approaches that would help to reduce relevant privacy risks.[62] But the Staff

59. Pub. L. No. 107-347, § 208(b)(1)(A), 116 Stat. 2899, 2921–22 (2002) (codified at 44 U.S.C. § 3501 (2006)).

60. *See* Joshua B. Bolten, *M-03-22, OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002*, OFF. MGMT. & BUDGET (Sept. 26, 2003), http://www.whitehouse.gov/omb/memoranda_m03-22 (further specifying when PIAs are required).

61. *See* Roger Clarke, *Privacy Impact Assessment: Its Origins and Development*, 25 COMP. L. & SECURITY 123, 129 (2009). Clarke defines a PIA as "a systemic process that identifies and evaluates, from the perspectives of all stakeholders, the potential effects on privacy of a project, initiative or proposed system or scheme, and includes a search for ways to avoid or mitigate negative privacy impacts." *See* Roger Clarke, *An Evaluation of Privacy Impact Assessment Guidance Documents*, 1 INT'L DATA PRIVACY L., no. 2, at 111 (2011), *available at* http://idpl.oxfordjournals.org/content/1/2/111.full.pdf. Clarke criticizes the Section 208 PIA process as mainly "checklist-based and almost entirely devoid of any content of significance to privacy protection, beyond the narrowly circumscribed legal requirements." *Id.* at 117.

62. INTERNET POL'Y TASK FORCE (IPTF), U.S. DEP'T OF COMMERCE (DOC), COMMERCIAL DATA PRIVACY AND INNOVATION IN THE INTERNET ECONOMY: A DYNAMIC POLICY FRAMEWORK 34–36 (2010), *available at* http://www.ntia.doc.gov/reports/2010/IPTF_Privacy_GreenPaper_12162010.pdf [hereinafter DOC GREEN PAPER]. The discussion cites a recent EC recommendation encouraging the RFID industry and relevant stakeholders to develop a framework to assess the privacy risks of using RFID applications, subject to endorsement by the Article 29 Working Party. *See Industry Proposal: Privacy and Data Protection Impact Assessment Framework for RFID Applications*, EUROPE'S INFO. SOC'Y (2010) (draft), *available at* http://ec.europa.eu/information_society/policy/rfid/documents/d31031industrypia.pdf. This 25-page proposed framework would require RFID operators to report the types of data that RFID tags and applications collect and process, including any personal or sensitive data; whether this information gives rise to particular privacy risks, such as tracking

Report discussion of privacy assessments is too brief to infer whether it concurs with the DOC's reasoning or would embrace the European model in which industry-wide PIAs must be reviewed and approved by privacy officials.

In sum, the Staff Report is best read as a first cut at agency guidance regarding privacy by design, with Sections V(B)(1) and (2) offering preliminary guidelines on how firms might integrate privacy safeguards into their development and data management practices. Other sources of guidance in the Staff Report include the discussion of "commonly accepted practices" in providing notice and choice[63] and how to increase transparency in data practices,[64] both of which suggest recommended practices in privacy by design.

Also instructive are some half-dozen "spyware" and "adware" enforcement actions suggesting prohibited design practices or required disclosure practices. In the prohibited category, the FTC has brought several cases involving the alleged practices of (1) installing software without a user's consent by exploiting security vulnerabilities; (2) bundling software with malware; and (3) installing root kit software. In the required category, several additional cases concern allegations of failing to clearly and conspicuously disclose (4) the bundling of free software with malware; (5) all the features of a program (such as content protection or "phone home" features); (6) the types of data that certain tracking software will monitor, record, or transmit; and (7) the means by which consumers may uninstall any adware or similar programs that monitor internet use and display frequent, targeted pop-up ads. These enforcement cases help flesh out the discussion in the Staff Report and constitute a down payment on privacy design guidelines in the form of prohibited, required, and recommended practices.[65]

---

an individual's movements; and to address the privacy and security features designed to minimize these risks, and whether the applications are ready for deployment (i.e., provide for suitable controls, practices, and accountability) or if a corrective action plan needs to be developed followed by a new PIA. Industry won the endorsement of the Working Party after revising its proposed framework in response to criticism. *See* Art. 29 Data Protection Working Party, 00327/11/EN, WP 180, *Opinion 9/2011 on the Revised Industry Proposal for a Privacy and Data Protection Impact Assessment Framework for RFID Applications* 3–4 (Feb. 11, 2011), *available at* http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp180_en.pdf.

63. *See* FTC STAFF REPORT, *supra* note 4, at 53–65.

64. *Id.* at 69–77.

65. The Appendix, *infra*, identifies the relevant cases and additional discussion in the FTC Staff Report and organizes them into a list of prohibited, required, or recommended privacy design practices.

Admittedly, none of this adds up to a complete version of what the FTC means by privacy by design, or—to use the broader notion—by CIPPs. But the Commission provides two hints of what future enforcement actions may bring. The first hint is discernible in the FTC's letter to Google closing the Street View investigation.[66] Despite its stated concerns regarding the adequacy of Google's internal review processes, the Commission chose to end this inquiry based on assurances that (1) Google neither had nor would use the Wi-Fi payload data and intended to delete it, and (2) that it would adopt certain practices "including appointing a director of privacy for engineering and product management; adding core privacy training for key employees; and incorporating a formal privacy review process into the design phases of new initiatives."[67] In addition, the Commission recommended that Google "develop and implement reasonable procedures" such as "collecting information only to the extent necessary to fulfill a business purpose, disposing of the information no longer necessary to accomplish that purpose, and maintaining the privacy and security of information collected and stored."[68] This closing letter clearly anticipates several themes in the Staff Report discussion of privacy by design. The second hint consists in the obvious similarities between CIPPs, as described in the Staff Report, and comprehensive information security programs ("CISPs"), as defined in the Safeguards Rule[69] and numerous FTC enforcement actions.[70] A recent consent agreement resolving allegations that Google engaged in deceptive trade practices when it launched its "Buzz" social networking service confirms that the Commission modeled CIPPs on CISPs, both as to their overall conception and specific elements.[71]

66. *See* Letter from David C. Vladeck, Director, Bureau of Consumer Protection, to Albert Gidari, Esq., Counsel for Google (Oct. 27, 2010), *available at* http://www.ftc.gov/os/closings/101027googleletter.pdf (closing Google inquiry). The Google Street View service displays panoramic images of many cities taken from cars equipped with specially adapted digital cameras and antennas. In April 2010, Google revealed that these cars had been inadvertently collecting data from Wi-Fi networks. *See* Kevin J. O'Brien, *New Questions over Google's Street View in Germany*, N.Y. TIMES, Apr. 29, 2010.

67. Letter from David C. Vladeck to Albert Gidari, *supra* note 66.

68. *Id.*

69. The Safeguards Rule, 16 C.F.R. pt. 314 (2010), implements the security and confidentiality requirements of the Gramm-Leach-Bliley Act ("GLBA"), Pub. L. No. 106-102, 113 Stat. 1338 (1999) (codified at 15 U.S.C. §§ 6801–6809 (2006)).

70. For a list of relevant cases, see FTC STAFF REPORT, *supra* note 4, at 10–11. For a recent example of an enforcement action defining a CISP, see Agreement Containing Consent Order, *In re* Twitter, Inc., File No. 0923093 (Fed. Trade Comm'n (FTC) June 24, 2010), *available at* http://www.ftc.gov/os/caselist/0923093/100624twitteragree.pdf.

71. *See* Agreement Containing Consent Order, *In re* Google, File. No. 102 3136 (FTC Mar. 30, 2011), *available at* http://www.ftc.gov/os/caselist/1023136/110330googlebuzz

Although both CISPs and CIPPs incorporate a mix of personnel and accountability measures, risk assessments (including consideration of product design), design and implementation processes, and ongoing evaluations, there are important respects in which the two programs differ. For example, privacy risk assessments are still in their infancy and have far fewer technical resources to draw upon than security risk assessments, which often take the form of threat modeling and rely on highly developed and well-established secure coding practices and testing tools.[72] Similarly, the FTC consent orders establishing CISPs and CIPPs require companies to submit periodic assessments from qualified professionals certifying that their programs operate effectively based on generally accepted procedures and standards. While in the security world such benchmarks exist, in the privacy world they do not, although this is changing.[73] Lastly, it is worth noting that while the

agreeorder.pdf. FTC consent orders resulting from data security incidents usually require the violating company to implement a comprehensive, written CISP that is (1) reasonably designed to protect the security, privacy, confidentiality, and integrity of personal information; and (2) contains administrative, technical, and physical safeguards appropriate to a company's size and complexity, the nature and scope of its activities, and the sensitivity of the personal information. *See In re Twitter*, Consent Order, *supra* note 70, at 3. Similarly, the Google consent order requires the company to implement a comprehensive, written CIPP that is (1) reasonably designed to address privacy risks and protect the privacy and confidentiality of personal information; and (2) contain privacy controls and procedures appropriate to the company's size and complexity. *See In re Google*, Consent Order, *supra*, at 4–6. Additionally, the five major constituents of each type of program are all but identical. The first element in both programs is "the designation of a responsible employee to coordinate and be accountable for" the program. The second element in both, "the identification of reasonably foreseeable, material risks," is similarly structured although each focuses on somewhat different dangers and requires assessments of different factors. The third element, the design and implementation of reasonable "safeguards" (CISPs) or "privacy controls and procedures" (CIPPs), and the "regular testing or monitoring of the effectiveness" of such safeguards or controls, is also the same in both. The fourth element in both programs calls for reasonable care in selecting and retaining service providers. The fifth element in both uses nearly identical language to require "the evaluation and adjustment" of the relevant program based on the results of the required "testing and monitoring . . . , any material changes to respondent's operations or business arrangements, or any other" relevant circumstances. *See In re Twitter*, Consent Order, *supra* note 70; *In re Google*, Consent Order, *supra*.

72. For a description of relevant tools and techniques, see generally MARK GRAFF & KENNETH VAN WYK, SECURE CODING: PRINCIPLES AND PRACTICES (2003); MICHAEL HOWARD & DAVID LEBLANC, WRITING SECURE CODE (2003); GARY MCGRAW, SOFTWARE SECURITY: BUILDING SECURITY IN (2006).

73. ISO/IEC 27002 is a widely acknowledged and well-established, certifiable information security standard published by the International Organization for Standardization ("ISO"). Although Subcommittee 27 ("SC 27"), IT Security Techniques, of the ISO's Joint Technical Committee 1 is working on several projects, including a "Privacy Framework," "Privacy Reference Architecture," and "Proposal on a Privacy Capability Assessment Model," international privacy standards remain at a very preliminary stage. *See IT*

Staff Report's discussion of CIPPs largely anticipates the obligations set forth in the Google settlement, the report endorses "privacy by design" while the consent decree avoids this language entirely, even though several of the prescribed elements of CIPPs include design aspects. It remains to be seen whether this omission is deliberate or signals a shift in how the FTC refers to and/or conceives of these requirements.

III. MARKET INCENTIVES

This Part addresses the question of whether the privacy market provides sufficient incentives for firms to invest in the elements of CIPPs (including privacy design and technology aspects) at a socially optimal level or if government intervention is needed to ensure appropriate investment. Many of the privacy regulators who endorse privacy by design seem confident that businesses will recognize the advantages of such investments and act accordingly. Thus, the U.K. Information Commissioner's Office ("ICO") insists that privacy by design will yield a "privacy dividend"[74] echoing Ann Cavoukian's earlier claim of a "privacy payoff" for firms that respect privacy and earn customer trust,[75] and her more recent assertion that "Full Functionality—*Positive-Sum*, not Zero-Sum" is a foundational principle of what she refers to as PbD.[76] But there are reasons to question their optimism.

To begin with, the orthodox economic view predicts that under perfect information, market forces will produce an efficient level of data collection and analysis. As a corollary, rational firms will invest in CIPPs in response to consumer demand for protection against the risks associated with data collection, unauthorized secondary use, processing errors, and improper access.[77] However, this view assumes that consumers understand how to recognize and protect themselves against both tangible harms, such as

Security Techniques, INT'L ORG. FOR STANDARDIZATION, http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_tc_browse.htm?commid=45306&development=on (last visited Nov. 6, 2011).

74. *See* U.K. INFO. COMM'R'S OFFICE, THE PRIVACY DIVIDEND: THE BUSINESS CASE FOR INVESTING IN PROACTIVE PRIVACY PROTECTION 3 (2010).

75. *See* ANN CAVOUKIAN & TYLER J. HAMILTON, THE PRIVACY PAYOFF: HOW SUCCESSFUL BUSINESSES BUILD CUSTOMER TRUST 36 (2002).

76. *See* Ann Cavoukian, Info. & Privacy Comm'r (Ontario, Canada), *Privacy by Design: The 7 Foundation Principles* 2 (revised Jan. 2011) (2009), *available at* http://www.ipc.on.ca/images/Resources/7foundationalprinciples.pdf. ("*Privacy by Design* seeks to accommodate all legitimate interests and objectives in a positive-sum 'win-win' manner . . . .").

77. *See* H. Jeff Smith & Sandra J. Milberg, *Information Privacy: Measuring Individuals' Concerns About Organizational Practices,* 20 MIS Q. 167 (1996) (identifying these four specific privacy dimensions, which represent the cognitive state of consumers towards corporate use of information).

identity theft or price discrimination, and intangible harms, which are harder to define in economic terms since they involve what Daniel Solove refers to as "digital dossiers" and the sense of "unease, vulnerability, and powerlessness" associated with them.[78] In fact, few consumers understand these risks and even fewer are familiar with PETs (or take the trouble to use them) or can easily identify firms with sound privacy programs.[79] Moreover, the weight of scholarly opinion suggests that this lack of awareness reflects information asymmetries and that this and related market failures are difficult to correct absent regulatory intervention.[80]

Second, firms contemplating how much to invest in privacy programs run up against several problems. In theory, establishing a CIPP, designing privacy into products and services, and/or deploying PETs should lower the risk of misuse or abuse of personal data, thereby reducing the probability and costs of any privacy breaches. Using a cost-benefit approach, firms would decide how much to invest by estimating and comparing the anticipated value of the benefits of avoiding such losses against the expected costs of privacy (and related security) safeguards. But the necessary data for these estimates is lacking and without it many firms instead lapse into a reactive mode, delaying needed investments until a privacy incident occurs or government regulation forces their hand.[81] Moreover, because firms profit from targeted advertising, personalization, and price discrimination, they are strongly motivated to collect and analyze as much customer data as possible with the fewest possible restrictions. Thus, certain PETs or privacy design decisions may impose opportunity costs that firms are reluctant to pay. Third, other reasons to make such investments—such as avoiding damage to

78. DANIEL J. SOLOVE, THE DIGITAL PERSON 149 (2004). More generally, Solove argues that privacy encompasses a range of problems that can create many different types of individual and societal harms, including financial losses, reputational harms, emotional and psychological harms, and relationship harms, to name a few. *See* DANIEL J. SOLOVE, UNDERSTANDING PRIVACY 174–79 (2008). *See also* Ryan Calo, The Boundaries of Privacy Harm (July 2010) (unpublished manuscript), *available at* http://works.bepress.com/m_ryan_calo/2 (arguing that privacy harms fall into two overarching categories: subjective harm (the unwanted perceptions of observation by others resulting in mental states such as anxiety, embarrassment, or fear) and objective harm (the "unanticipated or coerced use of information concerning a person against that person" such as "identity theft, the leaking of classified information that reveals an undercover agent, and the use of a drunk-driving suspect's blood as evidence against him")).

79. *See* LONDON ECON. STUDY, *supra* note 30, at 32–45.

80. *See, e.g.*, SOLOVE, THE DIGITAL PERSON, *supra* note 78, at 76–92; Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1265–68 (1998); Paul M. Schwartz, *Property, Privacy, and Personal Data*, 117 HARV. L. REV. 2055, 2076–84 (2004).

81. *See infra* Section III.C.

reputation and associated lost sales or customers—are not as compelling as they might seem.

As expected, industry defends its current practices quite vigorously, arguing that targeted ads provide consumers with useful information and underwrite free web content and services, and that advertisers use such information "anonymously."[82] Privacy advocates, on the other hand, strongly object to this rationale, calling attention instead to the potential harms associated with industry practices (such as the costs to consumers of price discrimination) and the advent of a dossier society.[83] In what follows, the goal is not to resolve these longstanding disputes or decide whether consumers would be better off if online advertisers were not only self-regulated but regulated by new privacy laws. Rather, the goal is to examine privacy investments in economic terms and decide if the market is or is not working.

A. WHY IS THERE WEAK DEMAND FOR CONSUMER PETS?

There is very little market data on the consumer demand for PETs, in part because they are not tracked as a separate product category. Anecdotal evidence exists regarding both substitute PETs and privacy-friendly PETs, and while inconclusive, it suggests that most PETs reach fewer than a million users.[84] The recent FTC Staff Report provided similar statistics on downloads or usage of popular ad-blocking tools.[85]

Are these numbers indicative of growing consumer demand for privacy tools to which companies should rationally respond by offering more PETs

82. *See* THOMAS M. LENARD & PAUL H. RUBIN, IN DEFENSE OF DATA: INFORMATION AND THE COSTS OF PRIVACY 2–3 (2009), *available at* http://www.techpolicyinstitute.org/files/in%20defense%20of%20data.pdf.

83. *See* Press Release, Ctr. for Digital Democracy (CDD), CDD and U.S. PIRG Call on FTC To Develop Stronger Online Privacy Framework (Feb. 18, 2011), http://www.democraticmedia.org/cdd-and-us-pirg-call-ftc-develop-stronger-online-privacy-framework.

84. *See* John Alan Farmer, *The Specter of Crypto-anarchy: Regulating Anonymity-Protecting Peer-to-Peer Networks*, 72 FORDHAM L. REV. 725, 754 (2003) (noting that an anonymity-protecting, peer-to-peer network had been downloaded over 1.2 million times since its launch in 1999); Steven Cherry, *Virtually Private*, IEEE SPECTRUM ONLINE (Dec. 1, 2006), *available at* http://spectrum.ieee.org/dec06/4744 (noting that an anonymous remailer had about 700,000 users in 1996); Kim Zetter, *Rogue Nodes Turn Tor Anonymizer into Eavesdropper's Paradise*, WIRED (Sep. 10, 2007), http://www.wired.com/politics/security/news/2007/09/embassy_hacks?currentPage=all (noting that "Tor has hundreds of thousands of users around the world").

85. *See* FTC STAFF REPORT, *supra* note 4, at 28. Mozilla.org estimates that AdBlock Plus has over 12 million "active daily users," which is much higher than anything in the FTC report. *See AdBlock Plus Statistics*, MOZILLA, https://addons.mozilla.org/en-US/statistics/addon/1865 (last visited Nov. 6, 2011).

or—alternatively—by building in privacy? Clearly, they are very small compared, for example, to popular anti-virus and related security products, which claim to have as many as 133 million users,[86] and miniscule compared to the nearly two billion worldwide Internet users.[87] The only contradictory data comes from a privacy official at Facebook, who recently indicated that almost thirty-five percent of the company's 350 million users customized their privacy settings when Facebook released new privacy controls in December of 2009.[88] This data may reflect user dissatisfaction with unpopular changes in Facebook's privacy controls; if not, it is an interesting development requiring further examination.

The most common explanation for the (apparently) weak demand for PETs is that due to information asymmetries, most individuals do not understand the risks to which they are exposed through sharing personal data.[89] Other commentators have noted the existence of a "privacy paradox" in that consumers both routinely state that they value their privacy highly yet behave as if their personal data has very little value.[90] Well-known examples of such behavior include consumers giving away personal data in exchange

86. *Internet Usage Statistics*, INTERNET WORLD STATS, http://www.internetworldstats.com/stats.htm. (last visited Nov. 6, 2011).

87. *Internet Usage in Europe*, INTERNET WORLD STATS, http://www.internetworldstats.com/stats4.htm. (last visited Nov. 6, 2011).

88. *See* FTC STAFF REPORT, *supra* note 4, at 28 n.68. For evidence that users will adjust their sharing behavior on social networks when user interfaces are augmented with visual or numerical displays of the size of the audience, see Kelly Caine et al., *Audience Visualization Influences Disclosures in Online Social Networks*, *in* PROCEEDINGS OF THE 2011 ANNUAL CONFERENCE EXTENDED ABSTRACTS ON HUMAN FACTORS IN COMPUTING SYSTEMS (ACM ed., 2011), *available at* http://dl.acm.org/citation.cfm?id=1979825&bnc=1.

89. *See* Alessandro Acquisti & Jens Grossklags, *What Can Behavioral Economics Teach Us About Privacy?*, *in* DIGITAL PRIVACY: THEORY, TECHNOLOGIES AND PRACTICES 363, 366–68 (Alessandro Acquisti et al. eds., 2007); *id.* at 364 ("Data subjects often know less than data holders about the magnitude of data collection and use of (un)willingly or (un)knowingly shared or collected personal data; they also know little about associated consequences.").

90. *See, e.g.*, Luc Wathieu & Allan Friedman, *An Empirical Approach to Understanding Privacy Valuation* (Harvard Bus. Sch., Working Paper No. 07-075, 2007), *available at* http://www.hbs.edu/research/pdf/07-075.pdf. In *U.S. West, Inc. v. Federal Communications Commission*, 182 F.3d 1224, 1234 (10th Cir. 1999), the court struck down on First Amendment grounds FCC regulations requiring customer opt-in approval prior to a telecommunications firm using their information for marketing purposes. In concluding that the FCC had failed to establish the protection of customer privacy as a "substantial interest," the court observed that it was insufficient to merely speculate that there are a substantial number of individuals who feel strongly about their privacy while at the same time assuming that they would not bother to opt-out even if given the chance. *Cf.* James P. Nehf, *Incomparability and the Passive Virtues of Ad Hoc Privacy Policy*, 76 U. COLO. L. REV. 1, 29–36 (2005) (arguing that a cost-benefit approach to valuing privacy inevitably favors the side seeking more data collection and sharing).

for loyalty cards, discounts, and other conveniences such as access to free content and services.[91] Privacy expert Alan Westin cites variable privacy sensitivities.[92] More recently, behavioral economists have developed explanations based on bounded rationality[93] and behavioral biases such as immediate gratification or optimism bias.[94]

Perhaps the most intuitively satisfying explanation of why people seem unwilling to look after their own privacy needs—whether through self-help or by demanding better privacy tools—comes from computer researchers Adam Shostack and Paul Syverson. They suggest that when people know they have a privacy problem (such as being on display to neighbors), they will pay for effective and understandable solutions (curtains and fences). But new situations like the Internet are harder to understand, a point they illustrate by reference to cookies:

> It is not trivial to understand what an http cookie is, as this requires some understanding of the idea of a protocol, a server, and statefulness. Understanding the interaction of cookies with traceability and linkability is even more complicated, as it requires understanding of web page construction, cookie regeneration, and non-cookie tracking mechanisms.[95]

---

91. *See* Janice Y. Tsai et al., *The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study*, 22 INFO. SYS. RES. 254, 255–56 (2011) (citing several relevant studies).

92. *See Opinion Surveys: What Consumers Have To Say About Information Privacy: Hearing Before the House Commerce Subcomm. on Commerce, Trade, and Consumer Protection of the H. Comm. on Energy and Commerce*, 107th Cong. 15–16 (2001) (testimony of Alan K. Westin, Professor, Public Law & Gov't Entities, Columbia Univ.; President, Privacy and Am. Bus.) (describing overall consumer privacy preferences as divided into three basic segments: *Privacy Fundamentalists* (25%), who reject offers of benefits, want only opt-in, and seek legislative privacy rules; *Privacy Unconcerned* (now down to 12% from 20% three years ago), who are comfortable giving their information for almost any consumer value; and . . . *Privacy Pragmatists* (63% or 125 million strong) [who] ask what the benefit is to them, what privacy risks arise, what protections are offered, and whether they trust the company or industry to apply those safeguards and to respect their individual choice").

93. *See* Acquisti & Grossklags, *supra* note 89, at 369–70 (noting that humans have limited ability to "process and act optimally on large amounts of data" and instead rely on simplified mental models).

94. *See* Alessandro Acquisti, *Privacy in Electronic Commerce and the Economics of Immediate Gratification*, *in* PROCEEDINGS OF THE 5TH ACM CONFERENCE ON ELECTRONIC COMMERCE 21, 22 (ACM ed., 2004) (highlighting "various forms of psychological inconsistencies (self-control problems, hyperbolic discounting, present-biases, etc.) that clash with the fully rational view of the economic agent").

95. Adam Shostack & Paul Syverson, *What Price Privacy? (and Why Identity Theft Is About Neither Identity Nor Theft)*, *in* ECONOMICS OF INFORMATION SECURITY 7 (L. Jean Camp & Stephen Lewis eds., 2004).

Unfortunately, it is all too easy to extend this analysis of the threat of cookies to other technologies consumers encounter in their everyday use of the Internet. In many cases, consumers lack awareness of tracking technologies or do not understand how the technology works when, for example, they visit a website that hosts "beacons" (invisible pixels that allow advertisers to track users as they surf the web), register for an online account, click on a banner ad, install a toolbar, use an ad-funded photo storage service, or use a mobile phone to locate a nearby store.[96] When they blog or share ideas, photos, or videos about themselves or their friends and relatives on a social network, they may have a better understanding about what they are doing while failing to fully appreciate the privacy implications of their actions. All of these cases require more insight and foresight about internet technology than most consumers have. Nor are there any "consumer reports" for privacy products and services that might assist them in evaluating the worth of a product or service.[97] This lack of an effective signaling mechanism to indicate "good" privacy practices has led one group of economists to conclude that online privacy suffers from adverse selection.[98]

B. WHY ARE FIRMS RELUCTANT TO INVEST IN PRIVACY BY DESIGN?

In deciding whether to invest in privacy by design, firms engage in a complex cost-benefit trade-off involving the direct, indirect, and opportunity costs of such investments; the effectiveness of various technologies and other privacy safeguards in reducing risks and associated losses; the demand for such technologies and safeguards; the competitive advantage gained by deploying them; and the opportunity costs associated with any technologies that may limit or prevent processing of personal data. The previous Section suggested that consumer demand is weak. This Section explores how firms go about budgeting for privacy expenditures in the face of weak consumer demand. An important caveat applies to this line of inquiry: most of the

96. For a discussion of the privacy (and security) implications of most of these activities, see GREG CONTI, GOOGLING SECURITY: HOW MUCH DOES GOOGLE KNOW ABOUT YOU? (2008).

97. Privacy seal programs seek to fill this role but have done so with only limited success.

98. *See* Tony Vila et al., *Why We Can't Be Bothered To Read Privacy Policies: Models of Privacy Economics as a Lemons Market*, *in* PROCEEDINGS OF THE 5TH INTERNATIONAL CONFERENCE ON ELECTRONIC COMMERCE 403, 404–05 (ACM ed., 2003), *available at* http://dl.acm.org/citation.cfm?id=948057 (suggesting this lemons market might be fixed by privacy signals that differentiate "good" sites and concluding that an efficient and reliable marketplace requires either privacy regulation or governments assuming the cost of testing signals; another possible solution is price discrimination).

relevant analysis and data originates in the literature on information security investments. This is unavoidable given the scarcity of reliable data on the costs of privacy.[99] For the sake of analysis, however, we will assume that firms approach both investments in roughly the same manner.[100]

This Section also examines a factor largely neglected when regulators make the business case for privacy by design—namely, the reasons that regulators think that businesses will benefit from adopting this approach. The missing factor is the opportunity costs businesses would incur if privacy design practices limit the scope of commercial exploitation of personal data.[101]

Economists who have analyzed how much firms should invest in information security generally agree on three points. The first is that cost-benefit analysis is a sound basis for decision making. Under this approach, firms must estimate both the costs and expected benefits of security activities, which in turn requires estimates of the potential losses from security breaches[102] and the probability of such breaches occurring. The second is that firms are more likely to utilize cost-benefit analysis if there is reliable data to inform the analysis. Here, however, that data on potential losses and their probability is hard to come by. The third is that in the absence of such data, many firms rely on alternatives to cost-benefit approaches such as incremental budget adjustments (i.e., adjusting the prior year's budget up or down based on possibly extraneous factors) or a more reactive approach (i.e., increasing investments in response to a breach event that makes security a must-do project).[103]

---

99. *See* LONDON ECON. STUDY, *supra* note 30, at 59.

100. This assumption may be justified given that at large corporations, chief privacy officers ("CPOs") and chief security officers ("CSOs") work closely and cooperatively and frequently sit in the same organization, or have similar reporting structures. Moreover, surveys of CPOs and CSOs suggest that in many cases, security issues may drive a CPO's objectives, while privacy issues may drive a CSO's. *See generally* ERNST & YOUNG, ACHIEVING SUCCESS IN A GLOBALIZED WORLD: IS YOUR WAY SECURE? (2006); IAPP/PONEMON, BENCHMARK PRIVACY: AN EXECUTIVE SUMMARY STUDY (2010). On the other hand, if weak security has clear economic costs and inadequate privacy does not, then perhaps firms will approach the relevant investment decisions in a different manner.

101. *But see* LONDON ECON. STUDY, *supra* note 30 (taking into account this factor).

102. These losses include direct losses, such as fraud, identity theft or interference with intellectual property rights; consequential losses, such as fines, penalties, and investigatory and remedial costs; and reputational damage, which may result in lost customers, sales, or profits.

103. *See, e.g.*, Lawrence Gordon & Martin Loeb, *Budgeting Process for Information Security Expenditures*, 49 COMM. ACM 121 (2006); Brent R. Rowe & Michael P. Gallaher, Private Sector Cyber Security Investment Strategies: An Empirical Analysis (Mar. 2006) (unpublished manuscript), *available at* http://weis2006.econinfosec.org/docs/18.pdf.

Assume for the sake of argument that these observations apply to privacy investment decisions as well. As noted, there is almost no data on the "benefits of privacy," i.e., any reliable estimates of the potential loss from a privacy incident or the probability that such incidents would occur. As for data on the "costs of privacy," the two available studies report very different results: the first suggests that large organizations spend from $500,000 to $22 million annually on overall privacy investments and that spending on privacy technology accounts for less than ten percent of the total budget (as compared to twenty-three percent and twenty-four percent devoted to a privacy office (staff and related overhead) and training programs, respectively).[104] The second study pegs this range at $500,000 to $2.5 million per year.[105] It is not clear if these figures are high or low when compared with the average security expenditures of a Fortune 500 firm.[106]

In the absence of data that would enable firms to use a cost-benefit approach in evaluating privacy investments, firms may decide not to invest in privacy by design due to opportunity costs, i.e., the costs attributed to technologies or other safeguards that may interfere with their current methods of collecting and analyzing customer data including such common practices as profiling and targeting. Indeed, opportunity costs might be thought of as the uninvited guests at the privacy by design pep rally. Standard economic doctrine teaches that firms will only care about privacy if that helps them increase their profits by attracting new customers.[107] There is some

104. *See* IBM & PONEMON INST., THE COSTS OF PRIVACY STUDY 13–14 (2004) (studying 44 large corporations, mostly Fortune 500 companies with between 5,000 and 75,000 employees).

105. *Cf.* IAPP/PONEMON, supra note 100 (finding that more than 70% of companies with over $10 billion in revenue reported privacy budgets between $500,000 and $2.5 million).

106. According to a recent survey that included data on total estimated information security budgets (including the cost of hardware, software, IT salaries, and consultants), 23% of respondents spent less than $100,000; 18% spent $100,000–$500,000; 18% spent $500,000–$2,000,000; and 10% spent $2,000,000–$10,000,000 (and the rest didn't know). *See* INFORMATIONWEEK, 2011 STRATEGIC SECURITY SURVEY: CEOS TAKE NOTICE 51 (2011), *available at* http://analytics.informationweek.com/abstract/21/6854/Security/research-2011-strategic-security-survey.html. Unfortunately, this data is not comparable to the privacy studies noted above because it is based on the responses of 1,084 business technology and security professionals at companies with more than 100 employees and does not segregate its findings by company size.

107. *See* LONDON ECON. STUDY, *supra* note 30, at 32–45; R. Böhme & S. Koble, On the Viability of Privacy-Enhancing Technologies in a Self-Regulated Business-to-Consumer Market: Will Privacy Remain a Luxury Good? (2007) (unpublished manuscript), *available at* http://www.inf.tu-dresden.de/~rb21/publications/BK2007_PET_Viability_WEIS.pdf; Joan Feigenbaum et al., Economic Barriers to the Deployment of Existing Privacy Technologies 2

experimental evidence of consumers' willingness to pay a privacy premium to online merchants with superior privacy practices even though they offer goods at higher prices.[108] Economists, including Alessandro Acquisti, have also speculated on whether privacy-enhanced identity management systems ("IDMs") may be used to enable consumers to interact pseudonymously with merchants while nevertheless allowing businesses to collect, analyze, and profitably exploit de-identified or aggregate data.[109] These are intriguing ideas, but Acquisti offers no evidence of commercial adoption despite the fact that the relevant technology has been available for many years.[110]

On the other hand, firms profit from collecting and analyzing customer data and are more likely than not to reject any privacy safeguards that would deprive them of this highly valuable information.[111] This data collection and analysis for online advertising purposes is big business.[112] According to

(2003) (unpublished manuscript), *available at* http://www.homeport.org/~adam/econbar-wes02.pdf.

108. *See* Serge Egelman et al., Timing Is Everything? The Effects of Timing and Placement of Online Privacy Indicators 1 (2009) (unpublished manuscript), *available at* http://www.guanotronic.com/~serge/papers/chi09a.pdf (follow-up study demonstrating that consumers are willing to pay more for a higher level of privacy when privacy indicators are presented alongside of search results); Tsai et al., *supra* note 91, at 255 (lab study demonstrating that consumers are willing to pay more to shop at websites that have better privacy policies).

109. *See* Alessandro Acquisti, *Privacy and Security of Personal Information: Economic Incentives and Technological Solutions*, *in* ECONOMICS OF INFORMATION SECURITY, *supra* note 95, at 7; Alessandro Acquisti, *Identity Management, Privacy, and Price Discrimination*, 6 IEEE SECURITY & PRIVACY 18 (2008); Böhme & Koble, *supra* note 107.

110. *See* Jan Camenisch et al., Position Paper, Credential-Based Access Control Extensions to XACML 4 (W3C Workshop on Access Control Application Scenarios, 2009), *available at* http://www.w3.org/2009/policy-ws/papers/Neven.pdf.

111. *See* Catherine E. Tucker, *The Economics Value of Online Customer Data* (WPISP & WPIE, Background Paper #1, Dec. 1, 2010), *available at* http://www.oecd.org/dataoecd/8/53/46968839.pdf (noting that online merchants and ad-funded Web businesses benefit from creating customer profiles based on clickstream data, cookies and Web bugs that track activities across the Web, demographic and behavioral data collected by specialized firms, data harvested from user-generated content on social networking and other Web 2.0 sites, and even more intrusive methods such as deep packet inspection).

112. *See id.* § 3.2.1 (citing a report by the Internet Advertising Bureau ("IAB") estimating that U.S. online advertising spending in 2009 reached $22.7 billion; a second IAB study suggesting that "ad-funded websites represented 2.1% of the U.S gross domestic product and directly employed more than 1.2 million people;" and a McKinsey study that used "conjoint" techniques to estimate that "in the U.S. and Europe consumers received 100 billion euros in value in 2010 from advertising-supported web services"). Similarly, Google published a study analyzing the total economic value received by U.S. advertisers, website publishers, and non-profits in 2010, which it estimated at $64 billion. *See* GOOGLE INC., GOOGLE'S ECONOMIC IMPACT: UNITED STATES | 2010 (2011), http://www.google.com/economicimpact/.

Tucker, online advertising is highly dependent on targeting, which uses customer profiles to find the particular ads most likely to influence a particular customer. Moreover, targeting increases the value of advertising to firms because they no longer have to pay for wasted eyeballs. Indeed, in 2009 the price of behaviorally targeted advertising was estimated at 2.68 times the price of untargeted advertising.[113]

In sum, ad targeting is valuable and privacy safeguards may increase opportunity costs to the extent that they diminish the economic value of online advertising, thereby creating an investment disincentive for firms dependent on advertising revenues. This disincentive may be offset by investments in privacy safeguards if they enable firms to attract new, privacy-sensitive customers or charge them higher prices, but there is scant evidence of this happening.

C. DO REPUTATIONAL SANCTIONS DRIVE PRIVACY INVESTMENTS?

Are firms sufficiently concerned about the reputational harms associated with high-profile privacy incidents to increase their investments in privacy technology? Although there is little data on firm expenditures in response to privacy meltdowns, the data on the reputational impact of security breach notifications is worth examining. More than forty-five states have enacted laws requiring that companies notify individuals of data security incidents involving their personal information.[114] These disclosures result in what Schwartz and Janger call "useful embarrassments" because they force businesses to invest ex ante in data security to avoid reputational sanctions including both diminished trust and potential loss of customers.[115] The Ponemon Institute has studied the costs of data breaches in the United States over the past several years and reports that in 2009, data breaches cost companies an average of $6.75 million per incident and $204 per

113. *See* Tucker, *supra* note 111, § 3.2.2 (citing Howard Beales, *The Value of Behavioral Targeting*, NETWORK ADVERTISING INITIATIVE, http://www.networkadvertising.org/pdfs/Beales_NAI_Study.pdf (last visited Mar. 5, 2012)); *see also* LENARD & RUBIN, *supra* note 82, at 14–18; Avi Goldfarb & Catherine E. Tucker, *Privacy Regulation and Online Advertising*, MGMT. SCI., Jan. 2011, at 57 (finding based on survey results that EU privacy regulation reduces the effectiveness of online advertising by restricting advertisers' ability to collect data on users for ad targeting purposes).

114. According to the National Council of State Legislatures (NCSL), "[f]orty-six states, the District of Columbia, Puerto Rico and the Virgin Islands have enacted legislation requiring notification of security breaches involving personal information." *See State Security Breach Notification Laws*, NAT'L CONF. ST. LEGISLATURES, http://www.ncsl.org/Default.aspx?TabId=13489 (last updated Oct. 12, 2010).

115. *See* Paul M. Schwartz & Edward J. Janger, *Notification of Data Security Breaches*, 105 MICH. L. REV. 913, 936 (2007).

compromised record.[116] Over 70% of the latter amount related to indirect costs including "abnormal turnover, or churn of existing and future customers" (down from 75% in 2008); these companies also suffered an average increased churn rate of 3.7% (up from 3.6% in 2008).[117] On the other hand, empirical evidence suggests that the cost of reputation loss (in terms of stock market impact) following incidents of data loss is statistically significant but relatively low in monetary terms and dissipates quickly.[118]

Although several commentators treat these studies as evidence that reputational sanctions pressure companies into improving their security practices,[119] Schwartz and Janger take a more cautious approach. As they note, the influence of reputational sanctions on data security can be quite complex. First, smaller firms and "bad apples" generally are less sensitive to reputational concerns.[120] Second, if sanctions rely on self-reporting, this may create a disincentive for reporting. Third, reputational sanctions are ineffective without "a well-functioning consumer-side market for data security."[121] Switching costs and lack of information about how firms manage data security undermine this market, notwithstanding whatever knowledge customers may derive from receiving, reading, and understanding breach notices.[122] In addition, there are a few drawbacks to the methods relied on in the Ponemon study—for example, it bases churn rates on company estimates, not on a survey of how many customers changed to another firm

116. *See* PONEMON INST., 2009 U.S. COST OF A DATA BREACH STUDY 5 (2009) (based on 45 respondents).

117. *Id.*

118. *See* Alessandro Acquisti, et al., Is There a Cost to Privacy Breaches? An Event Study 13–14 (2006) (unpublished manuscript), *available at* http://weis2006.econinfosec.org/docs/40.pdf (finding a cumulative drop in share prices per privacy incident of close to 0.6% on the day following the event, which equates to an average loss of approximately $10 million in market value).

119. *See* Kenneth A. Bamberger & Deirdre K. Mulligan, *Privacy on the Books and on the Ground*, 63 STAN. L. REV. 10, 147 (2011); SAMUELSON LAW, TECH. & PUB. POLICY CLINIC, SECURITY BREACH NOTIFICATION LAWS: VIEWS FROM CHIEF SECURITY OFFICERS 13–21 (2007), *available at* http://groups.ischool.berkeley.edu/samuelsonclinic/files/cso_study.pdf. *Cf.* Sasha Romanosky & Alessandro Acquisti, *Privacy Costs and Personal Data Protection: Economic and Legal Perspectives*, 24 BERKELEY TECH. L.J. 1061 (2009) (arguing that state breach disclosure laws have only a very weak effect on the incidence of data loss).

120. Schwartz & Janger, *supra* note 115, at 930–31.

121. *Id.* at 944.

122. *Id.* at 947.

following a breach disclosure,[123] and it fails to explain the variance from the pre-breach churn rate or what other possible co-factors might exist.[124]

Even assuming that reputational sanctions help bring about increased security expenditures, there is reason to question their impact on privacy investments. An obvious difference is that while unauthorized access to personal data triggers existing breach notification laws, there are no laws requiring notification of privacy incidents *other than* data breaches.[125] In the absence of laws mandating disclosure of such matters, businesses are disinclined to self-report their privacy failures. Although investigative journalists and privacy activists may take up the slack, even if they do a good job the net result is that less data is available on how customers react to privacy incidents and whether firms respond to customer backlash by investing more in privacy safeguards. And this lack of data makes empirical study quite difficult.[126] One result is that there are no studies on the costs of privacy failure akin to the Ponemon series on data breaches. At the same time, the other factors noted by Schwartz and Janger remain in place. Thus, small firms and "bad apples" will free ride on the reputational efforts of larger firms, while information asymmetries and behavioral biases prevent consumers from understanding how a privacy incident might affect them or

123. *See* PONEMON INST., *supra* note 116, at 11, 35 (noting that the study required each company contact person to estimate opportunity costs based on her professional experience).

124. *See* ADAM SHOSTACK & ANDREW STEWART, THE NEW SCHOOL OF INFORMATION SECURITY 190 (2008). For an interesting counterpoint to the Ponemon study, see Larry Dignan, *The TJX Data Breach: Why Loss Estimates Are Overblown*, ZDNET: BETWEEN THE LINES (May 8, 2007), http://www.zdnet.com/blog/btl/the-tjx-data-breach-why-loss-estimates-are-overblown/5000 (noting that anticipated "brand impairment" was less severe than expected); Jaikumar Vijayan, *One Year Later: Five Takeaways from the TJX Breach*, COMPUTERWORLD (Jan. 17, 2008), http://www.computerworld.com/s/article/9057758/One_year_later_Five_takeaways_from_the_TJX_breach (noting that TJX's comparable-store sales increased 4% in the year following the breach).

125. These other incidents may range from objectionable data collection practices (such as profiling or targeting) to unauthorized secondary use of personal data to various processing errors that may lead to economic or non-economic harm. *See* Smith & Milberg, *supra* note 77. Although Acquisti et al. title their study "Is There a Cost to Privacy Breaches? An Event Study," they limit their analysis to data breaches. As a result, their findings have little bearing on the reputational costs of privacy incidents that fall beyond the scope of security breach notification laws.

126. *See* LONDON ECON. STUDY, *supra* note 30, at 52 (noting that "[d]espite the importance of reputation, relatively little reliable empirical work has been undertaken to measure its value in the context of privacy. This is partly because it is difficult to measure the value of an intangible asset such as reputation. But, it is also difficult to obtain good quality data on the costs of reputation loss (e.g. through privacy breaches) since firms may be unwilling or unable to quantify their losses"). *See* SHOSTACK & STEWART, *supra* note 124, at 74–76, 149–53 (discussing the value of breach data in understanding information security).

what they can do about it. The point is that in the absence of a well-functioning consumer-side market for privacy safeguards, firms will remain reluctant to spend more on PETs or privacy by design, notwithstanding potential reputational sanctions.[127]

What about longstanding industry forecasts suggesting that firms lose billions of dollars in online sales due to privacy concerns?[128] It is unclear whether this truly happens. As noted, consumers' self-reported attitudes about the high importance of privacy to their online shopping decisions do not always match their actual behavior. To the contrary, many consumers (all of Westin's "unconcerned" and at least some of his "pragmatists") seem willing to trade away privacy for discounts or convenience.[129] This is not to say that firms are—or should be—indifferent to their reputation for privacy and trustworthiness. Firms do seem to care, not only because consumer perceptions have some impact on sales and profits, but because any rational firm would prefer to avoid the expenses associated with a major privacy incident. These include legal fees, call center staffing costs, lost employee productivity, regulatory fines, diminished customer trust, and potential customer desertions, all of which can be costly.[130]

And yet the impact of these reputational sanctions on investments in privacy technology remains ambiguous. A spate of recent privacy incidents—in years past, from Microsoft (Word, Windows Media Player, Passport), and more recently from Google (Gmail, Search, Street View, Buzz), Facebook (Beacon, Newsfeed), and Apple (iPhone locational-tracking data)—raises similar concerns about transparency, notice, choice, and data retention. Advocates respond to these incidents in similar ways, with public outcries, open letters, and complaints to regulators. Newspapers publish major stories and editorials, privacy officials open investigations and issue opinions, and a few customers file class action law suits. But the outcomes in terms of investments in privacy safeguards vary widely, suggesting that negative publicity may be important to increased investments only when accompanied by two additional factors: sustained attention by government officials and a blatant violation of users' expectations that provokes an

127. Of course, this may vary by business sector, with more regulated industries or professions showing a greater willingness to invest in remediation of privacy breaches than a typical Internet firm.

128. *See* ALESSANDRO ACQUISTI, THE ECONOMICS OF PERSONAL DATA AND THE ECONOMICS OF PRIVACY 21 (2010).

129. *See supra* Section III.A.

130. Google recently paid $8.5 million to settle lawsuits concerning its Buzz service, *see* Damon Darlin, *Google Settles Suit over Buzz and Privacy*, N.Y. TIMES, Nov. 23, 2010, and will no doubt incur additional costs in complying with the FTC consent agreement.

immediate outcry (as when firms cross the invisible boundary between appropriate and inappropriate data sharing).[131] This requires further empirical study and analysis but is beyond the scope of this paper.[132]

IV. RECOMMENDED REGULATORY INCENTIVES

The previous Part concluded that economic incentives are not enough to increase firm investments in privacy safeguards. To summarize, the weak consumer demand for PETs, the opportunity costs to businesses associated with many PETs, and a lack of relevant data needed for cost-benefit analyses of investments in privacy safeguards all work against the further implementation of PETs in the marketplace. As noted, reputational sanctions do play a role especially when firms are also subject to sustained attention by regulators or cross a subtle boundary beyond which certain data processing practices are vigorously opposed by the general public. In these cases, even internet giants like Microsoft, Google, Facebook, and Apple are forced to retreat and to modify or withdraw disputed features.

Does this imply that self-regulation is working, or is government intervention still needed? Over the past twelve months, Congress has considered or introduced new privacy legislation, ranging from narrow bills that would mainly protect consumers against online tracking to omnibus privacy bills.[133] In anticipation of these bills, industry has unveiled new self-regulatory initiatives including both voluntary codes of conduct from the advertising industry and privacy-friendly tools from search firms, network

131. *See* HELEN NISSENBAUM, PRIVACY IN CONTEXT: TECHNOLOGY, POLICY AND THE INTEGRITY OF SOCIAL LIFE (2010) (analyzing privacy as "contextual integrity," which she defines in terms of appropriate information flows).

132. The author is undertaking a series of studies relating to privacy by design including empirical work that may shed light on these issues.

133. Narrow bills: *See* Staff of Richard Boucher, Discussion Draft of House Bill To Require Notice to and Consent of an Individual Prior to the Collection and Disclosure of Certain Personal Information Relating to That Individual, 111th Cong. § 3(e) (May 3, 2010), *available at* http://dataprivacy.foxrothschild.com/uploads/file/Privacy_Draft_5-10.pdf (exempting network advertisers from having to obtain explicit, opt-in consent to engage in online tracking provided they allow consumers to access and manage their profiles); Do Not Track Me Online Act of 2011, H.R. 654, 112th Cong. (directing the FTC to develop standards for a "do not track" mechanism allowing individuals to opt out of the collection, use or sale of their online activities and requiring covered entities to respect the consumer's choice). Omnibus bills: *See* Best Practices Act, H.R. 611, 112th Cong. (2010); Commercial Privacy Bill of Rights Act of 2011 (the Kerry-McCain bill), S. 799, 112th Cong., *available at* http://www.govtrack.us/congress/bill.xpd?bill=s112-799.

advertisers, and browser vendors.[134] It remains to be seen whether these activities will be successful in warding off new legislation.[135]

On the other hand, privacy advocates reject these self-regulatory efforts as too little and too late. They argue that government intervention is needed to correct privacy market failures, implying that the demand for privacy safeguards will remain low and that firms will not increase their investments absent new legislation.[136] Accordingly, they insist that Congress at long last enact comprehensive legislation establishing baseline privacy requirements for online and offline data processing practices and empower the FTC to engage in rulemaking.[137] Of course, new default privacy rules may correct market failures but will also constrain profit-making activities at a significant cost to firms and the public.

In a recently published article, I suggested that self-regulation and prescriptive government regulation should not be viewed as mutually exclusive options from which policy makers are forced to choose. This is a false dichotomy and ignores the wide variety of co-regulatory alternatives that could be playing a larger role in the privacy arena.[138] Drawing on this earlier work and that of privacy scholars Kenneth Bamberger and Deirdre Mulligan, this Article concludes with a number of recommendations for how regulators might achieve better success in promoting the use of privacy by design by identifying best practices and/or situating these best practices within an innovative regulatory framework. This analysis considers co-regulatory solutions in two distinct environments: first, where Congress fails to enact new legislation but the FTC continues to play an activist role in defining CIPPs; and second, where Congress enacts a new privacy law

134. *See* Tanzina Vega, *Google and Mozilla Announce New Privacy Features*, N.Y. TIMES (Jan. 24, 2011), http://nyti.ms/y0g3fb (describing new "do not track" features by Google, Mozilla, and Microsoft as well as several self-regulatory programs).

135. *See* John Eggerton, *Q&A with FTC Chairman Jon Leibowitz*, MULTICHANNEL NEWS (Feb. 21, 2011), *available at* http://www.multichannel.com/article/print/464262-Privacy_Please_Q_A_With_FTC_Chairman_Jon_Leibowitz.php (noting that "the business community really has it in its hands to avoid regulation, it just has to step up to the plate").

136. *See* CDD Press Release, *supra* note 83.

137. *See* Juliana Gruenwald, *Lawmakers Looking for Right Balance on Privacy*, NATIONALJOURNAL TECH DAILY DOSE (Mar. 16, 2011), http://techdailydose.nationaljournal.com/2011/03/lawmakers-looking-for-right-ba.php.

138. *See* Ira S. Rubinstein, *Privacy and Regulatory Innovation: Moving Beyond Voluntary Codes*, 6 I/S: J.L. & POL'Y INFO. SOC'Y 355, 358 (2011) (noting that "[i]n co-regulatory approaches, industry enjoys considerable flexibility in shaping self-regulatory guidelines, while government sets default requirements and retains general oversight authority to approve and enforce these guidelines").

making FIPPs broadly applicable to firms that collect PII and possibly authorizing the FTC to establish a co-regulatory safe harbor program.[139]

A. THE FTC AS PRIVACY REGULATOR

If Congress enacts a new privacy law requiring firms to integrate privacy into their regular business operations and at every stage of the product development and data management lifecycle, and authorizing FTC rulemaking, then the Commission would address privacy by design by issuing implementing regulations. This would be quite analogous to the FTC drafting a rule covering the security and confidentiality requirements of financial institutions under the Gramm-Leach-Bliley Act ("GLBA").[140] If new legislation is not enacted, does the Commission already have authority under section 5 of the FTC Act to define the elements of CIPPs and require commercial firms to implement them? The short answer is yes—with a caveat.

Section 18 of the FTC Act grants the Commission *limited* authority to prescribe rules defining "unfair or deceptive acts or practices in or affecting commerce."[141] For better or worse, these procedures are burdensome and time-consuming as compared to conventional Administrative Protection Act ("APA") rulemaking.[142] As a result, the Commission often prefers to rely on strategic enforcement actions to achieve its regulatory goals, which is the procedure it followed in developing information security programs applicable to commercial firms.[143] In laying the foundations for CIPPs, the Commission has also relied on its section 5 powers to issue agency guidance regarding

139. For examples of a privacy bill providing safe harbor option, see the omnibus bills cited *supra* note 133. For a broader discussion of "co-regulatory" safe harbors and what they might contribute to the privacy debate, see generally Rubinstein, *supra* note 138, at 405–20 (arguing that such programs incentivize organizations to meet high standards of data protection by shielding safe harbor participants from various "sticks" such as a private right of action, and rewarding them with various "carrots" such as allowing greater flexibility in how they implement statutory requirements).

140. *See supra* note 69.

141. *See* 15 U.S.C. § 57a(a)(2) (2006).

142. *See* § 57a(b)(1), (2) (requiring that, before engaging in rulemaking, the FTC provide advance rulemaking notice to Congress and the public, hold public hearings at which interested parties have limited rights of cross-examination, and submit a statement of basis and purpose addressing both the prevalence of the acts or practices specified by the rule and its economic effect). Congress imposed these additional requirements on the Commission in 1980 in response to perceived abuses of the agency's rulemaking authority. *See generally* JULIAN O. VON KALINOWSKI ET AL., ANTITRUST LAWS AND TRADE REGULATION § 5.14 (1997).

143. *See* FTC STAFF REPORT, *supra* note 4, at 10–11.

commercial privacy practices. This has proven a flexible and effective tool.[144] Overall, this combination of strategic enforcement and agency guidelines developed in collaboration with industry demonstrates the FTC's "ability to respond to harmful outcomes by enforcing evolving standards of privacy protection" in keeping with changes in "the market, technology, and consumer expectations."[145]

Building from this foundation, the Commission can and should supplement the small number of enforcement cases related to privacy design practices by pursuing a strategic enforcement strategy. Indeed, it should look for cases that would further refine the core elements of CIPPs by establishing more prohibited, required, and recommended practices. This is necessary both because the analogy between CIPPs and CISPs is imperfect at best and the underlying design, coding, and testing practices for the former are far less developed than those of the latter. The Commission should consider several additional steps such as (1) convening a new round of workshops at which experts from industry, academia, and advocacy organizations identify useful PETs and discuss best practices in privacy by design, followed by a staff report and other guidance as appropriate; (2) supporting ongoing efforts by the ISO and others to define international privacy design standards; and (3) working with the National Institute of Standards or other federal agencies to fund research in requirements engineering, formal languages, and related tools and techniques that would transform privacy by design from a rallying cry into an engineering discipline.

B. REGULATORY INNOVATION

On the other hand, if Congress enacts new privacy legislation and authorizes the FTC to issue implementing regulations, this would open up several new pathways for regulatory innovation ranging from company-specific experimentation with new technologies and engineering practices to multi-stakeholder agreements on how to implement "do not track" practices to flexible safe harbor arrangements. This Section briefly examines several steps that the FTC should take if it is granted new regulatory authority.

1. Project XL for Privacy

The FTC should borrow a page from the environmental regulatory playbook by sponsoring a "Project XL for Privacy."[146] In a nutshell, Project

144. Bamberger & Mulligan, *supra* note 119, at 128–29.

145. *Id.*

146. *See* Rubinstein, *supra* note 138, at 374–76 (describing Project XL generally); *id.* at 406–10 (describing a modified version of Project XL attuned to the needs of privacy regulation).

XL is a program under which the Environmental Protection Agency ("EPA") negotiates agreements with individual firms to modify or relax existing regulatory requirements in exchange for enforceable commitments to achieve better environmental results. While these projects come in several flavors, the most useful for present purposes is the experimental XL project, in which the EPA takes the lead in identifying an innovative regulatory approach or technology and testing it out in a small number of pilot projects subject to rigorous evaluation by the EPA and other stakeholders. Conceived of as experiments from the outset, these projects may have industry-wide implications if they succeed or they may be abandoned if they fail to yield better results.

An obvious candidate for experimental XL projects for privacy might be in the area of privacy decision making. Several of the proposed privacy bills include lengthy and detailed notice requirements. These provisions are motivated by a desire to inform consumers of all relevant practices concerning personal data in a clear and conspicuous manner and to ensure that important information is not unduly vague or buried away. These efforts at ensuring rigorous and complete privacy notices are at once understandable and regrettable: no doubt many web sites and merchants engage in unfair or deceptive notice practices and yet more prescriptive notice requirements are not the remedy for the underlying problems, which range from asymmetric information to lack of readability to limited comprehension to consumer inertia.[147] However, researchers have developed a variety of tools to make privacy information more usable to consumers, such as standardized, easy-to-read privacy notices akin to nutrition labeling on food,[148] usability enhancements to P3P,[149] and a search engine that orders search results based on their computer-readable privacy policies.[150] The FTC should encourage firms to adopt these privacy-friendly PETs in exchange for regulatory relief on otherwise overly prescriptive notice requirements.

147. *See* Aleecia M. Mcdonald et al., *A Comparative Study of Online Privacy Policies and Formats*, *in* PROCEEDINGS OF THE 9TH INTERNATIONAL SYMPOSIUM ON PRIVACY ENHANCING TECHNOLOGIES (Ian Goldberg & Mikhail J. Atallah eds., 2009) and related studies cited therein; Egelman et al., *supra* note 108, at 1 (noting that "these policies rarely help consumers because they often go unread, or do not address the most common consumer concerns, [or] are difficult to understand").

148. PATRICK G. KELLEY ET AL., STANDARDIZING PRIVACY NOTICES: AN ONLINE STUDY OF THE NUTRITION LABEL APPROACH (Carnegie Mellon Univ., Report No. CMU-CyLab-09-014, 2010), *available at* http://www.cylab.cmu.edu/files/pdfs/tech_reports/CMUCyLab09014.pdf.

149. *See* PRIVACY BIRD, http://www.privacybird.org/ (last visited Nov. 8, 2011).

150. *See Frequently Asked Questions*, PRIVACY FINDER, http://www.privacyfinder.org/faq (last visited Nov. 8, 2011).

2. Negotiated Rulemaking

Congress may enact one of several pending bills that include a "do not track" requirement. If it does so and authorizes the FTC to promulgate a rule implementing a "do not track" provision, the FTC should forgo conventional rulemaking in favor of negotiated rulemaking.[151] In conventional FTC rulemaking—as exemplified by the rulemaking in the Children's Online Privacy Protection Act ("COPPA")[152]—the Commission first issues a Notice of Proposed Rulemaking ("NPR") soliciting comments from interested parties. Next, it conducts a review of the issues raised by these comments, which also includes holding a public workshop to obtain additional information regarding specific issues from industry, privacy advocates, consumer groups, and other government agencies. Lastly, the Commission publishes a Final Rule, which includes the agency's analysis of the public comments (which are also published) and its reasons for accepting or rejecting changes proposed pursuant to the NPR.[153]

Negotiated rulemaking, on the other hand, is a statutorily-defined alternative to conventional rulemaking in which agencies are granted the discretion to bring together representatives of the affected parties in a negotiating committee for face-to-face discussions. If the committee achieves consensus (defined as unanimous concurrence unless the committee agrees on a different definition such as general concurrence), the agency can then issue the agreement as a proposed rule subject to normal administrative review processes; but if negotiations fail to reach consensus, the agency may proceed with its own rule.[154]

Why might it be desirable to negotiate a "do not track" rule rather than rely on conventional rulemaking?[155] The core insight underlying negotiated rulemaking is that conventional rulemaking discourages direct communication among the parties, often leading to misunderstanding and even costly litigation over final rules. In contrast, the promise of negotiated

151. *See* Rubinstein, *supra* note 138, at 410–14 (describing negotiated rulemaking generally); *id.* at 377–80 (describing the application of negotiated rulemaking to the privacy issues associated with online behavioral advertising).

152. Pub. L. No. 105-277, 112 Stat. 2581 (1998) (codified at 15 U.S.C. §§ 6501–6506 (2006)).

153. *See COPPA Rulemaking and Rule Reviews*, FTC BUREAU CONSUMER PROTECTION BUS. CTR., http://business.ftc.gov/documents/coppa-rulemaking-and-rule-reviews (last updated Sept. 15, 2011).

154. *See generally* Negotiated Rulemaking Act of 1990 (NRA), Pub. L. No. 101-648, § 2(3)–(5), 104 Stat. 4969, 4969 (codified as amended at 5 U.S.C. §§ 561–570 (2006)).

155. The following treatment draws from a more detailed discussion of regulatory options, including their strengths and weaknesses, in Rubinstein, *supra* note 138, at 412–14.

rulemaking is that by enlisting diverse stakeholders in the rulemaking process, responding to their concerns, and reaching informed compromises, better-quality rules will emerge at a lower cost and with greater legitimacy. Negotiated rulemaking works best when the underlying rule requires information sharing between the regulators, the regulated industry, and other affected parties, and when the parties believe they have something to gain from working together and achieving a compromise.[156] Arguably, these conditions would be met if the FTC formed a negotiated rulemaking committee to tackle a "do not track" rule.

Clearly, the parties would come to the table with different views. Industry would hope to minimize any burdens on its ability to collect and analyze the data needed for ad targeting, thereby maintaining the free flow of information. For example, it might suggest that privacy-friendly PETs suffice to achieve legislative goals. Advocates seeking better and more effective protection against profiling and targeting might demand that any opt-out mechanisms be turned on by default as opposed to requiring user-initiated action,[157] or they might otherwise require that industry adopt privacy-preserving PETs. These differences are deep-seated and perhaps ideological, and thus not easily overcome. Yet there is reason to believe that all of the affected parties—the regulated industry, the advocates representing the public interest, and the regulators—might be highly motivated to engage in face-to-face negotiations and would benefit from the information sharing that inevitably occurs in this setting.

As to motivation, industry may be concerned about whether the FTC lacks the necessary expertise to understand the complex technologies and business models underlying online advertising; and, if not, whether the Commission might issue a "do not track" rule lacking in flexibility and nuance with highly negative results for industry revenues and profitability. They may also fear that in the wake of new legislation, the Commission will pursue a more aggressive enforcement strategy. Advocates may worry that even if Congress enacts "do not track" legislation, this is no guarantee of a successful rulemaking. To begin with, the online advertising industry will persist in arguing that profiling and tracking for advertising purposes cause little if any real consumer harm whereas new advertising restrictions (especially a default opt-out rule) will not only lower advertising revenues but

156. *Id.* at 373 nn.69–72.

157. *See* Art. 29 Data Protection Working Party, 00909/10/EN, WP 171, *Opinion 2/2010 on Online Behavioural Advertising* 15 (June 22, 2010), *available at* http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp171_en.pdf (arguing that default privacy-protective settings should require users "to go through a privacy wizard when they first install or update the browser").

imperil the subsidization of free online content and services, resulting in higher costs to consumers.[158] Moreover, advocates may worry that private factions will capture the conventional rulemaking process or that in implementing new legislation with unknown economic effects, the FTC will proceed very cautiously. In short, both sides may have something to gain from putting forward their best arguments in face-to-face negotiations, making reasonable concessions, and agreeing on a compromise.

As to information sharing, the negotiated rulemaking process by its very nature encourages more credible transmission of information among the parties. To begin with, the online advertising industry undoubtedly possesses greater expertise and insight into its own technology and evolving business models than either privacy advocates or FTC staff. In the past, this information has been shared or elicited mostly through one-sided communications—unilateral codes of conduct, complaints filed with the FTC, comments on FTC reports, or charges and countercharges at public forums. In a negotiated rulemaking process, however, the logic of Coasian bargaining prevails. In other words, each party seeks to "maximize its share of the gains produced by departure from standard requirements," and this requires that parties "educate each other, pool knowledge, and cooperate in problem solving."[159] In short, when both sides engage in explicit bargaining over priorities and tradeoffs, they are far more likely to achieve a satisfactory compromise than by relying on the indirect communications that characterize conventional rulemaking,[160] especially given their understanding that if negotiations fail, the FTC will proceed with its own rule.

3. Safe Harbor Programs

Finally, if Congress enacts into law either of the proposed bills that authorize safe harbor programs, the FTC should take a co-regulatory approach to rulemaking, i.e., one in which industry enjoys considerable flexibility in shaping self-regulatory guidelines in exchange for providing

158. *See, e.g.*, *'Do-Not-Track' Dissected: ClickZ Sends Feedback to FTC*, CLICKZ (Feb. 18, 2011), http://www.clickz.com/clickz/news/2027495/-track-dissected-clickz-sends-feedback-ftc.

159. *See* Jody Freeman & Laura I. Langbein, *Regulatory Negotiation and the Legitimacy Benefit*, 9 N.Y.U. ENVTL. L.J. 60, 69 (2000). For a very similar point, see Andrew P. Morriss et al., *Choosing How To Regulate*, 29 HARV. ENVTL. L. REV. 179, 201 (2005) (observing that "agencies may need the negotiation process to allow one set of interests to make credible commitments or disclosures to another set of interests that enable the regulation to be recognized as a Pareto improvement").

160. *See* DOC GREEN PAPER, *supra* note 62, at 5–6 (encouraging the development of codes of conduct using multi-stakeholder groups).

privacy protections that exceed default statutory requirements.[161] Section 5503 of the COPPA establishes an optional safe harbor that, in theory, would allow "flexibility in the development of self-regulatory guidelines" in a manner that "takes into account industry-specific concerns and technological developments."[162] In practice, the COPPA regulations are not very flexible, in part because the safe harbor approval process requires a side-by-side comparison of the substantive provisions of the COPPA rule with the corresponding sections of the self-regulatory guidelines. As a result, the four approved COPPA safe harbor programs are alike in reproducing the statutory requirements, and they show little differentiation by sector or technology. Nor do they benefit from face-to-face negotiations among the interested parties. The new privacy legislation provides a welcome opportunity to improve upon this first effort at implementing safe harbors.

For example, H.R. 611 specifically directs the FTC to implement safe harbor programs that allow for and promote "continued evolution and innovation in privacy protection, meaningful consumer control, simplified approaches to disclosure, and transparency" and provide "additional incentives" for participation in self-regulation.[163] One way for the Commission to accomplish this goal would be to permit the kind of experimentation described above. The Commission could then decide whether to allow an industry sector to comply with the notice requirements under Title I of the Act through some combination of "nutrition labels" for privacy, P3P user agents, and privacy search services. Or, even though subsections 403(1)(A) and (B) require that safe harbor programs provide consumers with a universal opt-out mechanism and various preference management tools, the Commission could decide whether firms satisfy these requirements (partially or in full) by adopting privacy-preserving targeted ad systems like Adnostic.

In addition, the Commission should treat safe harbor implementation as a perfect opportunity to experiment with negotiated rulemaking.[164] The Kerry-McCain bill should also be read as encouraging experimentation given that section 103 imposes a privacy by design requirement and section 501 requires the FTC to promulgate a rule establishing safe harbor programs that

161. *See* Rubinstein, *supra* note 138, at 414–20.

162. *See* Children's Online Privacy Protection Rule, 64 Fed. Reg. 59,888, 59,906 (Nov. 3, 1999) (codified at 16 C.F.R. pt. 312).

163. H.R. 611, 112th Cong. § 404(4), (5) (2010).

164. Under the NRA, agencies have discretion to determine whether to rely on negotiated rulemaking provided they determine that the use of this procedure serves the public interest based on consideration of the seven factors identified in 5 U.S.C. § 563(a) (2006).

implement the requirements of the Act with regard to certain uses of personal data, while subsection 701(1) requires the DOC to contribute to the development of commercial data privacy policy by "convening private sector stakeholders, including members of industry, civil society groups, academia, in open forums, to develop codes of conduct in support of applications for safe harbor programs."[165] This language amounts to an open invitation to appoint a negotiating committee to flesh out the meaning of privacy by design in the context of a safe harbor program.

V. CONCLUSION

The endorsement by privacy officials of PETs and privacy by design presents both exciting opportunities and serious challenges. While firms could improve their data practices by adopting appropriate PETs or building privacy into the design of new products and services, they are unlikely to seize the initiative as long as the economic incentives remain inadequate and the meaning of privacy by design or PET remains inexact. In the face of weak consumer demand, a lack of relevant data to engage in cost-benefit analyses, high opportunity costs for any voluntary restrictions on collecting and analyzing valuable personal data, and reputational sanctions that frequently are not compelling enough to drive new privacy investments, regulatory incentives are required.

In the coming years, Congress may or may not enact new privacy legislation. In the absence of new legislation, the FTC may continue to pursue strategic enforcement actions and, on its own or in conjunction with the DOC, convene experts from industry, advocacy groups, and academia to develop best practices for privacy by design. Alternatively, new legislation may authorize FTC-supervised experimentation with innovative regulatory approaches that relax one-size-fits-all requirements in exchange for better privacy results, negotiated solutions to emerging regulatory challenges such as how best to implement a "do not track" rule, and/or the use of safe harbor programs that permit flexible self-regulatory arrangements for implementing CIPPs subject to FTC oversight and enforcement. In short, regardless of which path Congress follows, a co-regulatory approach not only overcomes the false dichotomy of purely voluntary industry codes of conduct versus highly prescriptive government regulation, but it also helps encourage innovation and experimentation with privacy technology.

165. The Kerry-McCain bill, S. 799, 112th Cong., § 701(1) (2010).

APPENDIX

PRELIMINARY LISTING OF BEST PRACTICES IN PRIVACY DESIGN BASED ON FTC ENFORCEMENT CASES AND THE STAFF REPORT[166]

Prohibited practices. Companies shall not:

- Exploit any security vulnerability to download or install software.[167]
- Distribute software code bundled with "lureware" that tracks consumers' internet activity or collects other personal information, changes their preferred homepage or other browser settings, inserts new toolbars onto their browsers, installs dialer programs, inserts advertising hyperlinks into third-party web pages, or installs other advertising software.[168]
- Install content protection software that hides, cloaks, or misnames files, folders, or directories or misrepresents the purpose or effect of files, directory folders, formats, or registry entries.[169]

Required practices. Companies must:

- Clearly and conspicuously disclose when free software is bundled with harmful software (malware) creating security and privacy risks for consumers who install it.[170]

166. This preliminary listing of privacy design practices is premised entirely on a subset of FTC enforcement cases and the views stated in the Staff Report. It is therefore a work in progress and necessarily incomplete in its current form. For alternative lists of privacy best practices, see *Privacy and Security | Privacy Overview*, ACM US PUB. POLICY COUNCIL, http://usacm.acm.org/privsec/category.cfm?cat=7&Privacy%20and%20Security (describing 24 recommended practices for developing systems that utilize PII); Marilyn Prosch, *Protecting Personal Information Using Generally Accepted Privacy Principles (GAPP) and Continuous Control Monitoring To Enhance Corporate Governance*, 5 INT'L J. DISCLOSURE & GOVERNANCE 153 (2008) (describing the development of a privacy framework that resulted in the formulation of the GAPP, which consists in 10 privacy principles and 66 auditable criteria).

167. Stipulated Final Order for Permanent Injunction and Settlement of Claims for Monetary Relief, FTC v. Odysseus Mktg., Inc., Civil No. 05-CV-330 (D.N.H. Oct. 24, 2006), *available at* http://www.ftc.gov/os/caselist/0423205/061121odysseusstipfinal.pdf.

168. Stipulated Final Order for Permanent Injunction and Monetary Judgment as to Defendants Enternet Media, Inc., Conspy & Co., Inc., Lida Rohbani, Nima Hakimi, and Baback (Babak) Hakimi, FTC v. Enternet Media, Inc., Civil No. CV 05-7777 (C.D. Cal. Aug. 22, 2006), *available at* http://www.ftc.gov/os/caselist/0523135/060823enternetmediastlmnt.pdf.

169. Decision and Order, Sony BMG Music Entm't, Docket No. C-4195 (FTC June 28, 2007), *available at* http://www.ftc.gov/os/caselist/0623019/0623019do070629.pdf.

- Clearly and conspicuously disclose that the installation of software from a CD may limit a consumer's ability to copy or distribute audio files from the CD or other digital content; and, if such software causes information about consumers, their computes, or their use of a product to be transmitted via the Internet (so-called "phone home" features), then companies must disclose this prior to any such transmission and obtain the consumer's opt-in consent.[171]
- Provide a readily identifiable means for consumers to uninstall any adware or similar programs that monitor consumers' internet use and display frequent, targeted pop-up ads, where the companies deliberately made these adware programs difficult for consumers to identify, locate, and remove from their computers.[172]
- Clearly and prominently disclose the types of data that certain tracking software will monitor, record, or transmit prior to installing this software and separate from any user license agreement.[173]
- Provide prominent disclosures and obtain opt-in consent before using consumer data in a materially different manner than claimed when the data was collected, posted, or otherwise obtained.[174]

Recommended practices. Companies should do or adhere to the following:

- Develop and implement reasonable procedures concerning the collection and use of any personally identifiable information, including collecting information only to the extent necessary to fulfill a business purpose, disposing of the information no longer necessary to accomplish that purpose, and maintaining the privacy and security of information collected and stored.[175]
- Incorporate a formal privacy review process into the design phases of new initiatives.[176]

170. Decision and Order, Advertising.com, Docket No. C-4147 (FTC Sept. 12, 2005), *available at* http://www.ftc.gov/os/caselist/0423196/050916do0423196.pdf.

171. *Sony BMG*, Decision and Order, *supra* note 169.

172. Decision and Order, Zango, Inc., Docket No. C-4186 (FTC Mar. 9, 2007), *available at* http://www.ftc.gov/os/caselist/0523130/0523130c4186decisionorder.pdf.

173. Decision and Order, Sears Holdings Management Corp., Docket No. C-4264 (FTC Sept. 9, 2009), *available at* http://www.ftc.gov/os/caselist/0823099/090604searsdo.pdf.

174. *See* Decision and Order, Gateway Learning Corp., Docket No. C-4120 (FTC Sept. 10, 2004), *available at* http://www.ftc.gov/os/caselist/0423047/040917do0423047.pdf.

175. Letter from David C. Vladeck to Albert Gidari, *supra* note 66.

176. *Id.*

- Implement a "sliding scale" approach to access, taking into account the costs and benefits of access in different situations.[177]
- Provide "clear, comparable, and concise descriptions of a company's overall data practices" in privacy notices.[178]
- Seek affirmative express consent before collecting, using, or sharing any "sensitive information" including "information about children, financial and medical information, and precise geolocation data."[179]
- Where consumers elect not to have their information collected, used, or shared, "that decision should be durable and not subject to repeated additional requests from the particular merchant."[180]
- Where a company has a relationship with a consumer, it should offer a choice mechanism "at the point when the consumer is providing data or otherwise engaging with the company."[181]
- Where a company is engaged in online behavioral advertising, it should use a special choice mechanism consisting in "do not track."[182]
- Where a social media firm conveys consumer information to a third-party application developer, "the notice-and-choice mechanism should appear at the time the consumer is deciding whether to use the application and in any event, before the application obtains the consumer's information."[183]

177. FTC STAFF REPORT, *supra* note 4, at 72–73.
178. *Id.* at 71.
179. *Id.* at 61.
180. *Id.*
181. *Id.* at 58.
182. *Id.* at 63–69.
183. *Id.* at 59.

STRONG WILLS, WEAK LOCKS: CONSUMER EXPECTATIONS AND THE DMCA ANTICIRCUMVENTION REGIME

Krzysztof Bebenek†

TABLE OF CONTENTS

I. INTRODUCTION

Why do people sometimes choose to infringe copyrights? Why do they sometimes choose not to? This Note suggests they do so because they want to. That answer may sound flippant, but note the twist: it says nothing about

† J.D., 2011, University of California, Berkeley School of Law. Thank you to Kenneth Bamberger, Yan Fang, Daniel Farber, workshop participants at the Technology and Regulation Symposium, and the BTLJ editorial staff for their helpful comments and discussion.

law or technology. The anticircumvention regime of the Digital Millennium Copyright Act ("DMCA"),[1] which operates in the background of our everyday interactions with the technologies used to distribute copyrighted works, relies on a combination of digital rights management ("DRM") systems that restrict certain interactions and capabilities, and legal rules that impose liability for defeating these protections.[2] This regime is complex and, according to many commentators, overly burdensome, riding roughshod over traditional copyright law's careful balancing of rights between authors and the general public.[3] Such criticisms are apt, but I add to them another: because the DMCA's anticircumvention regime relies on a combination of complex law and porous technology that fails to reflect consumer expectations, there is good reason to believe that it is also fairly ineffectual.

Rather than actively complying with the DMCA's abstruse provisions or passively accepting the narrow range of interactions that DRM technologies typically allow, many copyright consumers seem to have a different lodestar—their own beliefs and intuitions about the kinds of interactions with copyrighted works that are desirable, appropriate, or natural. Following these intuitions, users do with works as they see fit. They may copy for personal use, to remix and criticize, to share with others, or to avoid paying a price. Should a DRM barrier stand in their way, they may very well circumvent it or they may not, but neither law nor technology seems to bear heavily on the choice. I do not offer this as a reductionist account of consumer behavior. I do suggest, though, that such intuitions or norms may play a greater role than many suspect in governing copyright consumers' behavior; that they may undermine the efficacy of both legal and technological restraints; and that market participants and lawmakers alike would do well to take them seriously.

The argument proceeds in three Parts. Part II surveys the DMCA regime under § 1201 and then offers a sketch of consumer norms as an alternative paradigm. Part III suggests why such norms may be the primary regulators governing the use (and abuse) of copyrighted works. It draws on both theoretical intuitions and insights from the Copyright Office triennial rulemaking on DMCA exemptions. Part IV explores two key implications of

1. Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2860 (1998) (codified as amended in scattered sections of 17 U.S.C.).

2. *See infra* Section II.A.

3. *See generally* LAWRENCE LESSIG, CODE VERSION 2.0, at 180–90 (2006); Timothy K. Armstrong, *Digital Rights Management and the Process of Fair Use*, 20 HARV. J.L. & TECH. 49, 67–68 (2006); Dan L. Burk & Julie E. Cohen, *Fair Use Infrastructure for Rights Management Systems*, 15 HARV. J.L. & TECH. 41, 43–54 (2001).

this view: first, norms force us to rethink our understanding of technology as a tool of governance; second, they hold information that may be vital to improving our copyright policy.

II. WHAT REGULATES?

Perhaps it is foolish to ask what *really* governs people's interactions with copyrighted works and with the technological means used to embed and distribute them. As Cass Sunstein has aptly noted, "what lies behind choices is not a thing but an unruly amalgam of . . . aspirations, tastes, physical states, responses to existing roles and norms, values, judgments, emotions, drives, beliefs, whims."[4] Still, it is worth asking whether the relevant legal regime in its current form at least plays a significant role in that amalgam—and, if it does not, what takes its place. This Part lays the groundwork: it surveys the anticircumvention provisions of the DMCA, which are intended to govern user behavior, and offers a technology-oriented concept of consumer norms as an alternative candidate.

A. THE DMCA DEFAULT

Section 1201 of the Copyright Act codifies the anticircumvention provisions of the DMCA.[5] It is a complex statutory scheme but is based on a simple principle: the technological measures that copyright holders use to protect digital embodiments of their works will define the contours of liability.

The statute contemplates two different types of technological protections: access controls and copy controls. Section 1201(a), which governs the circumvention of access controls, defines these protections as technology that, "in the ordinary course of its operation, requires the application of information, or a process or a treatment, with the authority of the copyright owner, to gain access to [a] work."[6] Section 1201(b), meanwhile, governs circumvention of copy controls. It defines such measures as technology that, "in the ordinary course of its operation, prevents, restricts, or otherwise limits the exercise of a right of a copyright owner under [the Copyright Act]."[7]

4. Cass R. Sunstein, *Social Norms and Social Roles*, 96 COLUM. L. REV. 903, 913 (1996).
5. 17 U.S.C. § 1201 (2006).
6. § 1201(a)(3)(B).
7. § 1201(b)(2)(B).

This distinction is central to the structure of § 1201's prohibitions.[8] Section 1201(a) provides for two different sources of liability related to the circumvention of access controls. First, it contains a direct prohibition, stating simply that "[n]o person shall circumvent" an access control.[9] Second, it forbids the "manufacture, import, offer to the public, provi[sion], or [other] traffic[king] in any technology, product, service, device, component, or part thereof that . . . is primarily designed or produced for the purpose of circumventing" access controls.[10] Section 1201(b), however, does not contain a direct prohibition on the circumvention of copy controls. Rather, it contains only a provision nearly identical to § 1201(a)'s prohibition on trafficking in circumvention tools or services, but targeted at tools that allow circumvention of copy controls.[11] To the extent one can tell apart the two types of technologies that control access and copying,[12] § 1201 privileges access controls by conferring on them a greater degree of protection: the defendant who defeats an access control to facilitate his infringement will face liability for violating both the traditional exclusive rights conferred on authors by § 106 of the Copyright Act[13] and the prohibition on circumvention of § 1201(a); the defendant who defeats a copy control in order to infringe, however, will be liable only for traditional copyright

8. Though it is a key feature of § 1201, some commentators have noted that the distinction between the two categories is unclear at best, and perhaps meaningless. As Aaron Perzanowski suggests, while "one can at least conceive of a protection measure that prevents copying without limiting access to the underlying copyrighted work, . . . such a measure may be difficult or impossible to engineer." Aaron K. Perzanowski, *Evolving Standards and the Future of the DMCA Rulemaking*, 10 J. INTERNET L., Apr. 2007, at 1, 12. Courts, too, have been somewhat inconsistent in making this distinction. *See, e.g.*, Universal City Studios, Inc. v. Corley, 273 F.3d 429, 435, 438 n.5 (2d Cir. 2001) (affirming the lower court's finding of a violation stemming from trafficking in tools for the circumvention of copy controls in spite of suggesting that the technological measure in question controlled only access); 321 Studios v. Metro Goldwyn Mayer Studios, Inc., 307 F. Supp. 2d 1085, 1097 (N.D. Cal. 2004) (dissolving the conceptual difference between access and copy controls by explaining that "the purpose of [an] access control [on a DVD] is to control copying").

9. § 1201(a)(1)(A).

10. § 1201(a)(2).

11. § 1201(b)(1).

12. For more on the difficulties of making this distinction, see *supra* note 8.

13. *See* 17 U.S.C. § 106 (defining an author's exclusive rights in a copyrighted work).

infringement.[14] Violations of § 1201 are subject to civil and, potentially, criminal penalties.[15]

Section 1201 layers two types of exemptions atop this scheme of liability. First, the statute designates seven specific categories of activity that it exempts from at least some circumvention liability. There are exemptions for uses by nonprofit libraries, archives, and educational institutions; law enforcement uses; reverse engineering; encryption research; the protection of minors; privacy protection; and security testing.[16] Some, like the law enforcement exemption, pertain to all prohibitions on circumvention and trafficking.[17] Others apply only to specific sources of liability: the nonprofit institution exemption, for instance, applies only to individual circumvention of access controls and does not affect liability for violating either of the trafficking prohibitions.[18]

Second, the statute grants the Librarian of Congress authority to engage in a triennial rulemaking to determine, upon the recommendation of the Register of Copyrights, whether "users of a copyrighted work are, or are likely to be in the succeeding 3-year period, adversely affected . . . in their ability to make noninfringing uses . . . of a particular class of copyrighted works."[19] Circumvention related to any class of works that the Librarian identifies will be exempt from liability for a three-year period.[20] Crucially, however, the references to adverse effects, as well as any resulting exemption, apply only to the prohibition on direct circumvention of access controls, and not to the prohibitions on trafficking.[21]

The Librarian of Congress has engaged in this rulemaking process four times to date.[22] The exemptions resulting from the initial two rounds were,

14. For more on the distinction between the different degrees of protection that § 1201 confers on access and copy controls, see R. Anthony Reese, *Will Merging Access Controls and Rights Controls Undermine the Structure of Anticircumvention Law?*, 18 BERKELEY TECH. L.J. 612, 622–24 (2003).

15. 17 U.S.C. §§ 1203–1204. Civil penalties under § 1203 are available for all violations, while criminal penalties under § 1204 apply only to violations done "willfully and for purposes of commercial advantage or private financial gain." *Id.* § 1204(a).

16. *Id.* § 1201(d)–(j).

17. § 1201(e).

18. § 1201(d)(1), (d)(4).

19. § 1201(a)(1)(C).

20. § 1201(a)(1)(D).

21. *See* § 1201(a)(1)(A), (B) (providing that "the prohibition in subparagraph (A)"—that is, the prohibition on "circumvent[ion of] a technological measure that effectively controls access to a work protected under this title"—"shall not apply" to works deemed to have an adverse effect on noninfringing uses).

22. James H. Billington, *Statement of the Librarian of Congress Relating to Section 1201 Rulemaking*, U.S. COPYRIGHT OFFICE (July 26, 2010), *available at* http://www.copyright.gov/

by and large, narrow and esoteric.[23] The last two rounds of rulemaking, however, give reason to suspect that the exemption provision might have more teeth.[24] In 2006, the Librarian created an exemption allowing film and media studies professors to circumvent access controls on DVDs in order to use clips from protected works in the classroom.[25] In 2010, the Librarian broadened that exemption to include use by other kinds of faculty, film and media students, documentary filmmakers, and, perhaps most significantly, the creators of "noncommercial videos."[26] In the same rulemaking, the Librarian also carved out a new exemption allowing smartphone users to circumvent access controls that prevent their devices from running third-party applications unapproved by their service providers, a practice known as "jailbreaking."[27]

One concept is conspicuously absent from the scheme of § 1201: on its face, § 1201 does not explicitly condition liability on any predicate finding of copyright infringement.[28] This absence has occasioned a fair amount of inconsistency.[29] The Federal Circuit, for instance, has gone some way toward reading in an infringement requirement, holding that plaintiffs must show a nexus between circumvention and infringement in order to establish liability.[30] Not all courts have adopted this reading, though. Some have held that defenses such as fair use, which if successful refutes a finding of

1201/2010/Librarian-of-Congress-1201-Statement.html (explaining that "[t]his is the fourth time" the Librarian of Congress has made the determination required by § 1201).

23. *See, e.g.*, Exemption to Prohibition on Circumvention of Copyright Protection Systems for Access Control Technologies (*Section 1201 Rulemaking*), 65 Fed. Reg. 64,556, 64,562 (Oct. 27, 2000) (creating exemption for "[l]iterary works, including computer programs and databases, protected by access control mechanisms that fail to permit access because of malfunction, damage or obsoleteness"); *Section 1201 Rulemaking*, 68 Fed. Reg. 62,011, 62,013 (Oct. 31, 2003) (creating exemption for "[c]omputer programs protected by dongles that prevent access due to malfunction or damage and which are obsolete").

24. *See, e.g.*, Perzanowski, *supra* note 8, at 20–25 (cataloguing both the improvements of the 2006 rulemaking over the prior two rounds in its ability to mitigate the adverse effects of § 1201 on noninfringing uses, as well as its continued shortcomings).

25. *Section 1201 Rulemaking*, Final Rule, 71 Fed. Reg. 68,472, 68,480 (Nov. 27, 2006).

26. *Section 1201 Rulemaking*, Final Rule, 75 Fed. Reg. 43,825, 43,839 (July 27, 2010).

27. *Id.*

28. *See* 17 U.S.C. § 1201 (2006) (imposing prohibitions using the blanket expression "no person shall . . . ," and making no explicit mention of a predicate infringement requirement).

29. For a thorough review of judicial approaches to this aspect of § 1201, see Timothy K. Armstrong, *Fair Circumvention*, 74 BROOK. L. REV. 1, 5–27 (2008) (summarizing two divergent strands of statutory construction).

30. *See* Chamberlain Grp. v. Skylink Techs., Inc., 381 F.3d 1178, 1203 (Fed. Cir. 2004) (requiring that circumvention be done "in a manner that . . . infringes or facilitates infringing a right *protected* by the Copyright Act").

infringement,[31] do not apply to violations of § 1201.[32] Most notably, the Court of Appeals for the Ninth Circuit last year explicitly rejected the Federal Circuit's approach, finding its nexus requirement unpersuasive and "contrary to the plain language of the statute."[33]

B. THE NORM-BASED ALTERNATIVE

That is, briefly, the law on circumvention. But perhaps this complex legal regime does not order people's conduct—perhaps their actions owe something to a more vague and gauzy set of strictures. I will refer to this regulator as "norms," though this use of the term is idiosyncratic. By "norms" I mean the set of expectations and intuitions that govern people's relationships with technology. This idea draws on the concept of mental models developed by cognitive scientists and theorists of human-computer interaction.[34] The mental models theory holds, broadly, that people create simplified mental "maps" of devices that guide their interactions with these objects.[35] It draws, too, on the work of Langdon Winner, who famously suggested that technical artifacts have a political dimension that manages to conceal itself below their humdrum surface.[36] Such objects, Winner argued, are almost inevitably "designed and built in such a way that [they] produce[] a set of consequences logically and temporally *prior to any of [their] professed uses*."[37] As a result, "seemingly innocuous design features . . . actually mask social choices of profound significance."[38] Winner's contention that this political dimension remains hidden from view assumes that users have strong preconceived notions of what is *ordinary* for a given technology: they expect their CDs, DVDs, iPods, cell phones, and websites to have certain

31. 17 U.S.C. § 107 ("[T]he fair use of a copyrighted work . . . is not an infringement of copyright.").

32. *See, e.g.*, Universal City Studios, Inc. v. Corley, 273 F.3d 429, 443 (2d Cir. 2001) ("[Appellants] contend that subsection 1201(c)(1), which provides that '[n]othing in this section shall affect rights, remedies, limitations or defenses to copyright infringement, including fair use, under this title,' can be read to allow the circumvention of encryption technology protecting copyrighted material when the material will be put to 'fair uses' exempt from copyright liability. We disagree that subsection 1201(c)(1) permits such a reading.").

33. MDY Indus., L.L.C. v. Blizzard Entm't, 629 F.3d 928, 950 (9th Cir. 2010).

34. *See generally* Stephen J. Payne, *Users' Mental Models: The Very Ideas*, *in* HCI MODELS, THEORIES, AND FRAMEWORKS: TOWARD A MULTIDISCIPLINARY SCIENCE 135, 135–56 (John M. Carroll ed., 2003) (comparing and contrasting various strands of the mental models theory of human-computer interaction).

35. *Id.* at 142.

36. Langdon Winner, *Do Artifacts Have Politics?*, 109 DAEDALUS, no. 1, 1980, at 121.

37. *Id.* at 125.

38. *Id.* at 127.

capabilities and to lack others; if these technologies do not defy expectations, users may overlook their subtle politics.

This is not the classic account of norms. Many writers on the relationship between law and norms conceive of norms as informal obligations followed in the pursuit of interpersonal esteem, or for the avoidance of social sanction.[39] Others, focusing more closely on copyright and the problems of peer-to-peer ("P2P") file sharing, have traced the erosion of the specific norm of compliance with the law.[40] Technology-based norms are not interpersonal, but they nonetheless share two important features with social norms: first, flowing as they do from expectations and intuitions, they are by nature informal; second, because they pertain to mass-market devices and media formats, they are shared by large groups of people and hence are common enough to govern conduct. Compliance, meanwhile, is a rather reductive framework. While it may be apposite in the P2P context, where the relevant legal rule is a fairly simple prohibition on outright copying,[41] as argued below, it is far less illuminating with respect to a statute like the DMCA, whose complexity calls the very notion of strict compliance into question.

Such technological norms might just as easily go by the name of "consumer expectations" or, when put into practice, "consumer behavior." I use all of these terms interchangeably. I use the term "norms," though, because I believe both my use of that term and its more traditional usage have a common root in the concept of social meaning. Social meaning refers to the expressive content of everyday behaviors, which is derived from the particular role these behaviors play in people's relations with one another.[42] Because the vast majority of copyrighted works are expressive, our interactions with such works—be they through authorship, consumption, commentary, remixing, sampling, appropriation, or theft—are surely

39. *See, e.g.*, LESSIG, *supra* note 3, at 120–37; Richard H. McAdams, *The Origin, Development, and Regulation of Norms*, 96 MICH. L. REV. 338 (1997); Sunstein, *supra* note 4, at 914–21.

40. *See, e.g.*, Ben Depoorter & Sven Vanneste, *Norms and Enforcement: The Case Against Copyright Litigation*, 84 OR. L. REV. 1127, 1139–43 (2005); Daniel J. Gervais, *The Price of Social Norms: Towards a Liability Regime for File-Sharing*, 12 J. INTELL. PROP. L. 39, 49–51 (2004).

41. Even in this context, though, more sophisticated models are *possible*. Lior Strahilevitz, for instance, has merged the interpersonal and compliance frameworks by arguing that norms of cooperation and reciprocity on P2P networks create a kind of "charismatic code" that helps to mask the illegality of file sharing from participants. Lior Jacob Strahilevitz, *Charismatic Code, Social Norms, and the Emergence of Cooperation on the File-Swapping Networks*, 89 VA. L. REV. 505, 547–75 (2003).

42. *See* Lawrence Lessig, *The Regulation of Social Meaning*, 62 U. CHI. L. REV. 943 (1995); Sunstein, *supra* note 4, at 925–28.

freighted with social meanings: the ways we produce, consume, and otherwise use them are all embedded in particular social contexts and serve specific social purposes. It is therefore no stretch to suggest that this social meaning extends to the rather more mundane and solitary technical work that is instrumental to those interactions. Buying a record, downloading an MP3 file, photocopying artwork from a book, ripping a film from an access-protected DVD—we do these things, generally, not for their own intrinsic value, but to serve some further purpose: as with the choice to follow social norms, here too we have our reasons.

III. THE CASE FOR NORMS

There is cause to believe that today, those reasons might at the very least be as important as the law in governing people's actions. This Part makes the case for viewing circumvention and anticircumvention through the prism of norms on both theoretical and empirical grounds.

A. THEORETICAL ARGUMENTS

1. The Weakness of Law and Architecture

Before considering whether the DMCA works, we would do well to ask how it is supposed to work. In a sense, it is a law like many another: it demands compliance and imposes sanctions for refusal. But, it would be rather implausible to claim that ordinary users actually *comply* with the DMCA. That would presuppose knowledge of what is, after all, a very complex statute. A user contemplating circumvention of the Content Scramble System ("CSS") protecting a DVD who wished to make a noninfringing use and comply to the letter with the DMCA would first have to understand the statute's differential treatment of access and copy controls and would have to determine which set of provisions applied to CSS. He would then have to know that the statute makes a further distinction between circumvention on one's own and trafficking in circumvention tools or services. Finally, he would have to be aware of the seven statutory exemptions and the additional exemptions promulgated by rulemaking, and would have to determine whether any applied to his conduct. We may lack the kind of perfect access to the minds of consumers that would allow a definitive rejection of the compliance paradigm. Nonetheless, to argue that ordinary consumers consciously comply with or flout the law when they choose whether or not to circumvent seems intuitively incorrect.

If the statute's complexity renders compliance with the law difficult for the ordinary copyright consumer, it would seem that the technological protections imposed on copyrighted works do the lion's share of the work in

the DMCA regime. But there is also reason to doubt this possibility. Were DRM technology effective on its own at preventing piracy, § 1201's prohibitions would be largely superfluous. In fact, this technology is imperfect and porous. Software is hackable,[43] and DRM has proven particularly vulnerable on this front.[44] Technological protection measures, it seems, protect very little.

But perhaps directly blocking copying and other such interactions is not the only way that DRM operates. Imagine in place of DRM a chain-link fence of average height surrounding a piece of land.[45] The fence is by no means insurmountable, though to be sure, climbing it takes some amount of effort for the would-be trespasser. Its ultimate significance, though, is symbolic: it serves as a clear marker of the line beyond which the law of trespass takes effect. The modest amount of extra effort required to scale the fence therefore deters not in and of itself, but because it serves as a reminder that the trespasser's act flouts both the will of the property owner and the command of the sovereign. The law communicates, and the fence, along with the kinds of physical interactions it requires, is one of the means by which it communicates.

Arguably, DRM systems function in much the same way as the fence. Few such digital fences are inherently insurmountable. Granted, these measures do increase the cost of the interactions they seek to block. At first, then, the cost of circumvention may be high to most. But to those with the right technological skills, scaling the digital fence is only a matter of time. And it is an axiom of information economics that once someone has made it over, presuming a willingness to share information, others will be able to replicate his results at no additional cost: the marginal cost of producing information is generally near zero.[46] Ultimately, then, DRM increases the cost

43. James Grimmelmann, *Regulation by Software*, 114 YALE L.J. 1719, 1742–43 (2005) (describing vulnerability to hacking as an inherent feature of software).

44. *Id.* at 1755–57 (cataloguing DRM vulnerabilities and concluding that "[i]t is not clear that any DRM system has withstood a serious attempt to crack it").

45. The analogy between tangible and intellectual property rights is in many ways imperfect. For this reason, Dan Burk and Julie Cohen reject comparing DRM systems to fences. *See* Burk & Cohen, *supra* note 3, at 52–54. While I largely agree with their arguments, this analogy remains instructive with respect to the ways people experience and interact with these two very different technologies.

46. *See, e.g.*, Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, *in* THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS 609, 615 (Nat'l Bureau of Econ. Research ed., 1962) ("In the absence of special legal protection, the owner cannot . . . simply sell information on the open market. Any one purchaser can destroy the monopoly, since he can reproduce the information at little or no cost.").

of the interactions it proscribes only modestly. In doing so, however, it has the power to put the would-be circumventor on notice—to bring about a moment of contemplation when she might reflect on her actions. Whether she creates a circumvention tool from scratch or merely downloads one, our circumventor is likely to take note that her interaction with the copyrighted work has moved from an easy, unimpeded channel to a more clandestine kind of space. She must, at some level, confront the significance of this fact.

But what significance does this fact hold? For the pirate who seeks to profit by reproducing the work of others on a mass scale, probably very little. But the remixer who wants to repurpose a brief clip from a Hollywood blockbuster may face a more difficult choice. A dogged few may attempt to wade their way to truth though § 1201. Some could assume DRM is there for good reason and dutifully stand down. Others, however, driven by some private sense that they are taking only a short clip, or that the remix is a cultural staple and could not possibly be illegal, or that their conduct harms the copyright holder negligibly, if at all, would likely forge ahead. As Lessig and other commentators have observed, using technological measures to implement legal protections is a way of "privatizing" legal decision-making.[47] They had in mind the way such technological measures let copyright *owners*, rather than lawmakers, regulate how their works are used.[48] There is just as much reason to believe, however, that the combination of a highly complicated legal regime with imperfect, porous technology actually puts the key decisions in the hands of the *users* whom it is meant to govern. That result, of course, can hardly be called "governance."

2. DRM Versus the Norms of Ownership

As the discussion above suggests, people tend to have different ideas about the appropriateness of circumvention and of the underlying uses of copyrighted works that circumvention may help further. But DRM does more than just leave consumers to rely on those ideas; it pushes them to confront the question whether or not to circumvent. At the same time, there is reason to believe that it also destabilizes some of the very ideas one might rely on to help arrive at an answer. This is because DRM brings into conflict with one another two strong intuitions about ownership: the notion of property as the owner's sole dominion and the prohibition on trespass.

47. *See, e.g.*, LESSIG, *supra* note 3, at 179 ("Trusted systems . . . are a privatized alternative to copyright law."); Daniel Benoliel, *Technological Standards, Inc.: Rethinking Cyberspace Regulatory Epistemology*, 92 CALIF. L. REV. 1069, 1114 (2004) (arguing that through the use of DRM, "decentralized content providers are . . . privatizing [copyright] enforcement authority").

48. *See, e.g.*, LESSIG, *supra* note 3, at 179.

William Blackstone famously described property as "that sole and despotic dominion which one man claims and exercises over the external things of the world, in total exclusion of the right of any other individual in the universe."[49] At least as far as chattels are concerned, the law does largely give the Blackstonian notion effect. Courts have traditionally been highly skeptical of servitudes or other encumbrances that run with personal property in a way that they have not been with similar restrictions on an owner's use of land, and contemporary authorities on the subject agree that such restrictions would likely be impossible to impose today.[50]

This general reluctance is heightened all the more in the intellectual property context, where well-established doctrines both stress the separateness of the owner's intangible entitlement from the physical object that embodies it and limit the entitlement holder's ability to exert control over the physical object beyond an initial sale. The first sale doctrine in copyright, for instance, allows the owner of a particular copy of a copyrighted work to sell or display that copy,[51] even though such uses are otherwise among the exclusive rights of the copyright holder.[52] Moreover, the first sale doctrine is not merely a statutory provision but has a far richer and more far-reaching common law pedigree.[53] As such, it should be properly understood not as an "idiosyncratic limit on the distribution right," but as a "principle hold[ing] that a fundamental set of user rights or privileges flows from lawful ownership of a copy of a work."[54] Patent law's exhaustion doctrine embodies a similar principle.[55]

All of this, of course, is law. But it is by and large settled, uniform, and simple law. Unlike the case of the DMCA, here, there is far more reason to believe that we have internalized these principles. In the popular

49. 2 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 2 (7th ed. 1775).

50. *See* Molly Shaffer Van Houweling, *The New Servitudes*, 96 GEO. L.J. 885, 906 (2008) (describing the "origins and development of [a] special hostility to chattel servitudes" and noting "[t]he conventional wisdom, as described by contemporary commentators, . . . that personal property servitudes are seldom enforceable").

51. *See* 17 U.S.C. § 109 (2006).

52. *See id.* § 106.

53. Aaron Perzanowski & Jason Schultz, *Digital Exhaustion*, 58 UCLA L. REV. 889, 912–25 (2011).

54. *Id.* at 912.

55. As the Supreme Court recently reaffirmed, under "[t]he longstanding doctrine of patent exhaustion . . . the initial authorized sale of a patented item terminates all patent rights to that item." Quanta Computer, Inc. v. LG Elecs., Inc., 553 U.S. 617, 625 (2008). Exhaustion is a common law doctrine not codified in the Patent Act. Its roots lie in nineteenth-century cases, many of which involved post-sale restrictions on patented items with an anticompetitive flavor. For a brief history of the doctrine, see *id.* at 625–28.

understanding that derives from this stable and fairly unequivocal set of laws, then, to own an object means to do with it as one pleases. However, DRM has brought this clear principle into tension with another venerable pillar of property law: the prohibition on trespass. This prohibition is not necessarily absolute,[56] but it is definitive. The Supreme Court has called the right to exclude "one of the most essential sticks in the bundle of rights that are commonly characterized as property."[57] Accordingly, this interest enjoys broad protection. Tort law allows a cause of action for trespass both to chattels and on land (and, unusually for tort, requires no showing of harm in the latter case);[58] furthermore, it favors injunctive relief on a finding of liability.[59] Criminal law imposes liability for similar conduct.[60] In short, the law takes special care to protect boundary lines, and, as with the Blackstonian conception, its prohibitions in this sphere are almost universally understood. That trespass, lock picking, and breaking and entering are forbidden is beyond common knowledge—it is second nature.

At first glance these two principles do not seem out of step with each other. The law's safeguards against trespass seem perfectly aligned with the Blackstonian notion of "sole and despotic dominion."[61] But DRM upends this balance. It limits the natural functionality of tangible objects that we own,[62] placing certain capabilities behind digital locks. So the professor who wishes to supplement her lecture with compiled clips of films culled from lawfully purchased DVDs must first defeat the CSS protections disabling copy functionality on each disc. And the iPhone owner who wants to install an application that Apple has deemed inappropriate faces a similar lock. Here, and in many analogous situations, such locks place before the user a

56. EDUARDO MOISÉS PEÑALVER & SONIA K. KATYAL, PROPERTY OUTLAWS: HOW SQUATTERS, PIRATES, AND PROTESTERS IMPROVE THE LAW OF OWNERSHIP 152–56 (2010) (discussing doctrine of necessity as an exception to liability for trespass).

57. Kaiser Aetna v. United States, 444 U.S. 164, 176 (1979).

58. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 158 (1977) (imposing liability for intentional trespass on land); *id.* § 163 (imposing liability for intentional trespass on land in the absence of any resultant harm); *id.* § 218 (imposing liability for trespass to chattels).

59. *See id.* § 937 cmt. a (explaining that while tort law now deems equitable remedies appropriate in other contexts as well, equity has historically "intervened . . . to protect property interests").

60. *See, e.g.*, MODEL PENAL CODE §§ 221.1–221.2 (1981) (defining liability for burglary and criminal trespass).

61. *See* BLACKSTONE, *supra* note 49, at 2.

62. To be sure, the question of copy *ownership* becomes complicated when works are distributed directly over digital networks, without a clear tangible embodiment or chattel that one might be said to own in the traditional sense. *See* Brian W. Carver, *Why License Agreements Do Not Control Copy Ownership: First Sales and Essential Copies*, 25 BERKELEY TECH. L.J. 1888 (2010).

dilemma that traditional intuitions about ownership cannot easily resolve. To abandon one's desired uses is to relinquish one's sole dominion over the object in question, while carrying on as planned would require a kind of trespass.

To anyone who has internalized traditional property expectations, neither option can be very satisfying. Writing about the closely related scenario where restrictive terms in shrinkwrap licenses clash with consumers' expectations about ownership, Mark Lemley describes the result as cognitive dissonance.[63] If anything, such dissonance is even stronger here. After all, Lemley refers to a conflict between user expectations and legal restrictions, whose force users might never know unless they find themselves hauled into court. Here, however, the restrictions act directly on the user—the conflict is unavoidable, the signals DRM sends decidedly mixed.

One might object to this reasoning. The first sale doctrine notwithstanding, has copyright law itself not always posed sharp limitations on the range of uses to which consumers could put tangible embodiments of works? Might we not expect consumers to have internalized such restrictions? While such limitations do exist, they have historically intersected very little with the kinds of uses that consumers might typically have wanted to make of their copies of copyrighted works. Thus, there is good reason to believe that they have had a limited impact on consumer expectations.

In the analog world, infringement of the sort that would deprive copyright holders of profits had to take place on an industrial scale: it called for printing presses, vinyl presses, mass duplication facilities—in other words, large capital investments.[64] Copyright law traditionally reflected this notion, treating smaller-scale, private uses with benign neglect, or specifically exempting them from liability.[65] Julie Cohen has called this schema copyright's public-private distinction. This distinction let consumers read, listen, and watch—and, along the way, develop their expectations—largely outside of copyright law's reach. Coupled with the (silent) limitations of analog consumer technology, it had little scope to upset the belief that their books, records, or cassettes were theirs to do with as they pleased.[66] Because

63. Mark A. Lemley, *The Law and Economics of Internet Norms*, 73 CHI.-KENT L. REV. 1257, 1274–75 (1998).

64. Jessica Litman, *Real Copyright Reform*, 96 IOWA L. REV. 1, 12 (2010) ("Until recently, mass distribution of copies of works of authorship required large capital investment.").

65. *See, e.g.*, 17 U.S.C. § 101 (2006) (excluding "a normal circle of family and its social acquaintances" from the definition of "publicly" in order to limit the scope of the public performance and display rights of § 106).

66. Julie E. Cohen, *Comment: Copyright's Public-Private Distinction*, 55 CASE W. RES. L. REV. 963, 964 (2005); *see also* Jessica Litman, *The Exclusive Right To Read*, 13 CARDOZO ARTS

DRM gives copyright owners control over many such private uses, its encroachment on those expectations is far more drastic. Our theoretical intuitions suggest, then, not just that norms are the only potential line of defense in the current DMCA regime, but that they are a rather unstable one at that.

B. CASE STUDY: THE DMCA RULEMAKING

Members of the motion picture, recording, and publishing industries, of related standard-setting bodies, and of other copyright-industry trade groups appear to recognize both the role that norms play in determining consumers' propensity to circumvent and the current instability of those norms. This seems to be an underlying theme, in any case, in these entities' submissions to and testimony in the triennial Copyright Office rulemaking on exemptions to circumvention liability. As might be expected, copyright industry participants in these proceedings often take the position that the Librarian of Congress should grant *no* exemptions to circumvention liability under § 1201(a)(1). Of greater interest is the narrative that emerges as these participants make their case. It is a story with a simple point: granting exemptions will confuse consumers about the legality of circumvention and erode the DMCA's protections; a blanket prohibition is the only appropriate solution.

The claim that granting exemptions will lead to confusion about the appropriateness of circumvention is an oft-repeated refrain. For instance, a representative of the Motion Picture Association of America ("MPAA") argued in a 2006 hearing that "[o]nce you start creating exceptions to the Prohibition Against [Circumvention], . . . [y]ou create confusion."[67] And in a 2009 comment, the MPAA expressed similar concern about an exemption sowing "widespread confusion as to what circumventions are and are not allowed and whether hacking tools are legitimate."[68] Likewise, the DVD Copy Control Association ("DVD CCA") has repeatedly argued that "once a

& ENT. L.J. 29, 34 (1994) (explaining that before the age of easy copying, "ordinary consumers could go about their business without ever encountering a copyright problem").

67. Transcript of Testimony at 76, *Section 1201 Rulemaking*, Docket No. RM 2005-11A (U.S. Copyright Office Apr. 3, 2006), *available at* http://www.copyright.gov/1201/2006/hearings/transcript-april03.pdf (testimony of Fritz Attaway, Motion Picture Ass'n of Am.) [hereinafter Attaway Testimony, Apr. 3, 2006].

68. Comments of Motion Picture Ass'n of Am. (MPAA) at 11, *Section 1201 Rulemaking*, Docket No. RM 2008-8 (U.S. Copyright Office Feb. 2, 2009), *available at* http://www.copyright.gov/1201/2008/responses/mpaa-46.pdf.

hacker is given an exemption, even for a limited purpose, it would become impossible to control or predict future hacks."[69]

Confusion begets erosion, such that any exemption could prove a slippery slope to the end of the DMCA regime, if not the entertainment industry. So, argued the DVD CCA, the supposed impossibility of distinguishing what is and is not lawful means that "even . . . 'limited exemptions' will essentially render CSS ineffective as a means of protecting copyrighted content."[70] Likewise, Time Warner has argued that "[g]ranting [an exemption] with respect to any particular [DRM] technology would be tantamount to outlawing the use of that technology."[71] A coalition of major copyright holders and industry members has sounded similar notes.[72] In 2006, for instance, the coalition warned college professors who sought permission to create digital compilations of film clips from DVDs for classroom use that "the very medium that educators have grown to appreciate and desire to use in their classrooms would be threatened by the recognition of the requested exemption."[73]

Instead of confusing exemptions, the copyright industries favor clarity and simplicity. In the words of a software industry representative, exemptions would interfere with the DMCA's "clear prohibition" on

69. Comments of DVD Copy Control Ass'n, Inc. at 2, *Section 1201 Rulemaking*, Docket No. RM 2002-3 (U.S. Copyright Office Feb. 19, 2003), *available at* http://www.copyright.gov/1201/2003/reply/028.pdf; *see also* Comments of the DVD Copy Control Ass'n, Inc. at 14, *Section 1201 Rulemaking*, Docket No. RM 2005-11 (U.S. Copyright Office Feb. 2006), *available at* http://www.copyright.gov/1201/2006/reply/01taylor_DVDCCA.pdf (making a nearly identical argument).

70. Comments of DVD Copy Control Ass'n, *supra* note 69, at 2.

71. Comments of Time Warner, Inc. at 10–11, *Section 1201 Rulemaking*, Docket No. RM 2005-11 (U.S. Copyright Office Feb. 2006), *available at* http://www.copyright.gov/1201/2006/reply/18aistars_TWI.pdf.

72. This coalition brought together members of major industrial-scale copyright owners and related trade associations in order to represent their joint interests in the rulemaking process. Members included the Association of American Publishers, the Association of American University Presses, the American Society of Media Photographers, the Authors Guild, the Business Software Alliance, the Directors Guild of America, the Entertainment Software Association, the Independent Film & Television Alliance, the Motion Picture Association of America, the National Music Publishers' Association, the Professional Photographers of America, the Recording Industry Association of America, the Screen Actors Guild, and the Software and Information Industry Association. Joint Reply Comments of Ass'n of Am. Publishers et al. at 1, *Section 1201 Rulemaking*, Docket No. RM 2005-11 (U.S. Copyright Office Feb. 2, 2006), *available at* http://www.copyright.gov/1201/2006/reply/11metalitz_AAP.pdf.

73. *Id.* at 30.

circumvention.[74] Similarly, an MPAA representative testified at a 2006 hearing that his organization sought to "maintain the simple proposition that it is illegal to circumvent."[75] In 2009, that same representative reiterated his client's concern about "erosion of the principle that circumventing a technological measure is against the law."[76]

Most telling, however, is the explanation of *why* such bright-line rules are desirable. As a representative of a copyright industry coalition explained at a 2006 hearing, beyond the legal effect of any proposed exemption, there is also "kind of a meta issue" at play regarding "how [an exemption is] interpreted and how this would be communicated to the public and what message the public would get from it."[77] In 2010, this same group made strikingly similar arguments in both its written comments and at hearings.[78] Extending the scope of an exemption that allowed circumvention of CSS protection on DVDs for classroom use from professors to students, the group argued, would "work at cross purposes" with its members' "extensive educational campaigns designed to instill a respect for copyright in young people."[79] Granting another exemption would interfere with "the average consumer's recognition that digital locks are not meant to be picked."[80] Similarly, in a 2009 hearing, the group's attorney expressed concern that expanding the circumstances where circumvention is allowed "starts to normalize the behavior."[81]

Two things are salient about such arguments. First and foremost, as industry participants readily admit, the stakes are higher than the possible legal effect of any given exemption because of the way exemptions might shape public perception and behavior. Cursory, non-expert coverage of the

74. Transcript of Testimony at 61, *Section 1201 Rulemaking* (U.S. Copyright Office May 2, 2003), *available at* http://www.copyright.gov/1201/2003/hearings/transcript-may2.pdf (testimony of Stevan Mitchell, Interactive Digital Software Ass'n).

75. Attaway Testimony, Apr. 3, 2006, *supra* note 67, at 76.

76. Transcript of Testimony at 317–18, *Section 1201 Rulemaking* (U.S. Copyright Office May 6, 2009), *available at* http://www.copyright.gov/1201/hearings/2009/transcripts/1201-5-6-09.txt (testimony of Fritz Attaway, Motion Picture Ass'n of Am.) [hereinafter Attaway Testimony, May 6, 2009].

77. Transcript of Testimony at 113–14, *Section 1201 Rulemaking*, Docket No. RM 2005-11A (U.S. Copyright Office Mar. 23, 2006), *available at* http://www.copyright.gov/1201/2006/hearings/transcript-mar23.pdf (testimony of Steven Metalitz, Joint Reply Commenters).

78. Joint Reply Comments of Ass'n of Am. Publishers et. al. at 34, *Section 1201 Rulemaking*, Docket No. RM 2008-8 (U.S. Copyright Office Feb. 2, 2009), *available at* http://www.copyright.gov/1201/2008/responses/association-american-publishers-47.pdf.

79. *Id.*

80. *Id.* at 66.

81. Attaway Testimony, May 6, 2009, *supra* note 76, at 302.

resultant regulations,[82] coupled with the imprimatur of the Copyright Office and the Librarian of Congress,[83] hold significant potential to affect public ideas about the legitimacy and desirability of circumvention. Those ideas are well poised to play a significant part in regulating consumer behavior due to the weakness of the other potential regulators.[84] This explains the copyright holders' efforts to hold the line and oppose any and all exemptions in a blanket fashion.

Second, it is worth noting just what sort of public perception the copyright industries want to maintain. Their representatives' talk of a "clear prohibition"—of principles and bright-line rules against circumvention[85]—stands in stark contrast to the DMCA's actual prohibition on circumvention, which has always been anything but clear or bright. Section 1201 bases the legality of circumvention on the nature of the DRM technology at issue, carves out seven separate exemptions from liability, and authorizes the Librarian of Congress to carve out others.[86] In short, judging by the stark difference between the perception of the law they wish to create and the law as it actually exists, the copyright industries want the public to believe that the law is more unequivocal and restrictive than it really is. New exemptions, however minimal their scope, would threaten that belief.

One way to view the copyright holders' efforts is as a struggle against what Lessig has called "ambiguation."[87] The term refers to the phenomenon whereby the polyvalence of a particular act or object comes to undermine the

82. Given the difficulty of explaining a complex regime like the DMCA succinctly, it should not be surprising that much of the coverage of the rulemaking tends to leave out subtle but significant details. For instance, few stories covering the 2010 smartphone "jailbreaking" exemption mention § 1201's important distinction between circumventing on one's own and making use of another's products or services to do so. *See, e.g.*, Jenna Wortham, *In Ruling on iPhones, Apple Loses a Bit of Its Grip*, N.Y. TIMES, Jul. 27, 2010, at B3, *available at* http://www.nytimes.com/2010/07/27/technology/27iphone.html; David Kravets, *U.S. Declares iPhone Jailbreaking Legal, Over Apple's Objections*, WIRED (July 26, 2010), http://www.wired.com/threatlevel/2010/07/feds-ok-iphone-jailbreaking/.

83. Industry representatives have repeatedly expressed concerns about the effect of these institutions' blessings on consumer perceptions. *See, e.g.*, Transcript of Testimony at 79–80, *Section 1201 Rulemaking*, Docket No. RM 2005-11A (U.S. Copyright Office Apr. 3, 2006), *available at* http://www.copyright.gov/1201/2006/hearings/transcript-april03.pdf (testimony of Steven Metalitz, Joint Reply Commenters) (expressing concern over "an exemption [being] given the imprimatur of the Copyright Office, the Librarian of Congress, and the law" and "how that would be portrayed to the public"); Joint Reply Comments, *supra* note 78, at 66 ("Granting the proposed exemption could confuse even law abiding consumers by placing the stamp of the Librarian's approval on the 'darknet' marketplace.").

84. *See supra* Section III.A.1.

85. *See supra* notes 74–76 and accompanying text.

86. *See* 17 U.S.C. § 1201 (2006); *see also supra* Section II.A.

87. *See* Lessig, *supra* note 42, at 1010–12.

fixed social meaning it previously held.[88] Here, the fact that a particular act of circumvention may or may not trench on the legal entitlement of the copyright owner dilutes the strong warning against transgression that these digital barriers were originally meant to send. Although the sources of beliefs about the appropriateness of circumvention are many and varied, the copyright holders may have a point: the lack of congruence between DRM technology, which protects in a blanket way, and the DMCA, which contains exemptions, could contribute to the disorder of those beliefs. This does not mean that eliminating exemptions, as the copyright holders suggest, is the right solution. But by stressing the importance of consumer beliefs, the copyright holders do seem to reach the same conclusion as the theoretical propositions sketched out above: norms play a key role in consumer behavior.

IV. THE IMPLICATIONS

Recognizing the significant role that technological norms play in governing people's interactions with copyrighted works raises two questions with implications that may stretch beyond the context of the DMCA. First, it forces us to revisit the orthodox appraisal of technology as a modality of regulation and ask whether the pessimism that marks this point of view is misplaced. Second, in recognizing the potential power of norms to structure conduct, it is worth considering to what extent the law ought to reflect such popular perceptions.

A. REASSESSING TECHNOLOGY

Commentators like Lessig tend to take a dim view of technology as a tool of regulation.[89] I do not disagree completely, but the case of DRM forces us to refine some of Lessig's insights. At least with DRM, the problem with regulation by technology seems not to be the fact that it is silent, anonymous, and nearly omnipotent, but rather the contrary: that it is visible and fairly ineffectual.

1. Technology's Transparency Deficit

Critics of regulation by technology routinely stress the significant transparency deficit presumably inherent in this approach to governance. As Lessig puts it, regulation by technology "is not seen as regulation" because it creates subtle changes to the field of play itself rather than openly changing

88. *Id.*

89. *See* LESSIG, *supra* note 3, at 113–14.

the rules of the game; it thus allows the regulator to "hide its agenda."[90] Grimmelmann echoes these concerns, arguing that "[i]t may not occur to those regulated by software . . . to conceive of a restrictive design decision as being a decision at all."[91] Moreover, he continues, in systems governed by software, "it may be nearly impossible to determine who made the relevant regulatory decision"[92]—a feature that significantly undermines accountability. Lee Tien has argued that the relative obscurity of such regulators allows them slowly and quietly to chip away at and alter existing norms.[93] Other commentators have voiced similar concerns.[94]

These observations are apt, but they do not apply uniformly to all attempts at regulation by technology. The case of DRM suggests that the expectations of users can suddenly bring a hitherto invisible technological restriction into full view. Grimmelmann offers the absence of a "record" button from streaming media players as an example of how technology can obscure what are often very conscious regulatory choices.[95] But many might expect to find such a feature, perhaps because they remember it from VCRs and cassette recorders, or are used to seeing it on cameras or professional digital video equipment. In other words, it is part of their mental model of such technology. These users will immediately notice the absence and perhaps seek out ways to remedy it. What makes a particular restriction visible or invisible thus need not be a function of how a technological system deploys that restriction. Rather, it has at least as much to do with the conceptual categories we bring to the interaction in question. Those who can conceive of a "record" or "copy" button might be wiser to technology's tricks.

Of course, this is not to deny Tien's suggestion that over time, the absence of these capabilities from our media technologies will lead to fewer and fewer people conceiving of, let alone desiring, record or copy capability, or any other features that manufacturers have chosen to suppress.[96] For the moment, however, this seems not to be the case: the prevalence of

90. *Id.* at 135.

91. Grimmelmann, *supra* note 43, at 1737.

92. *Id.*

93. Lee Tien, *Architectural Regulation and the Evolution of Social Norms*, 7 YALE J.L. & TECH. 1, 3–4 (2004).

94. *See, e.g.*, ROGER BROWNSWORD, RIGHTS, REGULATIONS, AND THE TECHNOLOGICAL REVOLUTION 16 (2008) ("On the East Coast, legalism at least lets regulatees know where they stand. By contrast, on the West Coast, those who are controlled stand only where their regulated environment allows them.").

95. Grimmelmann, *supra* note 43, at 1737.

96. *See* Tien, *supra* note 93, at 12 n.29.

circumvention suggests DRM's restrictions remain visible to many.[97] We might therefore refine the orthodox view on transparency by specifying that regulation by technology will suffer a transparency deficit only if the technology in question does not defy our expectations. In other words, there is an outer limit to the kinds of restrictions that technology can impose, and this outer limit remains somewhat under our (imperfect, unconscious) control.

2. Technology's Authority Deficit

Technology may not always need to operate silently in order to regulate. A keen observer of urban life who was quick to notice that Robert Moses had (allegedly) built highway overpasses in New York City low enough to prevent buses—and their typically poor, African American passengers—from reaching Long Island's beaches would nonetheless be powerless to raise those overpasses and let the buses through.[98] But where, as in the case of software, the technology in question is highly susceptible to individual hacking and other manipulation, where its restraints, physical or otherwise, can readily be overcome, whatever power to regulate it may have resides entirely in the absence of transparency surrounding the act of regulation. This is because, as a conceptual matter, a technological measure lacks the *authority* to regulate and must therefore rely for its power to do so on the twin ruses of restraint and concealment.[99]

Tellingly, technological regulations appear to lack authority as both the legal positivist and natural law traditions define this concept—an impressive feat given that legal positivists tend to view themselves in opposition to natural law theorists.[100] On the legal positivist account, the law is morally neutral. Thus, as Joseph Raz has argued, its claims to legitimacy derive from the way a putative authority lays down a prescription for conduct. Rather than letting an individual reason about and resolve the matter that it seeks to

97. *See, e.g.*, Eric Pfanner, *File-Sharing Site Violated Copyright, Court Says*, N.Y. TIMES (Apr. 17, 2009), http://www.nytimes.com/2009/04/18/world/europe/18copy.html (citing analyst reports that enforcement actions do not suffice to curb the volume of piracy); *see also infra* Section IV.B.1 (describing prevalence of "noncommercial videos" created through circumvention on video-sharing sites like YouTube).

98. *See* Winner, *supra* note 36, at 123–24 (recounting the story of Moses's low overpasses, and the "master builder's" alleged motivation in building them).

99. *See supra* notes 90–94 and accompanying text.

100. John Finnis, *Natural Law Theories*, *in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2011), http://plato.stanford.edu/entries/natural-law-theories/ ("Legal theorists who present or understand their theories as 'positivist,' or as instances of 'legal positivism,' take their theories to be opposed to, or at least clearly distinct from, natural law theory.").

regulate, the authority, if it is to be viewed as legitimate, must simply furnish its own answer, which must be followed.[101] However, the consumer who finds himself thwarted by DRM has no particular reason to accept the answer that DRM offers: the clash between what he has sought to do and what DRM permits practically forces an evaluation of the sought-after interactions on their merits. Applying Raz's model, then, one might say that DRM fails to forestall a moral deliberation and thus fails to govern with any legitimate authority.

Natural law theory, by contrast, makes morality central to law's authority.[102] Working in a modern strand of this tradition, Roger Brownsword draws a distinction between ordinary and "techno" regulation. Regulation by law, he argues, can make use of a "moral pitch," signaling either that "the regulatory position is morally legitimate, or that . . . compliance . . . is morally obligatory."[103] Regulation by technology, however, does not "engage in any kind of moral discourse with regulatees," nor does it "rely on moral discipline or obedience to authority" because such regulation "by-passes practical reason altogether."[104] To be sure, Brownsword is skeptical about how much resistance circumvention and similar hacks might allow, and he warns that relying on such means to counter the creep of techno-regulation would be complacent.[105] Nonetheless, it is noteworthy that Brownsword locates the power of such regulation entirely in its ability to restrain, physically or technologically.[106]

To Brownsword, then, DRM lacks authority because it fails to make a moral appeal. On the legal positivist view, DRM lacks authority because it fails to lay down a rule of conduct. Ultimately, both points of view suggest that DRM lacks authority because it fails to *engage* users in any meaningful way.

---

101. *See* JOSEPH RAZ, THE MORALITY OF FREEDOM 53 (1986) ("[T]he normal way to establish that a person has authority over another person involves showing that the alleged subject is likely better to comply with reasons which apply to him . . . if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to follow the reasons which apply to him directly."). For an application of Raz's argument to the authoritativeness of law, see JEREMY WALDRON, LAW AND DISAGREEMENT 95–96 (1999).

102. *See, e.g.*, DERYCK BEYLEVELD & ROGER BROWNSWORD, LAW AS A MORAL JUDGMENT 176 (1986) (explaining that " 'a law' [is] a Legally valid rule . . . if, and only if, it is not immoral to posit the rule for attempted enforcement," meaning that "no wrong is done by positing the rule for attempted enforcement").

103. BROWNSWORD, *supra* note 94, at 243–44.

104. *Id.* at 246–47.

105. *See id.* at 247.

106. *Id.*

3. A Market Solution

To summarize, technology appears to fail as a regulator in the copyright sphere both because the sought-after regulations are too out of step with what consumers expect of copyright technologies, and because the technology itself holds limited power to shape those expectations. If copyright holders' paramount goal is to decrease instances of circumvention, one way of doing so might be by catering more closely to consumer expectations. Removing digital locks from at least the benign capabilities consumers desire would decrease the need for circumvention, and could help to shore up a non-circumvention norm. Such a response seems eminently possible, but thus far, there has been little movement on this front.

As some commentators have shown, DRM need not be an all-or-nothing proposition.[107] Existing technology would allow copyright holders to define permissions in a far more subtle and responsive way. Granted, this sort of tailoring would approach the kind of balance that traditional copyright law creates between owners and the public only imperfectly. After all, as Lessig stresses, copyright operates largely through complex standards.[108] The fair use inquiry involves a careful balancing of four statutory factors.[109] Evaluating infringement of the reproduction right likewise can involve the ambiguous standard of substantial similarity.[110] Technology, which lacks the capacity for discretion and thus cannot transform even a highly complex bundle of rules into a standard,[111] cannot hope to reproduce this system. Nonetheless, by allowing more of the kinds of capabilities users expect to find, even if not *all* such capabilities, this approach would bring DRM closer into line with consumer expectations.

Such a strategy might see producers differentiate their products by the level of technological protection imposed on them. Indeed, some market participants have taken this approach, particularly in the online music space.

---

107. *See* Stefan Bechtold, *The Present and Future of Digital Rights Management—Musings on Emerging Legal Problems*, *in* DIGITAL RIGHTS MANAGEMENT: TECHNOLOGICAL, ECONOMIC, LEGAL AND POLITICAL ASPECTS 597 (Eberhard Becker et al. eds., 2003); Armstrong, *supra* note 3, at 99–108.

108. LESSIG, *supra* note 3, at 187 ("Fair use inherently requires a judgment about purpose, or intent. That judgment is beyond the ken of even the best computers.").

109. *See* 17 U.S.C. § 107 (2006). *But see* Pamela Samuelson, *Unbundling Fair Uses*, 77 FORDHAM L. REV. 2537, 2541 (2009) (arguing that "fair use law is both more coherent and more predictable than many commentators have perceived once one recognizes that fair use cases tend to fall into common patterns, or . . . policy-relevant clusters").

110. *See* 3 WILLIAM F. PATRY, PATRY ON COPYRIGHT § 9:64 (2011) ("As a result of the necessary vagueness of the substantiality inquiry, no a priori line can be fixed to determine when appropriation is substantial.").

111. *See* Grimmelmann, *supra* note 43, at 1732–34.

Most notably, in 2007, Apple announced that its iTunes store would begin to sell music without any DRM.[112] Major publishers of audio books, including Random House, Penguin Group, and Simon & Schuster, followed suit a year later.[113] And in response to the 2006 classroom movie-clip DMCA exemption for faculty members, the motion picture industry claims to have begun work on a web-based content delivery system that would allow educators to download materials for classroom use, making circumvention unnecessary.[114] Many other sectors, however, have not followed suit. For the majority of video consumers, the blanket restriction on copying imposed by CSS, for instance, remains the standard for protecting content on DVDs.

Of course, there are good public policy reasons to oppose such an approach: charging consumers more for a DVD that would allow them to copy small clips of material, for instance, could be a step on the way to so-called "fared use"—a system that would allow copyright holders to profit from uses that lie beyond the scope of their exclusive rights.[115] Nonetheless, it is noteworthy that so few copyright holders have opted for this path. After all, if one's goal is to prevent consumers from circumventing, one way to do so might be to sell them products that let them do what they would like without having to circumvent.

That this has not happened might stem from an unusual kind of market failure: a refusal to respond to market signals out of fear or distrust. Because few have risen to answer these fairly ordinary market incentives, a new kind of motivator seems to have sprung up on the other end. The roadblocks that DRM poses may prompt the technologically inclined to develop circumvention tools. Perhaps this kind of response is to be expected when law diverges too far from norms.[116] Perhaps it is to be expected all the more when regulatory *technology* diverges too far from norms, as the remedy then

112. *Apple Unveils Higher Quality DRM-Free Music on the iTunes Store*, APPLE.COM (Apr. 2, 2007), http://www.apple.com/pr/library/2007/04/02itunes.html.

113. Brad Stone, *Publishers Phase Out Piracy Protection on Audio Books*, N.Y. TIMES (Mar. 3, 2008), http://www.nytimes.com/2008/03/03/business/media/03audiobook.html.

114. Comments of MPAA, *supra* note 68, at 10.

115. *See generally* Tom W. Bell, *Fair Use vs. Fared Use: The Impact of Automated Rights Management on Copyright's Fair Use Doctrine*, 76 N.C. L. REV. 557 (1998) (arguing that the scope of users' ability to make fair uses shrinks as DRM allows copyright owners to collect licensing fees for a growing number of uses). Bell views the shift to fared use as a development to be encouraged. *Id.* at 579–600. While the specific merits and shortcomings of such a shift are beyond the scope of this Note, one must at the very least acknowledge that the move toward fared use would radically alter the division of rights between users and copyright owners under the current Copyright Act.

116. *See* Gervais, *supra* note 40, at 50–53 (describing the response of technologists to social norms surrounding file sharing by developing new P2P platforms).

becomes a problem of pure engineering. On this view, circumvention is the cost of ignoring one's customers.

B. GIVING NORMS THEIR DUE

1. The Register's Nod to Norms

Market participants are not the only ones who would do well to take heed of the norms surrounding circumvention and the appropriate uses of copyrighted works. Policymakers should pay attention, too. Indeed, the Register of Copyrights has begun to take notice of such norms when evaluating proposed exemptions from liability for circumvention of access controls, in an apparent attempt to craft exemptions that better reflect public expectations. This is not to say, of course, that a widespread but clearly infringing activity—say, P2P file sharing of copyrighted works—would or should win an exemption merely because of its popularity. Nor is the Register simply giving certain norms direct legal recognition. But where a particular use is clearly, or at least colorably, noninfringing, an exemption seems more likely if the use is a popular one.

The Register requires proponents of exemptions to demonstrate that the DMCA's prohibition on circumvention bears a "substantial adverse effect on noninfringing use" of the class of works at issue.[117] Because the statute authorizes the Librarian to issue exemptions when "noninfringing uses . . . are, or are likely to be, adversely affected,"[118] the Register's addition of the word "substantial" makes for a stricter standard.[119] The practical effects of this requirement are twofold. First, proponents must demonstrate that DRM has occasioned a high degree of harm in order for the Register to recommend an exemption.[120] Second, this showing will be easier to make when the use in question is a widespread, commonly accepted activity—

117. *Section 1201 Rulemaking*, 64 Fed. Reg. 66,139, 66,141 (Nov. 24, 1999).

118. 17 U.S.C. § 1201(a)(1)(D) (2006).

119. The Register has faced criticism for this interpretation. In justifying her interpretation of the standard, she has pointed to parts of the DMCA's legislative history. *See Section 1201 Rulemaking*, 64 Fed. Reg. at 66,141–42. Others have contested this reading, including the Assistant Secretary of Commerce for Communications and Information. *See Section 1201 Rulemaking*, 65 Fed. Reg. 64,556, 64,562 (Oct. 27, 2000) (responding to "NTIA's observation that the word 'substantial' does not appear in section 1201(a)(1)(C)"). Some commentators attribute this addition to the Register's "eagerness to construct a relatively high burden of proof for proponents." Bill D. Herman & Oscar H. Gandy, *Catch 1201: A Legislative History and Content Analysis of the DMCA Exemption Proceedings*, 24 CARDOZO ARTS & ENT. L.J. 121, 169 (2006).

120. *See* Herman & Gandy, *supra* note 119, at 167–68 (discussing effect of the Register's adding "substantial" to the standard on the proponents' burden of proof).

when it reflects a widely held understanding of what one should be able to do with copyrighted works.

This framework still leaves the Register a good deal of discretion. Thus, in the first two rounds of rulemaking, the Register tended to recognize as legitimate only uses that we might describe as passive consumption: engagement with works such as video games[121] or ebooks[122] in precisely the way their creators envisioned, as a player or reader, and not the kind of transformative or productive engagement that doctrines like fair use might allow—engagement as a user, or an author in one's own right. Since no one would contest that these very basic, consumptive uses are legitimate ways of engaging with copyrighted works, there was hardly any need to determine how widespread consumer expectations about being able to make such uses ran; presumably, they would be ubiquitous. Instead, it sufficed for the Register that proponents showed that the DRM technology at issue provided a complete impediment to access.[123]

But the Register now seems to have recognized a broader universe of uses as legitimate, including ones that are more productive and transformative.[124] Accordingly, she seems to pay closer attention to how prevalent the use at issue in a given exemption appears to be. Thus, in extending the DVD circumvention exemption from media and film professors to creators of "noncommercial videos," the Register took notice of proponents' evidence about the popularity of such videos, citing various studies that put the number of noncommercial videos uploaded to YouTube each day at anywhere from 2,000 to 15,000.[125] Taking a more qualitative turn, the Register likewise recognized that "motion pictures are so central to modern American society and the lives of individual citizens that the need to

121. *See, e.g.*, *Section 1201 Rulemaking*, 68 Fed. Reg. 62,011, 62,014 (Oct. 31, 2003) (creating exemption for "[c]omputer programs and video games distributed in formats that have become obsolete and which require the original media or hardware as a condition of access").

122. *See id.* (creating exemption for "[l]iterary works distributed in ebook format when all existing ebook editions . . . contain access controls that prevent the enabling of the ebook's read-aloud function and that prevent the enabling of screen readers").

123. *See, e.g.*, Memorandum from Marybeth Peters, Register of Copyrights, to James H. Billington, Librarian of Cong., 80 (Oct. 27, 2003), *available at* http://www.copyright.gov/1201/docs/registers-recommendation.pdf (basing the decision to approve the ebook exemption in part on a clear finding that "a significant number of ebook titles are distributed only in formats that are not perceptible to the blind and visually impaired").

124. *See supra* Section II.A.

125. Memorandum from Marybeth Peters, Register of Copyrights, to James H. Billington, Librarian of Cong., 39 (June 11, 2010), *available at* http://www.copyright.gov/1201/2010/initialed-registers-recommendation-june-11-2010.pdf.

comment upon and criticize these works has become an important form of social discourse."[126] This observation echoed the view of the Assistant Secretary of Commerce for Communications and Information, who participates in the rulemaking[127] and who noted in his recommendation to the Register that "the world of online video has grown significantly, creating new norms and expectations that did not exist in 2006 and is becoming a new form of communication used worldwide."[128] (Indeed, the Register took more explicit note of these comments elsewhere in her recommendation.)[129]

Though her analysis was less extensive, the Register made similar observations in assessing the merits of the iPhone jailbreaking exemption. Her analysis noted proponents' assertion that "a very large number of purchasers of iPhones have circumvented . . . restrictions"[130] on third-party applications—some 1.8 million, of whom "approximately 400,000 are located in the United States."[131] The lesson? For proponents of exemptions, there is strength in numbers, which is to say, there is strength in norms.

2. From Norms to Law

The Register has just begun to look to such norms, and they are of course one input among many others considered in the rulemaking proceedings. Nonetheless, this reliance prompts the question whether such norms *ought* to play any role in shaping copyright law. And while there may be many good reasons to proceed with caution, on the whole the use of norms as guidelines in changing the shape of the law is a net positive and ought to be encouraged.

Justice Brandeis famously denounced as anarchy a state where "every man . . . become[s] a law unto himself."[132] So, at first glance, the idea of giving norms legal status offends cherished notions about the rule of law and democratic procedure. Norms are not the product of democratic institutions—they have a far hazier provenance in the muddle of social relations. Moreover, norms tend to be sticky.[133] Cass Sunstein has suggested

126. *Id.* at 70–71.

127. *See* 17 U.S.C. § 1201(a)(1)(C) (2006).

128. Letter from Lawrence E. Strickling, Assistant Sec'y of Comm. for Commc'ns and Info., to Marybeth Peters, Register of Copyrights, 6 (Nov. 4, 2009), *available at* http://www.copyright.gov/1201/2010/NTIA.pdf.

129. *See* Memorandum from Marybeth Peters, *supra* note 125, at 43 (repeating the Assistant Secretary's observation nearly verbatim).

130. *Id.* at 79.

131. *Id.* at 79 n.268.

132. Olmstead v. United States, 277 U.S. 438, 485 (1928) (Brandeis, J., dissenting).

133. *See, e.g.*, Dan M. Kahan, *Gentle Nudges vs. Hard Shoves: Solving the Sticky Norms Problem*, 67 U. CHI. L. REV. 607 (2000).

that this stickiness can go so far as to pose a threat to individual freedom.[134] Because changing norms requires collective action on a potentially massive scale, he argues, people are relatively powerless to effect such change and may remain bound by norms that encourage inefficient or even harmful behavior.[135] Given these criticisms, it would seem foolhardy to build our laws on such shaky foundations.

Then again, if law and norms diverge too starkly, that result may also seem undemocratic.[136] It is thus no surprise that the law frequently does give effect to norms, generally by relying on norms of various kinds to give content to standards. For instance, tort law defines the standard of due care by reference to what a reasonable person would do in a given situation.[137] Likewise, in certain circumstances, contract law relies on the concept of trade usage to aid in the construction of contractual terms.[138] Such "delegation" to norms is not entirely foreign in the copyright context. Courts evaluating the first factor of the fair use inquiry—the "purpose and character of the [allegedly infringing] use"[139] of the copyrighted work—will sometimes consider the social role of that use. In the landmark case *Campbell v. Acuff-Rose Music, Inc.*, for example, the Supreme Court held that a parodic appropriation constitutes fair use.[140] In reaching this conclusion, it found significant the social meaning of parody.[141] Indeed, a Copyright Office study leading up to the 1976 overhaul of the Copyright Act noted that fair use as a whole was, among other foundations, "said to be based on custom," and that

134. *See* Sunstein, *supra* note 4, at 910 ("There can be a serious obstacle to freedom in the fact that individual choices are a function of social norms, social meanings, and social roles, which individual agents may deplore, and over which individual agents have little or no control.").

135. *Id.*

136. *See* Gervais, *supra* note 40, at 50 n.49 (describing how a practice might pass from norm and custom into common and perhaps even constitutional law).

137. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 283 (1977) (describing the proper standard of conduct to avoid liability for negligence as "that of a reasonable man under like circumstances").

138. *See, e.g.*, U.C.C. § 1-205(5) cmt. 2 (2011) ("[T]he circumstances of the transaction, including . . . usages of trade . . . may be material.").

139. 17 U.S.C. § 107 (2006).

140. 510 U.S. 569 (1994).

141. *Id.* at 580 (explaining that parody "can provide social benefit, by shedding light on an earlier work, and, in the process, creating a new one").

"fair use is such use as is 'reasonable and customary.' "[142] The Supreme Court has cited this characterization of the doctrine with approval.[143]

Eduardo Peñalver and Sonia Katyal suggest that norms might be particularly worth noting and incorporating into the law when they appear to be out of step with other regulatory tools and thus foment disobedience.[144] The authors argue that such transgression, whether by trespass or circumvention, occupation, or duplication, plays an important informational role in the law of ownership: it serves as a kind of informal channel of communication from the governed to the government—a channel capable of signaling moments and places where the law is at loggerheads with public desires and expectations.[145] Property disobedience, Peñalver and Katyal argue, "is instrumental in provoking productive legal transitions that . . . foster innovation or equity,"[146] and "can play a powerful role in correcting for democratic and imaginative deficits in law and policy."[147] Their model carves out a special spot for disobedience in the intellectual property context—here, the messengers are not outlaws but "altlaws," because the tendency of IP entitlements to be defined by notoriously fuzzy boundaries means many such "transgressors" will in fact claim that their conduct ought to be recognized as falling within the contours of existing law.[148] As does the outlaw in tangible property, these altlaws play an indispensable informational function, pointing out trouble spots for the reform agenda.

The more the DMCA rulemaking process takes heed of consumer expectations, the more it institutionalizes Peñalver and Katyal's insights. This is a badly needed development, as the kind of information that such focus on norms might generate has always been in short supply in the framing of copyright law. Major copyright legislation has typically seen Congress brokering grand compromises between the competing interests of various copyright-related sectors, including the motion picture, music, and publishing industries—and, increasingly, software and telecommunications.[149] The voice

142. Alan Latman, *Fair Use of Copyrighted Works*, *in* STUDIES PREPARED FOR THE SUBCOMM. ON PATENTS, TRADEMARKS AND COPYRIGHTS, Study No. 14, at 7 (1960).

143. *See* Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 548 (1985) (citing Latman study).

144. For a summary of their argument, see PEÑALVER & KATYAL, *supra* note 56, at 11–16.

145. *See id.* at 15–16.

146. *Id.* at 171.

147. *Id.* at 172.

148. *Id.* at 76–82.

149. *See generally* Jessica D. Litman, *Copyright, Compromise, and Legislative History*, 72 CORNELL L. REV. 857 (1987) (explaining framing of 1976 Copyright Act as a process of compromise among major industrial stakeholders). On the growing role of the software and

of the public has always been absent from this process.[150] This notwithstanding the fact that the Constitution, with its vision of patents and copyrights promoting the "Progress of Science and useful Arts,"[151] designates the public as the ultimate beneficiary of intellectual property law.

How much of the Register's limited observations about copyright norms will influence legislators the next time Congress takes up significant copyright reforms remains an open question. However, the growing breadth of the DMCA exemptions and the appearance of norms among the justifications for these carve-outs both indicate that the copyright practices and visions of ordinary consumers would for the first time have the chance to offer an alternative to the well-financed and well-represented views of the various copyright industries, backed by the imprimatur of the Copyright Office and the Librarian of Congress. The Register's increased focus on norms could therefore be nothing but welcome. Of course, consumer norms cannot displace the legitimate interests of other copyright stakeholders outright. There are also potentially harmful norms, such as those that have helped infringement on file-sharing networks to proliferate. Some sort of limiting principle would thus be desirable. But perhaps it is too early to think about limiting a voice that has only recently begun to emerge.

V. CONCLUSION

A host of implications flows from the view of norms that this Note has advanced. Some come at a very granular level. If the DMCA rulemaking is to continue its recent turn toward assessing user expectations, the Copyright Office would do well to heed the suggestions of the Copyright Principles Project to hire economists and technologists to work alongside its lawyers.[152] The Copyright Office might also take up the suggestion of the Electronic Frontier Foundation and conduct independent fact-finding.[153] More broadly,

telecommunications industries, see Pamela Samuelson, *Intellectual Property and the Digital Economy: Why the Anti-circumvention Regulations Need To Be Revised*, 14 BERKELEY TECH. L.J. 519 (1999) (describing the competing interests of Hollywood and Silicon Valley in the framing of the DMCA); *see also* Peter S. Menell, *Envisioning Copyright Law's Digital Future*, 46 N.Y.L. SCH. L. REV. 63, 129–39 (2002) (describing the role of the content industries, as well as other countervailing interests, in the crafting of copyright legislation during the 1990s).

150. Litman, *supra* note 64, at 53 ("Right now the copyright-legislation playing field is completely controlled by its beneficiaries. . . . They are unlikely to countenance a statute that disempowers them in meaningful ways.").

151. U.S. CONST. art. I, § 8, cl. 8.

152. Pamela Samuelson et al., *The Copyright Principles Project: Directions for Reform*, 25 BERKELEY TECH. L.J. 1175, 1205–06 (2010).

153. *See* ELEC. FRONTIER FOUND., DMCA TRIENNIAL RULEMAKING: FAILING THE DIGITAL CONSUMER 8 (2005), *available at* http://w2.eff.org/IP/DMCA/copyrightoffice/

if notice and comment rulemaking has allowed policymakers to take heed of such customer expectations in a way unprecedented in the copyright sphere, then it is worth considering whether to expand this model beyond the narrow confines of the Librarian's current authority under § 1201.

At a higher level, however, this is a story of missed signals. Market participants have too frequently ignored the preferences of users, seemingly acting on the belief that instead of catering to their consumers' tastes, they must protect themselves against them. Policymakers have only just begun to pay attention, but this focus would be worthwhile in both spheres. It would also be worthwhile among commentators and scholars. Perhaps copyright scholarship is due for an ethnographic turn—a synthesis of empirical accounts of user behavior with searching reflection on what weight the law ought to give such facts on the ground.

DMCA_rulemaking_broken.pdf (suggesting directions for reform of the § 1201 rulemaking process).

MEDICARE AS TECHNOLOGY REGULATOR: MEDICARE POLICY'S ROLE IN SHAPING TECHNOLOGY USE AND ACCESS

April M. Elliott[†]

TABLE OF CONTENTS

† J.D., 2011, University of California, Berkeley School of Law. Ms. Elliott is an associate at Wilmer Cutler Pickering Hale & Dorr and served as a congressional aide from 2005 to 2008. The author wishes to thank Professor Kenneth Bamberger for his thoughtful guidance and support, and the Berkeley Technology Law Journal members and staff for their insightful and thorough assistance with this Article, for the opportunity to contribute to this esteemed publication, and for being a wonderful community throughout her years at Boalt and beyond.

I. INTRODUCTION

New technologies can present great hope, but they come with significant costs. Individuals, organizations, and the government regularly confront choices about how best to take advantage of technology. They pin their hopes on the promise of the next big breakthrough while trying to assess the value of new advances, the risks and reliability of using new technologies, and the cost of implementation—all while keeping pace with the latest innovations. Technological advancements, both dramatic and incremental, in communications, warfare, agriculture, transportation, and especially in medicine continue to change modern society in myriad ways.

It is axiomatic that scientific and technological progress drives modern medicine.[1] In the medical context, the push to develop and deploy new technologies ranging from cutting-edge pharmaceuticals to hardware often runs headlong into questions about the value, risks, and costs of new therapies as well as the breadth of public access to those therapies, including which individuals should have access at what stage of treatment.[2] Decisions

1. For example, the National Institutes of Health ("NIH"), the leading national public medical research agency, describes its purpose as funding research that helps Americans "liv[e] longer and healthier" by supporting "revolutionary ideas." NIH's slogan is "NIH . . . Turning Discovery Into Health." *About NIH*, NAT'L INSTS. OF HEALTH, http://nih.gov/about/ (last visited Nov. 17, 2011).

2. *See* U.S. FOOD AND DRUG ADMIN. (FDA), STRATEGIC PRIORITIES 2011–2015: RESPONDING TO THE PUBLIC HEALTH CHALLENGES OF THE 21ST CENTURY (2011), *available at* http://www.fda.gov/AboutFDA/ReportsManualsForms/Reports/ucm227527.

about cost, access, and regulation of health care technology are at the core of the Medicare reform debate. This debate implicates closely held values about life, health, choice, and the role of government as well as the strong financial interests of physicians, hospitals, insurance companies, pharmaceutical companies, and others in the health care industry.[3]

The tension between the promise of new technologies and the challenges inherent in putting those technologies to use permeates the debate about health care policy generally. This is particularly true in the world of Medicare, which as a national health insurance program not only subsidizes health care for millions of Americans and shapes health policy but also determines the kinds of therapies, treatments, and other services that it will fund, and at what level and under what circumstances it will fund them.[4] Effectively, Medicare's coverage determinations and benefits policies shape millions of Americans' access to types of care by determining how medical technologies can be used, by whom, and in some cases whether providers will adopt the use of new technologies at all.[5] Medicare regulates technology in obvious ways—for instance, by requiring a transition to electronic medical record keeping—but also in ways that may not be readily considered technology regulation, such as selecting the medical procedures for which it will provide reimbursement.[6]

Substantively, these reform efforts broadly concern three areas: rising costs, access, and quality.[7] The cost of health care services and the share of federal spending devoted to Medicare have consistently increased since Medicare's inception, growing from 0.7 percent of gross domestic product

htm (discussing goal of and strategies for "effectively oversee[ing] the translation of breakthrough discoveries in science into innovative, safe, and effective products and life-saving therapies for the people who need them most"); U.S. CONG. BUDGET OFFICE (CBO), TECHNOLOGICAL CHANGE AND THE GROWTH OF HEALTH CARE SPENDING (2008), *available at* http://www.cbo.gov/ftpdocs/89xx/doc8947/01-31-TechHealth.pdf.

3. *See* discussion *infra* Part III.

4. *See Fee Schedule—General Information*, CTRS. FOR MEDICARE & MEDICAID SERVS., http://www.cms.gov/FeeScheduleGenInfo/ (last modified Feb. 16, 2011).

5. For a discussion of the proposed national coverage determination ("NCD") for Computed Tomography Angiography ("CTA"), see *infra* Part IV.

6. *See* discussion *infra* Section III.C and Part IV.

7. *See* Michael Chernew, Professor of Health Policy, Harvard Medical Sch., *Physician Payment Post SGR* (May 5, 2011) (written testimony to *The Need To Move Beyond the SGR [Medicare Sustainable Growth Rate]: Hearing Before the Subcomm. on Health, H. Comm. on Energy & Commerce* (*H. Comm. Hearing on the Need To Move Beyond the SGR*), 112th Cong. (2011)), *available from Hearing: The Need To Move Beyond the SGR*, HOUSE ENERGY & COM. COMMITTEE, http://energycommerce.house.gov/hearings/hearingdetail.aspx?NewsID=8536 [hereinafter *Health Subcomm. SGR Hearing*].

("GDP") in 1970, to 1.9 percent in 1990, to 3.6 percent in 2010.[8] Medicare spending as a share of GDP is projected to continue to grow to 3.9 percent in 2020 and 5.1 percent in 2030.[9] Medicare spending in fiscal year 2010 is expected to reach $524 billion, which is fifteen percent of the federal budget and twenty percent of national health care expenditures.[10] This trend has fostered mounting concerns about the viability of Medicare spending in the future and garnered efforts across the political spectrum to rein in Medicare spending or reduce the rate of increase in Medicare spending (known as "bending the [cost] curve").[11] Costs have also increased for individuals (both beneficiaries and non-beneficiaries) and employers who provide health insurance to their employees, feeding debate about the second major issue in Medicare reform: access.

Access is a core concern for Medicare, and as a health insurance provider for the elderly and disabled Medicare has always endeavored to ensure that vulnerable populations have access to health care.[12] This vision for a basic, core level of access to Medicare appears to have quite broad political support,[13] but there are consistent challenges in defining the contours of Medicare access and determining whether the pool of beneficiaries should be expanded to provide care for more Americans in the face of the rising burden of costs on individuals and employers.

---

8. *Medicare Spending as a Percent of Gross Domestic Product, Fiscal Years 1968–2010*, *in* KAISER FAMILY FOUND. (KAISER), MEDICARE CHARTBOOK 79 (4th ed. 2010), *available at* http://facts.kff.org/chart.aspx?cb=58&sctn=169&ch=1796.

9. *Medicare Spending as a Percent of Gross Domestic Product 2000–2030*, *in* KAISER, *supra* note 8, at 84, *available at* http://facts.kff.org/chart.aspx?cb=58&sctn=169&ch=1806.

10. *Overview*, *in* KAISER, *supra* note 8, at 1, *available at* http://facts.kff.org/chartbook.aspx?cb=58.

11. *See* Peter Orszag, Office of Mgmt. and Budget (OMB), *Medicare Trustees to America: Bend the Curve!*, OMBLOG (May 12, 2009, 5:09 PM), http://www.whitehouse.gov/omb/blog/09/05/12/MedicareTrusteestoAmericaBendtheCurve.

12. Upon signing Medicare into law, for example, President Johnson proclaimed: "No longer will older Americans be denied the healing miracle of modern medicine." *President Lyndon B. Johnson's Remarks with President Truman at the Signing in Independence of the Medicare Bill, July 30, 1965*, LYNDON BAINES JOHNSON LIBRARY AND MUSEUM, http://www.lbjlib.utexas.edu/johnson/archives.hom/speeches.hom/650730.asp (last updated June 6, 2007).

13. *See* Kaiser, *Kaiser Health Tracking Poll: Public Opinion on Health Care Issues* 1 (Apr. 2011), http://www.kff.org/kaiserpolls/upload/8180-F.pdf (finding a total of eighty-nine percent of Americans support only minor spending reductions (32%) or no spending reductions at all (57%) to Medicare in order to reduce the deficit); *id.* at 3 (finding sixty-two percent of seniors support keeping the same level of Medicare benefits); *see also* Phil Galewitz, *Few Seniors Support GOP Plan To Restructure Medicare*, KAISER HEALTH NEWS (Apr. 27, 2011), http://www.kaiserhealthnews.org/Stories/2011/April/27/kaiser-poll-on-Medicare.aspx (discussing survey results).

The last major focus of reform is quality of care. Despite devoting a larger share of spending towards health care than other industrialized nations,[14] numerous studies demonstrate that health care outcomes and the quality of care provided in the American health system, on critical measures such as life expectancy, are worse than in other nations.[15] Many health care policy experts and organizations agree that the quality of care can be improved. They further agree that at least some Medicare spending goes towards care that does not improve health outcomes or even provides incentives for unnecessary care.[16] Thus, changes to Medicare policies require both investment in higher quality care and reduction in unnecessary spending.[17]

While issues of cost, access, and quality drive efforts to change Medicare policy, proposed reforms often run into stiff resistance. Conflicts over the value of care, the costs of care, and the role of government in shaping access, quality, and costs have consistently proven difficult to resolve. Changes to Medicare policy have significant effects on providers as well as beneficiaries, who collectively are quite active in pushing Congress and regulators to protect their interests.[18] Medicare policy is constantly in flux due to

14. *See* CHRIS L. PETERSON & RACHEL BURTON, CONG. RESEARCH SERV., RL34175, U.S. HEALTH CARE SPENDING: COMPARISON WITH OTHER OECD COUNTRIES 2–4 (2007).

15. *See id.* at 45–55.

16. *See Medicare Physician Payment Reform: Hearing Before the Subcomm. on Health, H. Comm. on Energy & Commerce*, 112th Cong. 2 (2011) (Internal Memorandum) [hereinafter *Health Subcomm. SGR Hearing Memorandum*] (stating that current Medicare payment policies "reward[] physicians for the volume of services they provide, not value or medical outcomes").

17. *See, e.g.*, Chernew, *supra* note 7; *Medicare Spending*, DARTMOUTH ATLAS OF HEALTH CARE, http://www.dartmouthatlas.org/keyissues/issue.aspx?con=1339 (last visited Nov. 20, 2011).

18. For example, of the lobbying organizations in the United States from 1998–2011, the second-highest-spending one was the American Medical Association (representing providers) and fourth-highest was the American Association of Retired Persons (representing beneficiaries) who often lobby on health care, in addition to Social Security and other matters. *See* Press Release, American Ass'n of Retired Persons (AARP), AARP to Members of Congress: Don't Cut Medicare, Social Security Benefits (Oct. 12, 2011), *available at* http://www.aarp.org/about-aarp/press-center/info-10-2011/aarp-to-members-of-congress-do-not-cut-medicare-social-security-benefits.html; *Seniors Groups Lobby Super-Committee*, SENIOR MARKET ADVISOR (Aug. 17, 2011), http://www.seniormarketadvisor.com/Exclusives/2011/8/Pages/Seniors-groups-lobby-supercommittee.aspx (discussing AARP lobbying Congress on Medicare and Social Security); Ctr. for Responsive Politics (CRP), *Lobbying: Top Spenders*, OPENSECRETS.ORG, http://www.opensecrets.org/lobby/top.php?indexType=s (last visited Nov. 16, 2011). Indeed, the top ten highest-spending lobbying organizations for the same period include the American Hospital Association (providers) and Blue Cross/Blue Shield (a private insurance company that provides Medicare Advantage and Medicare Prescription Drug Plans). CRP, *supra*; *The Blues and Medicare*, BLUECROSS

continuous legislative and regulatory changes[19] that often arise in response to pressure from providers or beneficiaries to address new developments in medicine[20] and even to adjust prior reform efforts.[21] But despite the regularity of Medicare policy change, the more significant reform efforts to address the core issues of cost, access, and quality have been sidelined[22] or have not been implemented at all due to pressure from providers or beneficiaries.[23]

Many of the core debates at issue involve exactly the kinds of difficult decisions we see elsewhere in technology regulation. Health care policy, and Medicare policy specifically, is unique in the degree to which it directly

---

BLUESHIELD ASS'N, http://www.bcbs.com/already-a-member/the-blues-and-medicare/ (last visited Nov. 16, 2011). The biggest spender, the U.S. Chamber of Commerce, *see* CRP, *supra*, has also engaged in significant lobbying on Medicare. *See Medicare Reform*, U.S. CHAMBER OF COMMERCE, http://www.uschamber.com/issues/health/medicare-reform (last visited Nov. 16, 2011). *See also* Christopher Rowland, *On Health Care, Lobbyists Flex Muscle: Medicare Overruled on Bone Scan Tests*, BOSTON GLOBE (May 31, 2010), http://www.boston.com/news/nation/washington/articles/2010/05/31/on_health_care_lobbyists_flex_muscle/ (describing providers and patient groups successfully lobbying Congress to reverse a CMS payment policy).

19. *See* Timothy Stoltzfus Jost, *Governing Medicare*, 51 ADMIN. L. REV. 39, 44 (1999).

20. *See id.* at 77.

21. *See Health Subcomm. SGR Hearing Memorandum*, *supra* note 16, at 3. Congress passed the Sustainable Growth Rate trigger, which took effect in 2002 and aimed to rein in Medicare spending. Pub. L. No. 106-113, § 211(b), 113 Stat. 1501, 1501A-348 (1999) (codified at 42 U.S.C. § 1395w-4(f) (2006)). However, with the exception of 2002, Congress has blocked the trigger's mandated cost reductions through legislation. *See Health Subcomm. SGR Hearing Memorandum*, *supra* note 16, at 3.

22. For example, the annual recommendations of the Medicare Payment Advisory Commission to Congress to reduce costs and improve the quality of care in Medicare are rarely implemented. *See* Ezra Klein, *News Break: How the White House Hopes To Control Health Care Costs*, WASH. POST EZRA KLEIN BLOG (June 3, 2009), http://voices.washingtonpost.com/ezra-klein/2009/06/breaking_how_the_white_house_p.html. While the landmark 2010 Patient Protection and Affordable Care Act ("PPACA") delegated some measure of increased authority to the Commission in the form of a newly established independent board, the Independent Payment Advisory Board ("IPAB"), many health experts have called for even greater authority in order to control costs and improve care. *See* Pub. L. No. 111-148, §§ 3403, 10320, 124 Stat. 119, 489, 949 (2010); sources cited *infra* note 39; *see also* Ezra Klein, *Four Ways To Improve the Medicare Board*, WASH. POST EZRA KLEIN BLOG (Apr. 20, 2011), http://www.washingtonpost.com/blogs/ezra-klein/post/four-ways-to-improve-the-medicare-board/2011/04/13/AFYsasDE_blog.html.

23. For example, while IPAB was included in PPACA, its authority has faced continued political opposition. Robert Pear, *Medicare Panel Meets Bipartisan Opposition*, N.Y. TIMES, Apr. 19, 2011, at A3; Ezra Klein, *Of Course Many in Congress Don't Want To Control Medicare Costs*, WASH. POST EZRA KLEIN BLOG (Apr. 20, 2011), http://www.washingtonpost.com/blogs/ezra-klein/post/of-course-many-in-congress-dont-want-to-control-medicare-costs/2011/04/13/AFDXcwBE_blog.html; *see also* Kathleen Sebelius, *IPAB Will Protect Medicare*, POLITICO (June 23, 2011), http://www.politico.com/news/stories/0611/57639.html (defending IPAB from congressional criticism).

affects human life and implicates deeply moral decisions. As a result, health care decision making is often paralyzed by partisan controversy and emotional debate. But examining Medicare policy from the perspective of technology regulation affords a useful framework for assessing the challenges inherent in Medicare reform.

This paper proceeds in Part II by providing a basic overview of Medicare operations and describes the role of Medicare as understood in academic literature. Part III describes the contours of debate in American health care policy generally and Medicare in particular, demonstrating the importance of Medicare policy, the importance of Medicare policy reform, and the significant challenges that Medicare reform efforts face. Part IV analyzes an example of a controversial proposal regarding Medicare coverage determination from a technology regulation perspective. Part V explains the implications from the technology regulation analysis in Part IV for current debates about payment reform in Medicare and ultimately suggests that Medicare shift to a different payment model.

II. MEDICARE OPERATIONS AND REGULATIONS

Medicare is a crucial provider of health care to millions of Americans and is a significant source of income for doctors, hospitals, and other health care providers.[24] On one end, Medicare provides medical insurance to eligible Medicare beneficiaries—individuals sixty-five and older, individuals with certain disabilities, and individuals with permanent kidney failure.[25] On the other, Medicare reimburses hospitals, physicians, and other providers for their services, often through insurance companies that act as intermediaries.[26] Medicare currently includes four programs. Two come from the original Medicare legislation: Part A for inpatient, hospital care, and Part B for outpatient care.[27] Congress added the other two parts in the past fifteen years: Part C or Medicare Advantage, which offers inpatient and outpatient care through private insurance companies,[28] and Part D for prescription drug coverage (prescription drugs were not previously covered by Medicare).[29]

---

24. *See generally* KAISER, *supra* note 8, *available at* http://facts.kff.org/chartbook.aspx?cb=58.

25. *See Medicare Program: General Information*, CTRS. FOR MEDICARE & MEDICAID SERVS., http://www.cms.gov/MedicareGenInfo/ (last modified Dec. 14, 2005).

26. "Intermediaries" or insurance companies under contract with CMS process payments and claims. *See* Jost, *supra* note 19, at 44.

27. Social Security Amendments of 1965, Pub. L. No. 89-97, 79 Stat. 286 (codified as amended at 42 U.S.C. §§ 1395c–1395i-5 (2006)).

28. The Balanced Budget Act of 1997 ("BBA") first introduced Part C. Pub. L. No. 105-33, 111 Stat. 251 (1997) (codified at 42 U.S.C. §§ 1395w-21 to -29 (2006)); Jost, *supra*

The Centers for Medicare and Medicaid Services ("CMS") is the principal agency charged with administering Medicare, although other agencies within the Department of Health and Human Services ("HHS"), as well as in the Social Security Administration ("SSA"), play a role in Medicare administration.[30] CMS organizes its operations in four consortia, which focus on Medicare health plans, Medicare financial management and fee-for-service operations, Medicaid and children's health operations, and quality improvement.[31] CMS issues regulations governing providers, such as services covered by Medicare through national coverage decisions, as well as regulations regarding claims processing and eligibility for payment.[32] CMS also enrolls beneficiaries and provides beneficiaries with information about their coverage options.[33]

But beyond these administrative responsibilities, CMS's power is relatively limited, especially with respect to policy-making, and there is scant legal scholarship on Medicare institutions. Timothy Stoltzfus Jost, who has written on legal issues in Medicare, describes Medicare as an "impossibly complex and technical regulatory program" that leaves little policy discretion to CMS.[34] Jost explains that because Medicare is perceived as a benefits administration institution with little policy-making authority, scholars have largely overlooked Medicare in favor of studying institutions with greater

note 19, at 44. The Medicare Prescription Drug Improvement and Modernization Act of 2003 (known as the Medicare Modernization Act or "MMA") modified Part C. *See* Pub. L. No. 108-173, § 221, 117 Stat. 2066, 2180–93 (2003) (amending Part C). These plans sometimes include Part D coverage. *See Medicare Benefits*, MEDICARE.GOV, http://www.medicare.gov/navigation/medicare-basics/medicare-benefits/medicare-benefits-overview.aspx (last visited Nov. 16, 2011).

29. MMA § 101, 117 Stat. at 2071–152 (codified at 42 U.S.C. §§ 1395w-101 to -152 (2006)).

30. *See generally* Social Security Amendments of 1965, 79 Stat. 286; *see also* Jost, *supra* note 19, at 82–88 (discussing the history of roles of CMS's precursor agency (the Health Care Financing Administration), the HHS Office of the Inspector General, and Social Security administrative law judges in Medicare administration).

31. *See Overview, CMS Consortia*, CTRS. FOR MEDICARE & MEDICAID SERVS., http://www.cms.gov/Consortia/ (last modified Mar. 8, 2011).

32. *See, e.g.*, *The CMS Quarterly Provider Update, April 2011*, CTRS. FOR MEDICARE & MEDICAID SERVS., http://www.cms.gov/quarterlyproviderupdates/emailupdates/itemdetail.asp?itemid=CMS1246822 (last modified Aug. 23, 2011).

33. *See, e.g.*, *Finding Plans*, MEDICARE.GOV, http://www.medicare.gov/default.aspx (last visited Nov. 17, 2011) (providing information to potential beneficiaries and current beneficiaries on enrollment, benefits, changing coverage, finding Medicare-eligible providers, etc.).

34. Jost, *supra* note 19, at 65–66.

regulatory power.[35] The relatively small amount of legal Medicare scholarship has focused on the (limited) role of judicial review in Medicare policy,[36] adjudication of payment disputes,[37] payment policies,[38] and the difficulties in effectuating reform through existing political institutions.[39]

However, much of the existing Medicare scholarship focuses on Congress, which is in fact the most active Medicare policy-making institution. Jost argues that Congress excessively micromanages and is "constantly tinkering" with Medicare, and he concludes that Congress should give bureaucrats more discretion in setting Medicare policy to better effectuate the policy goals of increasing quality of care and cost-containment.[40] Jost and other scholars argue that congressional involvement in Medicare for budgetary and ideological reasons has hampered rational policy-making in Medicare. Thus, they have called for the creation of an independent entity to implement quality enhancement and cost containment policies,[41] a version of

35. *Id.* at 42 (noting that even when scholarship has focused on benefit administration, it has principally been in the area of Social Security). For an example of such scholarship, see, e.g., JERRY L. MASHAW, BUREAUCRATIC JUSTICE, MANAGING SOCIAL SECURITY DISABILITY CLAIMS (1983) (examining the administration of Social Security benefits and its implications for the study of administrative law).

36. *See, e.g.*, Timothy Stoltzfus Jost, *Health Care Rationing in the Courts: A Comparative Study*, 21 HASTINGS INT'L & COMP. L. REV. 639 (1998); Eleanor D. Kinney, *The Role of Judicial Review Regarding Medicare and Medicaid Program Policy: Past Experience and Future Expectations*, 35 ST. LOUIS U. L.J. 759 (1991).

37. *See, e.g.*, Phyllis E. Bernard, *Empowering the Provider: A Better Way To Resolve Medicare Hospital Payment Disputes*, 49 ADMIN. L. REV. 269 (1997); Phyllis E. Bernard, *Social Security and Medicare Adjudications at HHS: Two Approaches to Administration of Justice in an Ever-Expanding Bureaucracy*, 3 HEALTH MATRIX 339 (1993).

38. *See, e.g.*, David M. Frankford, *The Complexity of Medicare's Hospital Reimbursement System: Paradoxes of Averaging*, 78 IOWA L. REV. 518 (1993); David M. Frankford, *The Medicare DRGs: Efficiency and Organizational Rationality*, 10 YALE J. ON REG. 273 (1993); Eleanor D. Kinney, *Making Hard Choices Under the Medicare Prospective Payment System: One Administrative Model for Allocating Medical Resources Under a Government Health Insurance Program*, 19 IND. L. REV. 1151 (1986); Eleanor D. Kinney, *Medicare Payment to Hospitals for a Return on Capital: The Influence of Federal Budget Policy on Judicial Decision-Making*, 11 J. CONTEMP. L. 453 (1985).

39. *See, e.g.*, LEN M. NICHOLS, NEW AM. FOUND., A SUSTAINABLE HEALTH SYSTEM FOR ALL AMERICANS 8 (2007), *available at* http://www.policyarchive.org/handle/10207/bitstreams/8826.pdf; PAUL N. VAN DE WATER, NAT'L ACAD. OF SOC. INS., DESIGNING ADMINISTRATIVE ORGANIZATIONS FOR HEALTH REFORM (2009), *available at* http://www.rwjf.org/files/research/vandewater.pdf; Timothy Stoltzfus Jost, *The Independent Medicare Advisory Board*, 11 YALE J. HEALTH POL'Y L. & ETHICS 21, 22 (2011).

40. Jost, *supra* note 19, at 44.

41. *See* Jost, *supra* note 39, at 22 (describing push by health policy experts, politicians, and scholars to create IPAB, which was initially named the "Independent Medicare Advisory Board"); David M. Cutler, *Health Reform Passes the Cost Test*, WALL ST. J. (Mar. 9, 2010), http://online.wsj.com/article/SB10001424052748703936804575108080266520738.html (describing an independent commission to control Medicare costs as an important method

which was included in the 2010 Patient Protection and Affordable Care Act ("PPACA").[42] But providers, lawmakers, and commentators have strongly criticized this approach, expressing concern about abdicating congressional control of Medicare policy to appointed "experts" who will play a large role in determining access to medical services.[43]

Other scholarship suggests that Medicare regulation is important not only for budgetary reasons and for ensuring quality of care provided through Medicare but also for the administration of private health care in America. Medicare payment rates, for example, can influence payment rates in the private sector,[44] suggesting that the impact of Medicare coverage determinations extends beyond Medicare beneficiaries and providers to the American health care system more broadly.

Medicare regulation is thus an important component of both national health and budgetary policy. The difficulty in overcoming values conflicts—discussed in Section III.A, *infra*—and pursuing policies that effectively promote quality care, cost containment, and access is evident in both the current structure of Medicare regulation as well as in debates about how to

for bending the cost curve); David Leonhardt, *Rockefeller: The Economists' Man in the Senate?*, N.Y. TIMES ECONOMIX (Dec. 8, 2009), http://economix.blogs.nytimes.com/2009/12/08/rockefeller-the-economists-man-in-the-senate (reprinting a letter from twenty-six "top health economists" at top universities and think tanks to Senate Majority Leader Harry Reid in support of establishing IPAB in order to control costs and improve quality of care by limiting congressional interference); Uwe E. Reinhardt, *A German Import That Could Help U.S. Health Reform*, N.Y. TIMES ECONOMIX (July 24, 2009), http://economix.blogs.nytimes.com/2009/07/24/a-german-import-that-could-help-american-health-reform/ (describing proposal for board of outside experts to control Medicare costs as "a very big deal"); *cf.* Henry J. Aaron, *The Independent Payment Advisory Board—Congress's "Good Deed,"* 364 NEW ENG. J. MED. 2477 (2011), *available at* http://www.nejm.org/doi/full/10.1056/NEJMp1105144 (supporting creation of IPAB to avoid congressional interference and describing challenges IPAB faces in overcoming congressional involvement); Sarah Kliff, *Experts Defend Medicare Board*, POLITICO (May 25, 2011), http://www.politico.com/news/stories/0511/55622.html (quoting health economists including "notable centrist" Alice Rivlin defending IPAB in connection with a May 2011 congressional hearing).

42. *See* PPACA, Pub. L. No. 111-148, § 3403(a)(1), 124 Stat. 119, 489–506 (2010); *see also id.* § 10320, 124 Stat. at 949 (changing name of the Board from "Independent Medicare Advisory Board" to "Independent Payment Advisory Board").

43. *See, e.g.*, Letter from AIDS Action Baltimore et al. to Senator Harry Reid, Majority Leader, U.S. Senate, and Representative Nancy Pelosi, Speaker, U.S. House of Representatives (Jan. 11, 2010), *available at* https://www.aamc.org/download/113668/data/group_letter_to_house_speaker_and_senate_majority_leader_re_sena.pdf; Joe White, *Getting the "IMAC" Proposal Right: Some Health Care Homework for the Recess*, ROLL CALL (Aug. 4, 2009), http://www.rollcall.com/news/37522-1.html.

44. *See* Kevin J. Hayes, Julian Pettengell & Jeffrey Stensland, *Getting the Price Right: Medicare Payment Rates for Cardiovascular Services*, 26 HEALTH AFFAIRS 124, 124 (2007); MEDICARE RBRVS: THE PHYSICIANS' GUIDE (P.E. Gallagher, Am. Med. Ass'n. ed., 2000).

improve Medicare regulation. An alternate perspective that examines Medicare regulation—in particular, coverage decisions that directly touch on issues of quality, cost, and access—from the perspective of technology regulation illustrates that many Medicare regulation challenges can be considered technology regulation challenges. Part III examines this perspective.

III. HEALTH CARE AND MEDICARE POLICY: DEBATING VALUES, POLICY ALTERNATIVES, AND TECHNOLOGY

American health care policy is a perennial source of debate and target of reform. Questions of access, cost, degree of government involvement, and quality of care fuel intense debates moored in deeply held values.

One debate centers on aspects of the value of life and health to individuals and to society: the value of providing individuals with choice among treatment options, the value of care that has the potential to prevent life-threatening or life-altering conditions down the road, the value of care that prevents untimely death, and the value of care that prolongs or increases comfort at life's end. These values are difficult to measure, monetize, and balance, but these are exactly the tasks that health care systems, and in particular Medicare, face.

Another values debate concerns the proper role of government in providing services and access to care. To what degree should the government be involved in the provision of care? What is the scope of services for which the government should pay? For whom and to what extent should the government ensure access to care? And what role should the government take in regulating the private health insurance market? The persistent challenge of answering these questions gives rise to the near-constant salience of health care policy in American politics. At times, the health care policy debate is front-and-center, such as during the debate over health care reform in 2010, discussed in Section III.A, *infra*. However, even when health care policy is not at the forefront of national political debate, Congress, the health care lobby, and CMS constantly debate the contours of Medicare regulation.

As these debates play out, another important question emerges: what is the proper role of Medicare as a technology regulator? Unlike the traditional "high-level" public discourse around Medicare regulation, discussed above, the questions raised by considering Medicare as a technology regulator are subtler. However, questions of how to regulate rapidly evolving technologies, manage the costs of new technologies, and regulate access to medical technologies are inherent in debates about Medicare regulation.

A. VALUES CLASH: HALLMARK OF THE HEALTH POLICY DEBATE

As previously described, the values tensions in the health policy debate exist in the context of the three main reform drivers: cost, access, and quality issues. More specifically, when one or more of these three drivers spur reform proposals, values conflicts often shape provider and beneficiary responses to these proposals. With such difficult-to-address values conflicts, and clear provider and beneficiary incentives to protect their own interests, proposals often run into strong resistance even when the proposals respond to significant needs for reform.

Recent Medicare reform debates illustrate these conflicts, including those debates that dominated the 2008 presidential campaign, the 2010 Congressional elections, and much of the first half of the Obama administration during debate over the PPACA. The public, government officials, candidates, members of the medical community and insurance industry, and an array of other interested parties fiercely argued over the role of government in providing access to care, the role of private insurers, and ways to ensure access to quality care while reining in the increasingly large costs of health care on the government, individuals, and employers. Some of the most hotly contested issues centered on the proper role of the government in resolving the direct conflict between the value of life, the amount that the government should pay to provide care that may prolong life, and how such decisions should be made. For instance, the controversies regarding so-called "death panels," the creation of the Independent Payment Advisory Board ("IPAB"), and more generalized concerns about Medicare "rationing" care nearly derailed critical pieces of health care reform legislation, demonstrating the difficulty in reconciling tensions between the value of life, the proper role of government, and the proper use of public funds.

Opponents of the health care reform bill coined the term "death panels" to criticize provisions that proposed optional end-of-life counseling sessions for seniors.[45] They characterized such provisions as unwanted and intrusive bureaucratic interference into the highly sensitive affairs of private individuals. Allegations that proposed reforms would establish "death panels" exemplify the extent to which conflicting values impact the health care reform debate. On one hand there is a very strong public sentiment regarding the value of life that individuals and their families should retain full autonomy in making end-of-life decisions and a strong resistance to

45. Jim Rutenberg & Jackie Calmes, *False 'Death Panel' Rumor Has Some Familiar Roots*, N.Y. TIMES, Aug. 14, 2009, at A1.

government being involved in such decisions. On the other hand, there is a very pragmatic concern that government funds should be spent on effective, proven treatments, and that decisions about end-of-life care should take into account the cost and projected benefits of that care.[46] While the health care reform did not include the controversial end-of-life counseling proposal[47] and included language preventing care from being "rationed,"[48] end-of-life coverage issues remain highly controversial.[49]

The controversy surrounding the creation of IPAB reveals similar underlying themes and conflicts within the health care debate. Unlike what happened to the end-of-life counseling provision, the final 2010 health care bill did include a version of the Board. The legislation calls for the establishment of an independent board of fifteen health experts from a variety of professional backgrounds to make recommendations for Medicare payment reform which, if not actively opposed by Congress, would go into

46. *See, e.g*, Ezra Klein, *Is the Government Going To Euthanize Your Grandmother?*, WASH. POST EZRA KLEIN BLOG (Aug. 10, 2009), http://voices.washingtonpost.com/ezra-klein/2009/08/is_the_government_going_to_eut.html; Robert Pear, *Obama Returns to End-of-Life Plan That Caused Stir*, N.Y. TIMES, Dec. 26, 2010, at A1 (discussing the debate over end-of-life counseling including the congressional debate, views of Obama administration officials, research, and interested citizens groups); Darius Lakdawalla et al., *Medicare End-of-Life Counseling: A Matter of Choice*, AEI OUTLOOK SERIES (Am. Enter. Inst., Washington, D.C.), Aug. 2011, http://www.aei.org/outlook/101070 (describing key issues in the end-of-life counseling debate, including impact on cost of end-of-life care, access to end-of-life counseling, and end-of-life care options).

47. *See* Foon Ree, *Senators Eliminate End-of-Life Provision, Respond to Charge of 'Death Panels*,*'* BOSTON GLOBE (Aug. 14, 2009), *available at* http://www.boston.com/news/nation/washington/articles/2009/08/14/senators_eliminate_end_of_life_provision/; *see also* Robert Pear, *Obama Returns to End-of-Life Plan That Caused Stir*, N.Y. TIMES, Dec. 26, 2010, at A1 (explaining that the final PPACA did not include the end-of-life counseling provision).

48. PPACA, Pub. L. No. 111-148, sec. 3403, § 1899A(c)(2)(A)(ii), 124 Stat. 119, 490 (2010) ("The [IPAB] proposal[s] shall not include any recommendation to ration health care, raise revenues or Medicare beneficiary premiums under section 1818, 1818A, or 1839, increase Medicare beneficiary cost-sharing (including deductibles, coinsurance, and copayments), or otherwise restrict benefits or modify eligibility criteria."); *see also* Sebelius, *supra* note 23 (Secretary of Health and Human Services explaining that the Independent Advisory Board does not have authority to ration care).

49. For example, in January 2011 the Department of Health and Human Services rescinded regulations permitting the inclusion of end-of-life counseling in Medicare visits, despite coverage of such discussions under Bush administration regulations. *See* Jason Millman, *White House Attempts To Quiet Revived Talk of 'Death Panels*,*'* HEALTHWATCH, THE HILL'S HEALTHCARE BLOG (Dec. 27, 2010), http://thehill.com/blogs/healthwatch/health-reform-implementation/135167-white-house-tries-to-smother-new-death-panel-talk ("[A]n administration spokesman said the regulation . . . is actually a continuation of a policy enacted under former President George W. Bush."); Robert Pear, *U.S. Alters Rule on Paying for End-of-Life Planning*, N.Y. TIMES (Jan. 4, 2011), http://www.nytimes.com/2011/01/05/health/policy/05health.html.

effect.[50] The President, several members of Congress, and various health policy experts advocated for the provision, asserting that independent experts were better suited to set policies that could effectively contain costs while improving the quality of care provided under Medicare.[51] The proposal faced harsh criticism, however, for giving too much control to unelected and unaccountable "experts" and for suggesting that Congress was ill-equipped to set Medicare policy.[52] The final provision included in the bill did not go as far as advocates had hoped in providing a mechanism to control Medicare costs based on data and expertise without congressional interference.[53] Yet, the provision remains under fire by various members of Congress and others who are concerned that the Board will be unaccountable, will "ration" care, and will not act in the best interest of Medicare and its beneficiaries.[54] Thus at the core of this conflict are different views on how to best control costs in Medicare to bring Medicare spending to a sustainable level. Ultimately, both supporters and opponents of an independent Medicare board argue that their approach is better for the health of Medicare beneficiaries and the fiscal health of Medicare, but the two sides disagree as to who can be trusted to promote these interests, with supporters looking to independent experts to make data-driven policy changes and opponents favoring greater control by Congress, physicians, and providers.[55]

As both the end-of-life counseling and IPAB debates illustrate, concerns about the government "rationing" care are particularly salient and limit the range of options available for controlling costs in Medicare. One prominent health economist, Uwe Reinhardt, described the tension between the public's general distaste for policies that "ration" care and the goal of controlling costs in Medicare as a battle between two competing "common sense" beliefs.[56] Reinhardt speculated that the conflict between these "common sense" beliefs made it nearly impossible to devise a health reform package

50. Sebelius, *supra* note 23.

51. *See* sources cited *supra* note 41; Letter from President Barack Obama to Senators Edward Kennedy and Max Baucus (June 2, 2009), *available at* http://www.whitehouse.gov/the_press_office/Letter-from-President-Obama-to-Chairmen-Edward-M-Kennedy-and-Max-Baucus/.

52. *See, e.g.*, Letter from AIDS Action Baltimore et al. to Senator Reid and Representative Pelosi, *supra* note 43; White, *supra* note 43.

53. *See, e.g.*, Jost, *supra* note 39, at 22; Klein, *supra* note 22; Leonhardt, *supra* note 41.

54. *See, e.g.*, Jost, *supra* note 39, at 22; Klein, *supra* note 23 (discussing criticisms of the Board); Pear, *supra* note 23; Sebelius, *supra* note 23 (addressing criticisms of the Board).

55. *See* Pear, *supra* note 23.

56. Uwe E. Reinhardt, *A 'Common Sense' American Health Reform Plan*, N.Y. TIMES ECONOMIX (July 31, 2009), http://economix.blogs.nytimes.com/2009/07/31/a-common-sense-american-health-reform-plan/.

capable of cutting government health care expenditures without bumping up against "rationing" fears.[57] Since these "common sense" beliefs appear impossible to disentangle from values that shape how Americans view the role of government in health care, Reinhart's discussion supports the inference that the debates about end-of-life counseling and IPAB indicate a broader clash in values that has been and likely will remain a hallmark of health care decision making.

B. CLASHING APPROACHES TO MEDICARE GOVERNANCE: AN EVER-CHANGING POLICY

Most of the running health care policy debates at the federal level center on Medicare. Congress passed major Medicare payment reforms in 1998 and 2003 and has regularly included changes to Medicare payment and coverage policies in annual budget reconciliation bills as well as in the PPACA.[58] In contrast to many other areas of federal regulation, congressional modification of Medicare is so common that many consider it a viable method for effectuating policy change.[59]

Federal health care policy debates often focus on Medicare because of its significant and expanding share of federal spending, which consistently thrusts Medicare into the center of debates about balancing the budget and reducing the federal deficit.[60] The major Medicare reforms of the past fifteen years have all included prominent measures at least nominally intended to control costs in Medicare. These measures strive to reform payments to health care providers[61] by cutting payments to physicians (and certain other providers) when the growth rate of payments exceeds the growth rate of

57. *Id.*

58. *See, e.g.*, PPACA, Pub. L. No. 111-148, 124 Stat. 119 (2010); MMA, Pub. L. No. 108-173, 117 Stat. 2066 (2003); Balanced Budget Act (BBA) of 1997, Pub. L. No. 105-33, 111 Stat. 251; Jost, *supra* note 19, at 67–70.

59. *See* Jost, *supra* note 19, at 43; Jost, *supra* note 39, at 22.

60. *See, e.g.*, H. COMM. ON THE BUDGET, THE PATH TO PROSPERITY: RESTORING AMERICA'S PROMISE, FISCAL YEAR 2012 BUDGET (2011), *available at* http://budget.house.gov/UploadedFiles/PathToProsperityFY2012.pdf [hereinafter RYAN BUDGET PROPOSAL] (proposing significant changes in Medicare to address the budget deficit).

61. *Compare* BBA, 111 Stat. 251 (setting the Sustainable Growth Rate system to automatically reduce Medicare expenditures when expenditures exceed the target growth rate), *with Health Subcomm. SGR Hearing Memorandum*, *supra* note 16, at 3 (describing repeated congressional action to prevent the cuts called for by the BBA).

GDP.[62] They also increase the role of independent experts in providing guidance to Congress on how to control Medicare expenditures.[63]

A number of officials and commentators on both sides of the aisle emphasize that the long-term fiscal health of the nation requires "bending the curve" of Medicare spending (i.e., reducing the growth rate of Medicare costs to more sustainable levels in the medium and long term).[64] Many cost-cutting proposals, however, run into such strong opposition that supportive lawmakers ultimately drop them.[65] Recent examples of this phenomenon include proposals for changing the nature and extent of Medicare coverage,[66] expanding the role of government in making treatment decisions,[67] and expanding the government's role in providing coverage.[68] Other reforms such as reduced physician payments do succeed but receive such harsh backlash upon implementation that they are never fully permitted to take effect; instead, Congress regularly votes to delay implementation of the

62. BBA, 111 Stat. 251; *see also Health Subcomm. SGR Hearing Memorandum*, *supra* note 16, at 2.

63. BBA established the Medicare Payment Advisory Commission (MedPAC), which advises Congress on Medicare policy, providing semi-annual reports to Congress with recommendations on a broad range of issues within Medicare. *About MedPAC*, MEDICARE PAYMENT ADVISORY COMM'N, http://www.medpac.gov/about.cfm (last visited Nov. 16, 2011).

64. *See, e.g.*, THE NAT'L COMM'N ON FISCAL RESPONSIBILITY AND REFORM, THE MOMENT OF TRUTH 36–42 (2010) (calling for significant Medicare funding changes); Karen Davis et al., *Bending the Cost Curve: Focusing Only on Federal Budget Outlays Won't Solve the Problem*, THE COMMONWEALTH FUND BLOG (Jan. 28, 2011), http://www.commonwealthfund.org/Content/Blog/2011/Jan/Bending-the-Health-Care-Cost-Curve.aspx; Orszag, *supra* note 11.

65. *See* Carl Hulse & Jackie Calmes, *G.O.P. Rethinking Bid To Overhaul Medicare Rules*, N.Y. TIMES, May 5, 2011, at A1 (noting that Ryan proposal was effectively dropped from legislative consideration). The PPACA did not adopt end-of-life counseling or the single payer model. *See* PPACA, Pub. L. No. 111-148, 124 Stat. 119 (2010).

66. *See, e.g.*, Michael McAuliff & Sam Stein, *Paul Ryan's Budget Becomes Bogeyman Uniting Progressives, Democrats*, HUFFINGTON POST (Apr. 25, 2011), http://www.huffingtonpost.com/2011/04/25/paul-ryans-budget-angry-democrats-town-halls_n_853553.html; Jake Sherman & Marin Cogan, *Some in GOP Squirm over Paul Ryan Budget*, POLITICO (Apr. 12, 2011), http://www.politico.com/news/stories/0411/53075.html.

67. For example, see the debate over proposals for end-of-life counseling in health care reform, which opponents called "death panels." *See Palin v. Obama: Death Panels*, THE FACT CHECK WIRE (Aug. 19, 2009), http://www.factcheck.org/2009/08/palin-vs-obama-death-panels/.

68. *See* SARA COLLINS ET AL., THE COMMONWEALTH FUND, AN ANALYSIS OF CONGRESSIONAL HEALTH CARE BILLS, 2007–2008: PART I, INSURANCE COVERAGE (2009), http://www.commonwealthfund.org/Content/Publications/Fund-Reports/2009/Jan/An-Analysis-of-Leading-Congressional-Health-Care-Bills--2007-2008--Part-I--Insurance-Coverage.aspx (finding that a proposal for single-payer health coverage would garner more savings than other models for reform).

cuts.[69] While there is widespread agreement on the need to ensure quality of care, including access to advanced medical technologies, and control costs, including managing costs associated with technological advances, Congress and CMS have largely failed to implement workable solutions.[70]

1. Medicare and the Federal Spending Debate

The annual budget-making process, especially in the summer of 2011, highlights the salience of Medicare in the federal spending debate. House Budget Committee Chairman Paul Ryan's proposal to substantially change the way Medicare provides coverage by reducing the role of the federal government and increasing reliance on private providers largely defined this year's debate about the federal budget deficit.[71] While the House voted to approve the Ryan budget,[72] the Chairman of the House Committee with jurisdiction over the kinds of Medicare changes included in the Ryan budget stated that he does not intend to even hold hearings on the Ryan reforms.[73] The Ryan budget proposal failed to pass the Senate, effectively killing the proposal for the time being.[74] Not surprisingly, however, Medicare reforms remain hotly contested in the national political discourse, in particular with regards to the debate over the national debt in the summer of 2011.[75] Like health reform, the proposed changes to Medicare have proven to be a political lightning rod, invigorating lawmakers and voters on both sides of the aisle.[76]

---

69. *See* Merrill Goozner, The Fiscal Times, *Medicare 'Doc Fix' Put on Life Support by AMA Lobby*, KAISER HEALTH NEWS (May 6, 2011), http://www.kaiserhealthnews.org/Stories/2011/May/05/Fiscal-Times-Medicare-Doc-Fix.aspx ("It never happened. . . . [E]very year, Congress voids the SGR-mandated cuts.").

70. *See* Cecil B. Wilson, Am. Med. Ass'n, *Statement* (May 5, 2011) (written testimony to *H. Comm. Hearing on the Need To Move Byond the SGR*, 112th Cong. (2011)), *available from Health Subcomm. SGR Hearing*, *supra* note 7; Jost, *The Independent Medicare Advisory Board*, *supra* note 39.

71. *See* RYAN BUDGET PROPOSAL, *supra* note 60.

72. H.R. Con. Res. 34, 112th Cong. (2011) (budget approved in the House by a vote of 235 to 193).

73. *See* Hulse & Calmes, *supra* note 65 (quoting Ways and Means Committee Chairman as stating he "had no plans" to consider the Ryan reforms in committee).

74. H.R. Con. Res. 34, 112th Cong. (2011) (budget rejected in the Senate by a vote of 57 to 40).

75. *See, e.g.*, David Rogers, *Senate Rejects Ryan Budget*, POLITICO (May 25, 2011), http://www.politico.com/news/stories/0511/55721.html (discussing the Senate "Gang of Six" debt negotiations regarding Medicare); Noam N. Levy & David S. Cloud, *Debt Deal Raises Pressure on Medicare Providers*, L.A. TIMES (Aug. 3, 2011), http://articles.latimes.com/2011/aug/03/nation/la-na-debt-impact-20110803 (describing debt ceiling compromise enacted Aug. 2, 2011 which will cut payments to Medicare providers if an alternative debt reduction plan is not adopted by the end of 2011).

76. *See* Hulse & Calmes, *supra* note 65.

2. Frequent Congressional "Tinkering" with Medicare

Congress is very active in setting Medicare policy, through both large- and small-scale interventions. In addition to attesting to the salience of Medicare policy for both budgetary and health care policy reasons, the near-constant congressional involvement in Medicare demonstrates two important things. First, making significant fundamental changes to Medicare's core functions and institutional policy-making processes is very challenging. Second, congressional tinkering with Medicare regulations has frequent and disruptive effects on the provision of Medicare benefits.

As discussed above, Congress has actively shaped Medicare through sweeping legislative provisions, such as the PPACA, MMA, and BBA, smaller provisions tucked into budget reconciliation and omnibus spending bills, as well as through stand-alone legislation.[77] Members of Congress also use a range of other oversight mechanisms to exert influence over Medicare policy, from holding formal hearings in the various committees with jurisdiction over various aspects of Medicare[78] to sending letters to CMS weighing in on Medicare regulations.[79] This level of involvement has provoked criticism that Congress is, in fact, over-involved in Medicare regulation and inhibits rational, expert-driven policy-making.[80]

3. Resistance to Large-Scale Change: Influence of Beneficiaries and Providers

Significant changes to Medicare are prone to invoking strong public resistance. Despite regular "tinkering," changes to the fundamental nature of

77. *See, e.g.*, statutes cited and text accompanying *supra* note 58.

78. In the House, jurisdiction over Medicare is split between the Ways and Means Committee and the Energy and Commerce Committee. *See Committee Jurisdiction*, H. COMM. ON WAYS & MEANS, http://waysandmeans.house.gov/About/Jurisdiction.htm (last visited May 8, 2011); *Subcommittees*, HOUSE ENERGY & COM. COMMITTEE, http://energycommerce.house.gov/subcomms/subcommittees.shtml (last visited May 8, 2011). The House Budget Committee also addresses Medicare policy in the annual budget. *See* RYAN BUDGET PROPOSAL, *supra* note 60. The Senate has a similar split in jurisdiction between the Finance Committee and the Health, Labor, Education, and Pension Committee, with the addition of the Special Committee on Aging which does not have primary jurisdiction but can hold hearings, draft legislation, etc. *See Committee Background*, SENATE SPECIAL COMMITTEE ON AGING, http://aging.senate.gov/about/index.cfm (last visited May 8, 2011).

79. *See, e.g.*, Letter from Rep. Peter Roskam et al. to Donald Berwick, CMS Administrator (Mar. 11, 2011) (regarding proposed regulations for "short-cycle dispensing" of pharmaceuticals under PPACA); Letter from Rep. Earl Pomeroy et al. to Donald Berwick, CMS Administrator (Aug. 9, 2010) (expressing concerns regarding proposed payment cuts for outpatient physical therapy services); Letter from Rep. Heath Shuler et al. to Kerry Weems, CMS Administrator (Dec. 16, 2008) (requesting that CMS delay a final rule regarding payment of home oxygen suppliers).

80. *See* discussion *supra* Part II and statutes cited *supra* note 58.

Medicare and to the guarantee of coverage it provides to millions of beneficiaries are widely considered politically infeasible.[81] The American Association of Retired Persons ("AARP"), for example, represents millions of those beneficiaries and consistently advocates to preserve[82] and, in the case of prescription drug benefits, to expand[83] Medicare benefits. The reaction to the Ryan Medicare proposal speaks to this resistance: Republican leadership in the House, despite expressing support for the proposal, dropped its support once it became clear the bill could not pass the Senate.[84] As Republican supporters of the proposals faced stark criticism, Democrats promptly began describing the proposal as a bid to "end Medicare" and commentators questioned whether the proposal would motivate seniors and baby-boomers approaching Medicare-eligibility to use their votes to oppose the plan and then punish Republicans at the ballot box in the next election cycle. These speculations were shown to carry water when, in a special election for a vacant congressional seat in New York, the claim that Republicans support policies to "end Medicare as we know it" appeared to be largely responsible for significantly cutting the Republican frontrunner's lead in the polls.[85]

Providers have also opposed large-scale changes to Medicare. Perhaps the clearest example is pressure exerted by providers on Congress to "fix" doctor payments every year to prevent reductions in payment rates.[86] While the American Medical Association ("AMA"), the preeminent physicians' organization in the United States, has recently begun to support more

81. For example, a poll published by the Kaiser Family Foundation in April 2011 found that "62 percent of seniors said they wanted Medicare to be left alone with the program continuing to guarantee the same benefits to all enrollees." Phil Galewitz, *Few Seniors Support GOP Plan To Restructure Medicare*, KAISER HEALTH NEWS (Apr. 27, 2011), http://www.kaiserhealthnews.org/Stories/2011/April/27/kaiser-poll-on-Medicare.aspx.

82. *See* Press Release, Am. Ass'n of Retired Persons, AARP to Members of Congress: Don't Cut Medicare, Social Security Benefits (Oct. 12, 2011), *available at* http://www.aarp.org/about-aarp/press-center/info-10-2011/aarp-to-members-of-congress-do-not-cut-medicare-social-security-benefits.html; Michael Muskal, *AARP to 'Super Committee': Hands Off Social Security and Medicare*, L.A. TIMES (Sept. 21, 2011), http://articles.latimes.com/2011/sep/21/news/la-pn-aarp-super-committee-20110921.

83. *See* Jonathan Cohn, *Republicans to AARP: Payback Time*, THE NEW REPUBLIC (Mar. 29, 2011, 5:01 PM), http://www.tnr.com/blog/jonathan-cohn/85941/house-republican-hearings-aarp-affordable-care-act (AARP supported the creation of the Medicare Prescription Drug Program in the Medicare Modernization Act of 2003).

84. *See* Hulse & Calmes, *supra* note 65; *cf.* H.R. Con. Res. 34, 112th Cong. (2011) (budget rejected in the Senate).

85. Raymond Hernandez, *G.O.P. Medicare Plan Shakes Up Race for House Seat*, N.Y. TIMES, May 5, 2011, at A1, *available at* http://www.nytimes.com/2011/05/06/nyregion/medicare-heats-up-house-race-in-upstate-new-york.html.

86. *See Health Subcomm. SGR Hearing Memorandum*, *supra* note 16, at 3.

comprehensive payment reform, the AMA's focus is still largely directed at preventing payment cuts.[87] Similarly, the American Hospital Association is principally focused on protecting or increasing payment rates to hospitals under Medicare,[88] and it supported legislation that would kill IPAB out of concern that hospitals would lose their ability to prevent cuts to provider payments.[89] The American Hospital Association has, however, expressed support for raising the Medicare eligibility age, in order to prevent payment cuts to hospitals.[90]

C. TECHNOLOGY REGULATION AS AN ALTERNATE PERSPECTIVE

As discussed above, the Medicare policy debate involves clashing values and approaches to governance characterized by weighty fiscal pressures, a strong public attachment to the core functions of Medicare, a resistance to an increased role for the government in making coverage decisions, and constant congressional tinkering both to address fiscal concerns and to respond to provider and beneficiary interests. And while the lenses of health care policy and budget policy are understandably the most common ways of viewing Medicare policy debates, they often result in the intractable conflicts discussed above. Thus, stepping out of the health and budget policy contexts to examine Medicare policy through the lens of technology regulation provides a useful perspective in addressing hot-button issues in Medicare policy.

There are two ways in particular that Medicare policy comprises a form of technology regulation. First, the policy directly regulates traditional technology matters: the substitution of electronic health records for paper records, the promotion of increased use of information technology, and the consequent challenges of investment in software, training, interoperability,

87. *See* Wilson, *supra* note 70, at 2; Jost, *supra* note 39, at 23.

88. *See, e.g*, *Expiring Medicare Provider Payment Provisions: Hearing Before the Subcomm. on Health, H. Comm. on Ways & Means*, 112th Cong. 2 (2011) (statement of Richard Umbdenstock, President and CEO, American Hospital Association) (calling for extension of current Medicare provider payment rates); Press Release, Am. Hosp. Ass'n, Statement of Rich Umbdenstock on President Obama's Debt Reduction Proposal to the Joint Select Committee on Deficit Reduction (Sept. 19, 2011), http://www.aha.org/presscenter/pressrel/2011/110919-pr-debt-reduction.pdf (criticizing proposed cuts to provider payments).

89. Susan Ferrechio, *American Hospital Association Backs Bill To Strike Medicare Panel from Health Law*, WASH. EXAMINER BELTWAY CONFIDENTIAL (Oct. 25, 2010), http://washingtonexaminer.com/blogs/beltway-confidential/2010/10/american-hospital-association-backs-bill-strike-medicare-panel-hea.

90. Susan Jaffe, *Medicare Eligibility Age Should Go Up, Hospitals Say*, POLITICO (Sept. 8, 2011), http://www.politico.com/news/stories/0911/63020.html.

and privacy.[91] Second, Medicare's national coverage determinations indirectly regulate technology by forcing decisions about the adoption and sustained use of certain technologies. The current Medicare reimbursement process requires a separate coverage decision for each new piece of technology available to providers. Therefore, providers often make decisions on whether to adopt new technologies on the basis of Medicare's national coverage determinations, which set the level of reimbursement when using a particular technology and the particular conditions under which it may be used.[92]

These indirect technology regulation issues are crucial to Medicare policy and benefit administration but may be less likely to be considered technology regulation matters. These types of decisions, however, are effectively technology regulation decisions. A national coverage decision for the use of a particular imaging technology or pharmaceutical therapy can impact which beneficiaries have access to it and under what circumstances, how much providers are reimbursed for the service, and ultimately whether providers determine they can offer the service at all. This kind of decision brings the conflict in values that often characterizes technology and health care decisions to the forefront: it contrasts the promise of new, potentially life-saving technologies with pragmatic decisions about the relative benefit in terms of outcomes and cost of a technology's use compared to other available options.

IV. EXAMINING A RESCINDED PROPOSED NATIONAL COVERAGE DECISION THROUGH THE LENS OF TECHNOLOGY REGULATION

The dispute over Medicare coverage for use of a particular cardiac imaging technology, Computed Tomographic Angiography ("CTA"), demonstrates the value of a technology regulation perspective. In December 2007, CMS proposed a national coverage determination ("NCD") for CTA that would have limited the use of CTA to symptomatic patients in the context of approved clinical trials.[93] Both CMS's coverage determination and providers' strong opposition centered on familiar concerns about the appropriate use of new technologies, the importance of ensuring that new

91. *See E-Health General Information Overview*, CTRS. FOR MEDICARE & MEDICAID SERVS., http://www.cms.gov/EHealthGenInfo/ (last modified Jan. 26, 2011).

92. *See Medicare Coverage Determination Process Overview*, CTRS. FOR MEDICARE & MEDICAID SERVS., http://www.cms.gov/DeterminationProcess/ (last modified Jan. 28, 2011); *see also* discussion of a failed proposed national coverage determination *infra* Part IV.

93. *See* Burt Cohen, *Medicare Reverses Decision To Limit Coverage of Cardiac CT Heart Scans*, ANGIOPLASTY.ORG (March 15, 2008), http://www.ptca.org/news/2008/031502_CT.html.

technologies are put to sound use, and the difficulty in managing the costs of new technologies. The rapid pace of CTA technology's evolution meant that data on effectiveness and risks swiftly became outdated, and the technology had dynamic and difficult-to-predict effects on the approach physicians took to cardiac diagnostics. Also, the costs of CTA equipment were substantial. Ultimately, CMS backed off of the proposed NCD and allowed CTA coverage determinations to be made at the local level.[94]

The conflict itself shows the technological dimensions of Medicare coverage decisions and the interplay between the drivers for reform and values conflicts. The CTA debate suggests that policies that better account for technological advances and efficient incorporation of technologies would be better not only from a technology regulation perspective, but also for addressing Medicare policy goals and values.

A. CMS RATIONALE: IMPROVE PAYMENT EFFICIENCY, REDUCE RISKS, AND ENSURE APPROPRIATE USE OF THE TECHNOLOGY

CMS explained that its proposal to limit Medicare support for CTA was based on data regarding the effectiveness and risks associated with the use of CTA. It was particularly concerned that physicians were performing CTA scans in addition to other cardiac imaging and diagnostic procedures, leading to unnecessary spending without corresponding improvements in treatment as well as increased risks to patient health.[95] CMS argued that CTA scans were prone to being used as an extra layer of imaging services that in most cases did not add value to other imaging methods or procedures. Under the existing model, physicians had an incentive to offer patients CTA scans, even if they did not actually add value in diagnosing patients or replace other existing services, because Medicare reimbursed providers per scan. CMS was therefore concerned that Medicare was paying for and encouraging unnecessary scans. Additionally, CMS worried that attendant radiation risks to patients from the use of CTA scans were unjustifiable.[96]

---

94. *See* Cohen, *supra* note 93 (reporting that by declining to issue a National Coverage Decision, CMS left "current coverage in place," which implicitly means coverage decisions remained at the local and regional level); *Medicare Coverage Determination Process Overview*, CTRS. FOR MEDICARE & MEDICAID SERVS., http://www.cms.gov/DeterminationProcess/ (last modified Jan. 28, 2011) ("In the absence of a national coverage policy, an item or service may be covered at the discretion of the Medicare contractors based on a local coverage determination (LCD).").

95. *See* Phurrough et al., *Decision Memo for Computed Tomographic Angiography (CAG-00385N)*, CTRS. FOR MEDICARE & MEDICAID SERVS. (Mar. 12, 2008), http://go.cms.gov/tNBB0c.

96. *Id.*

CMS's concerns arose in the context of a trend of increasing utilization of cardiac imaging services, as well as the emergence of data suggesting that more providers were offering more cardiac imaging services because they were well-reimbursed.[97] In cardiac imaging, Medicare sets prices for reimbursement for very specific technologies and uses, creating significant opportunities for price distortion, especially as changes in technology and clinical practice undermine the assumptions underlying payment rates.[98] The combination of financial incentives for providers and patients' strong desire for access to cutting-edge technologies exacerbated concerns that such new technologies were prone to excessive use from both a financial and clinical perspective.[99]

Examining data on CTA scans, CMS determined that the benefits did not justify broad use of the technology and thus proposed only paying for CTA scans used on patients with symptoms of angina and risk for coronary artery disease within designated clinical trials.[100] CMS determined this would promote quality goals,[101] ensuring both proper use of a new technology and the efficient allocation of Medicare funds.

B. PROVIDER RESPONSE: CTA IS A WORTHWHILE INVESTMENT WITH POSITIVE DYNAMIC EFFECTS ON DIAGNOSTIC ASSESSMENT, IMPROVED ACCURACY, AND REDUCED RISKS

The provider community forcefully opposed CMS's decision. CMS received nearly 700 comments on the proposed NCD, the vast majority of which voiced opposition to the proposal.[102] These comments included letters

97. *See* Hayes, Pettengell & Stensland, *supra* note 44, at 125 (noting that growth in cardiac services performed in physician offices raises questions about accuracy of Medicare payment levels for such services); *id.* at 131–32 (noting that use of cardiac imaging services increased faster than use of other cardiac services between 1999 and 2004, with services in physician offices increasing most quickly); *id.* (noting that physicians chose to increase in-office cardiac imaging services, but cardiac imaging in other settings did not decrease to offset the increase); *cf.* Uwe E. Reinhardt, *Fees, Volume and Spending at Medicare*, N.Y. TIMES ECONOMIX (Dec. 24, 2010), http://economix.blogs.nytimes.com/2010/12/24/fees-volume-and-spending-at-medicare/ (describing increased rates of authorization of imaging services by self-referring physicians with financial interests in provision of such services).

98. *See* sources cited *supra* note 97.

99. *See* Cohen, *supra* note 93 ("CMS is afraid that everybody's going to stack tests—that you'd get a CT[A], then a nuclear stress test, then an invasive cath[eterization].").

100. *See* Phurrough et al., *Proposed Decision Memo for Computed Tomographic Angiography (CAG-00385N)*, CTRS. FOR MEDICARE & MEDICAID SERVS. (Dec. 13, 2007), http://go.cms.gov/u7pE80; Alicia Ault, *CMS Won't Limit Reimbursement for CT Angiography*, AM. COLLEGE OF EMERGENCY PHYSICIANS NEWS (May 2008), http://www.acep.org/content.aspx?id=37798.

101. *See* Phurrough et al., *supra* note 100.

102. *See* Ault, *supra* note 100.

from numerous professional medical organizations[103] and nearly ninety members of Congress.[104] Six national physicians' organizations, for example, asserted that CMS based its findings on outdated metrics, as the technology had developed significantly since the studies were conducted.[105] Opponents of the NCD claimed that the new data demonstrated that CTA scans were now more accurate, less risky to patient health, and capable of replacing existing imaging services.[106] Providers asserted that the newer CTA technology had become "the clinical standard for diagnosing [coronary artery disease]"[107] and was "advancing every month."[108] They argued that CMS relied on data from far less advanced technology—four-, eight-, and sixteen-slice imaging—whereas sixty-four-slice and higher had become the standard,[109] and a 320-slice technology was also available on the market.[110] Further, providers stated that new data suggested that CTA scans replaced the need for other tests[111] and thus would both "save money and reduce the number of invasive tests."[112] Additionally, providers pointed to data suggesting that the most advanced technology significantly reduced radiation exposure as compared both to prior CTA technology and existing alternative diagnostic tools.[113]

It is likely that the significant investment many providers had made in CTA technology partially drove their response. The machines are expensive,[114] so smaller providers who had recently purchased the technology in reliance on local coverage rates faced a significant risk of financial loss if

103. *See id.*

104. *See MITA Thanks Congress for CCTA Coverage Plea to CMS*, DIAGNOSTIC & INTERVENTIONAL CARDIOLOGY (Mar. 10, 2008), http://www.dicardiology.com/article/mita-thanks-congress-ccta-coverage-plea-cms.

105. Ault, *supra* note 100 (describing comments on the proposed NCD submitted to CMS by the American College of Cardiology, the American Society of Nuclear Cardiology, the American College of Radiation, the Society for Cardiovascular Angiography and Interventions, the North American Society for Cardiac Imaging, and the Society of Cardiovascular Computed Tomography); Cohen, *supra* note 93.

106. *See* Cohen, *supra* note 93.

107. Ault, *supra* note 100.

108. Cohen, *supra* note 93.

109. Ault, *supra* note 100.

110. Cohen, *supra* note 93.

111. *Id.*

112. Ault, *supra* note 100.

113. Cohen, *supra* note 93.

114. For example, in 2005, a 64-slice CT scanner by Phillips cost $1.5 to $2 million. *See Brilliance 64-Slice CT Scanner by Phillips*, MEDGADGET CARDIOLOGY/RADIOLOGY (Apr. 4, 2005, 5:04 AM), http://www.medgadget.com/2005/04/brilliance_64sl.html.

they could no longer bill Medicare for CTA scans.[115] As had been the case with other payment reforms, there were concerns that the coverage change would make the technology available for use in some settings but not others for pure cost, not clinical, reasons.[116] Providers were also able to make the case that the change in payment policy would limit beneficiaries' access to a promising technology.

Providers thus attacked the proposal from cost, quality, and access angles, tapping into values arguments about the promise of cutting-edge technology, the role of physicians in determining the appropriate use of new treatments, and the scientific soundness of CTA technology from clinical use and risk perspectives.

C. LESSONS FROM THE CTA DEBATE: ADVANCES IN TECHNOLOGY POSE SIGNIFICANT CHALLENGES TO SETTING APPROPRIATE PAYMENT AND COVERAGE POLICIES

CMS and the provider community's dramatically different perspectives are largely a function of challenges in keeping pace with and predicting how technologies will affect the practice of medicine, combined with Medicare's practice of setting very specific coverage and payment policies. While here CMS dropped the proposed policy, the concerns driving the policy remained unaddressed, such as the risks of payments distorting incentives to encourage unnecessary use of technology and unnecessary spending, and the risk that the side effects would outweigh the benefits of use. CMS did not acknowledge the need for policies to reflect rapidly advancing technologies, nor did it address either the dynamic potential of new technologies to reshape the status quo in patient care or the need to provide some level of security in investments in expensive technologies.

1. Keeping Pace with New Data

Perhaps the most apparent lesson applicable to Medicare policy-making is that in order for policies governing technology to address policy goals, they must be responsive to current data. Since CTA technology was developing at

115. This was one of the primary concerns communicated by physicians to members of Congress and congressional aides in order to garner signatures of support for the "Dear Colleague" letter sent to CMS. *See MITA Thanks Congress*, *supra* note 104. *Cf.* Reed Miller, *CMS Set To Cut Medicare Physician Fees for Cardiovascular Imaging*, HEARTWIRE (Nov. 3, 2009), http://www.theheart.org/article/1018537.do (paid subscription) (describing reaction in provider community to cuts proposed in the 2010 Medicare Physician Fee Schedule, including speculation that cardiologists would not be able to afford providing certain CT imaging services in non-hospital settings).

116. *See* Hayes, Pettengell & Stensland, *supra* note 44; Miller, *supra* note 115.

such a rapid pace when CMS announced its NCD, CMS's decision could not account for CTA's true benefits, costs, and risks. Where risks can counteract or overpower the health benefits of a new technology and costs are paid through public funds, certainly risks and costs must be considered and weighed against the anticipated benefits. The CTA case demonstrates that this can be difficult: because the data no longer accurately reflected the nature of the technology in use, CMS's assessment of the benefits and drawbacks of the use of CTA was flawed.[117]

While CMS would certainly have been better off grounding its initial decision in more current data, providers' descriptions of the technological advances suggest that "current" is really a moving target.[118] Relying on data at any particular time can be risky when the technology is undergoing rapid development. Here, however, it appears that a reasonable approach would have been to base the assessment on the iteration of the technology the industry considered to be "the clinical standard."[119] This suggests that one critical element of effective Medicare regulation of medical technology is the ability to keep pace with the most recent data, with an eye towards current industry use and standards. Doing so enables policies to better serve the underlying goals of promoting cost-efficiency and quality. Further, the more accurate the underlying metrics, the less likely policies are to distort provider incentives. Better data use would in turn serve quality goals by reducing unnecessary risks from overutilization, cost goals by reducing unnecessary spending, and access goals by preserving access to the technology in clinically appropriate settings.

2. Dynamic (and Possibly Distortive) Effects of Coverage Decisions

The CTA debate also suggests that Medicare policy should account for the dynamic effects new technologies can have on treatment. Because Medicare coverage decisions can incentivize providers to invest in or abandon a particular technology, Medicare decision making must not only keep pace with new data, but it must also account for distortive effects on utilization that payment policies may themselves create. Therefore, policymakers must also be sensitive to inadvertently creating incentives for over- or under-utilization of certain medical technology.

While a key presumption underpinning the current fee-for-service model of payment in Medicare is that providers will only authorize medically useful services, data suggest this presumption does not always hold. In the area of

117. *See* Ault, *supra* note 100.
118. *See* Cohen, *supra* note 93 (noting that CTA was "advancing every month").
119. *See* Ault, *supra* note 100.

imaging, in particular, even indirect financial incentives have been linked to increased imaging services usage.[120] For example, data have shown that physicians with a financial incentive to perform imaging services (i.e., where they own or lease imaging equipment) are more likely to authorize imaging services than those who do not.[121] Further, while Congress and the Medicare Payment Advisory Commission have been actively involved in the attempt to curtail incentives for overutilization of imaging services stemming from self-referrals, policymakers have found it challenging to keep pace with new fee arrangements that continue to provide incentives for overutilization.[122]

With CTA, the question of how to encourage productive use of the technology was central to the debate. Risks of overutilization were a concern because of a combination of the payment method (payment at time of service), the trends in increased utilization of diagnostic cardiac services, and the degree to which the promise of cutting-edge technology may generate demand for the technology even if it is not more effective.[123] In contrast, providers argued that CTA was a clinically superior option to the status quo and thus would displace technologies whose results were less accurate or that carried higher radiation side effects.[124] Part of this dispute rested on the issue of the age of the data discussed above. But the heart of this conflict goes to the inability of the Medicare model to set a coverage and payment policy on a very specific service-by-service basis capable of adequately accounting for possible duplicative or substitutive uses while minimizing financial incentives for excessive (or insufficient) utilization. Technologies like CTA have the potential to be used to provide both more efficient and less efficient care, but the payment model is not conducive to properly assessing the full potential impact of new technologies or incentivizing efficient and effective use in lieu of duplicative or even harmful uses.[125] As discussed in Part V, *infra*, a bundled payment model has the potential to address these concerns.[126]

120. Reinhardt, *supra* note 97 (discussing concerns identified by the Medicare Payment Advisory and Congress regarding self-referrals and utilization rate growth in imaging services).

121. *Id.*

122. *Id.*

123. *See* CBO, *supra* note 2; *see also* Reinhardt, *supra* note 97.

124. *See* Cohen, *supra* note 93.

125. This goes to one of the major critiques of Medicare payment policy from both a cost and quality perspective. The fee-for-service model is prone to incentivizing more care but not necessarily quality care. *See, e.g.*, *Medicare Spending*, *supra* note 17; *Health Subcomm. SGR Hearing Memorandum*, *supra* note 16, at 2.

126. *See* Reinhardt, *supra* note 97.

3. Security in Investments

In addition to demonstrating the need for Medicare policies to keep pace with current data on rapidly changing technologies and account for the dynamic impact of technological advances, the CTA debate demonstrates the need for policies that account for the often significant costs of acquiring and integrating new technologies into medical practices. While CMS was greatly concerned about distorted incentives and unnecessary costs, the sunk costs in new technologies on the provider side helped to drive providers' resistance to the proposed NCD.

For clinics or physicians who had recently invested in CTA technology, the prospect of receiving reduced or even no Medicare payments for CTA services posed a significant threat to their ability to perform diagnostic cardiac imaging services and possibly even the financial viability of their practice.[127] From a cost and quality perspective, this would not necessarily be a bad outcome if CMS's concerns about distorted incentives and unnecessary provision of services were well-founded (although more current data suggested that the cost and quality concerns were not well-founded). But when clinics close and providers cut back, patients may lose access to health care more generally because these clinics and providers invested in technology they can no longer afford to support. While the impact of a clinic closure in an urban area with numerous provider options for beneficiaries[128] could largely be limited to the physicians and staff of the clinic, in rural areas with fewer provider options[129] a clinic closure could have a significant impact on access to care and beneficiaries' ability to see providers of their choice. The degree of impact would necessarily vary according to the number and accessibility of other providers in that particular community, but it is clear that beneficiaries' options could be significantly constrained if policies were changed to make recent investments in technology unsupportable.

Moreover, the threat of losing Medicare reimbursement for the use of new technologies could discourage research and development in medical advancements, adversely impacting the quality of patient care in the future.[130] Certainly this, too, is not an approach Medicare should endeavor to promote.

127. *See MITA Thanks Congress*, *supra* note 104; Miller, *supra* note 115.

128. *See Hospital and Physician Capacity*, DARTMOUTH ATLAS OF HEALTH CARE, http://www.dartmouthatlas.org/data/topic/topic.aspx?cat=24 (last visited May 8, 2011) (noting that generally there are many more providers per capita in urban areas than in rural areas).

129. *See id.*

130. The Proposed National Coverage Determination would have rescinded coverage for procedures that had previously been approved in local coverage determinations. *See MITA Thanks Congress*, *supra* note 104; Miller, *supra* note 115.

Due to the rapid pace of development, some level of uncertainty in the value of technological investments is unavoidable. However, as was evident in the CTA debate, both providers and beneficiaries gain from investments in promising new technologies, as long as Medicare gives providers some level of security in their ability to recover costs. At the same time, the evolving contours of the benefits, costs, and risks of new technologies caution against incentivizing investment without adequate confidence in the technologies' ability to improve the quality of care or cost-efficiency. Medicare thus faces a real challenge in navigating the uncertainties attendant on new technologies while at the same time promoting high-quality, accessible, and cost-efficient care.

4. A Balanced Approach

There are two traditional ways in which Medicare has addressed the sorts of challenges implicated by the CTA debate. CMS has experimented with clinical trials, enabling some beneficiaries and providers to utilize new technologies while learning more about the benefits, costs, and risks associated with the given technology.[131] CMS has also delegated coverage determinations of new technologies to local officials[132] and has declined to make national coverage decisions unless and until the technology's benefits, costs, and risks are well-established. In effect, CMS allows local and regional Medicare intermediaries and carriers to set coverage policies for services for which CMS has not issued a national coverage determination.[133] Thus a Medicare-authorized insurance carrier in Arizona could decide to authorize coverage of a new service while authorized carriers in California could chose not to cover the service, or authorize the service under different conditions than the Arizona carrier.

But both the clinical trial and local coverage approaches run contrary to critical values and policy goals in Medicare. Clinical trials and local coverage decisions lead to inconsistent access to new technologies by providing beneficiaries only in certain areas with federally funded care that is not

131. This is a common approach in both CMS and Congress-driven Medicare policy, as well as in other areas of health and technology regulation, such as with FDA-approved clinical pharmaceutical trials.

132. Local coverage determinations are coverage decisions made by a "fiscal intermediary or a carrier . . . respecting whether or not a particular item or service is covered on an intermediary- or carrier-wide basis." *Local Coverage Determinations*, CTRS. FOR MEDICARE & MEDICAID SERVS., http://www.cms.gov/DeterminationProcess/04_LCDs.asp, (last modified July 12, 2011).

133. *Id.*

available elsewhere, thereby risking underutilization of beneficial technologies as well as wasteful spending on ineffective technologies.

In terms of both policy and values, a preferable resolution to the challenge posed by the significant costs of investing in new technologies strikes a balance between incentivizing national investment in promising technologies, discouraging investment in excessively risky or inefficient technology, and supporting the goal of ensuring broad access to as many promising technologies as possible. Unsurprisingly, finding a policy that adequately encompasses this balancing act is challenging.

Examining the CTA debate through the lens of technology regulation thus suggests three principal areas where Medicare policy should be particularly attentive to technology regulation concerns: (1) grounding coverage decisions in up-to-date information on the benefits, costs, and risks of a given technology; (2) accounting for dynamic effects that coverage polices for advancing technologies can have on medical practice; and (3) promoting smart investments in promising technologies with an eye towards broadly distributed access.

V. IMPLICATIONS OF THE TECHNOLOGY REGULATION PERSPECTIVE FOR MEDICARE REFORM: SHIFT TO A BUNDLED PAYMENT STRUCTURE

The lessons from the CTA debate apply beyond cardiac imaging coverage decisions, coverage decisions of rapidly developing technologies, or even coverage decisions themselves. The whole enterprise of Medicare administration revolves around identifying the technologies to which beneficiaries should receive access, determining the level of provider reimbursement, and deciding which beneficiaries should have access. This is certainly true for decisions related to the use of technology in providing or enhancing patient record-keeping services, such as electronic health records. It is also true for decisions about coverage for and access to prescription drugs and durable medical equipment.[134] Even determinations regarding the use of general care services not directly tied to technology, such as

134. Certainly in this context the Food and Drug Administration plays a prominent role in shaping what therapies and products are available to the public, but Medicare decisions about what to cover, at what levels, and under what circumstances can still be highly controversial and raise the same issues regarding the proper role of Medicare in shaping access to technology as in other areas. *See* Andie King et al., *ESA Controversy Continues To Draw Attention*, 36 DIALYSIS & TRANSPLANTATION 462 (2007) (discussing debate over the proper Medicare response, in terms of funding and access, to FDA warnings about the use of erythropoiesis stimulating agents in certain patients with kidney disease).

preventative care consultations, can be seen as essentially decisions that optimize the use of medical technologies. Modern healthcare is inexorably linked with the use of technology, and considering how healthcare policy implicates the use of technology is a necessary aspect to any policy decision.

Medicare's current approach to national coverage determinations on a fee-for-service basis is too rigid to respond to the challenges of effectively administering patient care in a swiftly-changing technological landscape. But a technology-oriented perspective may help Medicare to design policies that are more flexible and comprehensive in defining the scope of access to technologies that Medicare covers. These sorts of policies would promote meaningful improvements to quality, cost-effectiveness, and access. More flexible Medicare policy-making should also help navigate difficult values conflicts by giving providers greater discretion in determining the proper course of treatment, which helps address concerns about the role of the government in determining what treatments are available. For example, a more flexible payment system would (1) incentivize physicians to assess potential courses of treatment based on current research on medical effectiveness as well as comparative costs; (2) enable them to use new, pricier technologies when they are likely to be the most effective approach; and (3) limit incentives to use them if older, less expensive approaches were just as likely to be effective. By ensuring that physicians and their patients retain significant control over evaluating treatment options on a patient-by-patient basis based on current data, a more flexible system would avoid concerns about centralized "rationing" while encouraging more rational payment levels and effective courses of treatment.

Especially in the context of current debates about Medicare policy, these lessons line up with recent efforts to reshape Medicare payment reform by moving towards a bundled payment system.[135] Under a bundled system, instead of paying for each procedure or service performed, Medicare provides a set payment for treating a particular condition.[136] A recent hearing before the Health Subcommittee of the House Energy and Commerce Committee on payment reform focused largely on efforts to move away from the fee-for-service payment model.[137] Proposals included moving towards "a more bundled system, that pays for an episode of care or provides a global

135. *See, e.g.*, Mark McClellan, The Brookings Inst., *Prepared Testimony for the House Energy & Commerce Committee* 2–3 (May 5, 2011) (written testimony to *H. Comm. Hearing on the Need To Move Beyond the SGR*, 112th Cong. (2011)), *available from Health Subcomm. SGR Hearing*, *supra* note 7; Chernew, *supra* note 7, at 4.

136. *See id.*

137. *See Health Subcomm. SGR Hearing Memorandum*, *supra* note 16.

budget," which would "allow more flexibility for providers and obviate the need for purchasers (such as Medicare or private insurers) to micromanage payment systems."[138] Replacing the fee-for-service model with a bundled payment system would shift the focus of the payment model from the volume of services provided to the quality and efficiency of the care provided. Providers would enjoy wide discretion in allocating the amount per episode of care (whether in the form of a lump sum, per episode payment, or global budget) to best serve the beneficiary. Some advocates of a bundled system propose special incentives for efficient, effective treatment in the form of "quality bonuses," which reward those who provide excellent care, to be determined against dozens of set performance metrics.[139] Providers could then make cost-effective determinations, for instance, on whether the use of a new imaging technology is appropriate or whether an older, less expensive technology would provide adequate patient care.

A bundled payment model, if properly designed, would address the three lessons of the technology regulation analysis. First, a bundled payment system would grant providers more flexibility in determining when to adopt new technologies—incentivizing providers to monitor new data on available technologies and base decisions on the most current data. As suggested by the technology regulation analysis, this would promote adoption of new technologies only when medically sound and cost-effective. Second, by moving away from coverage decisions specific to individual technologies, bundled payments would allow providers to account for the dynamic effects payment policies have on new technologies, incentivizing providers to adjust their practices according to efficiency and effectiveness. It would also dramatically reduce, if not completely eliminate, the incentives under the fee-for-service model to "stack" services to increase payment.[140] Lastly, bundled payments would grant providers greater financial security in their decisions to invest in new and often costly technologies. Because providers would not be subject to coverage determinations altering payment rates, they would be incentivized to purchase new technologies only when expected quality benefits and cost-effectiveness justify the expenditure, while being discouraged from investing in unnecessarily risky or redundant technologies.

138. Chernew, *supra* note 7, at 4.

139. *Id.* at 5. Additional funds to cover high-risk patients for whom quality treatment exceeds the lump sum payment would also be available under a bundled payment model. *See id.* at 6.

140. "Stacking" services refers to the practice of performing duplicative procedures in order to receive payment for each service provided. CMS is concerned about this phenomenon. Cohen, *supra* note 93 ("CMS is afraid that everybody's going to stack tests—that you'd get a CT[A], then a nuclear stress test, then an invasive cath[eterization].").

Providers across the country would be able to make these determinations based on their medical judgment, the needs of their particular beneficiary pool, and any other relevant factors.

A bundled payment system may also ameliorate the values clashes that accompany Medicare policy decision making in several ways. For example, this model should appeal to libertarian sensibilities by shifting more control over individual treatment decisions to providers and beneficiaries. Similarly, a bundled payment system would limit the government's role in picking winners and losers among different medical technologies and therapy options. Also, by improving the cost-effectiveness of care delivery, more resources would be available to serve a broader patient pool—an outcome that would satisfy people and institutions concerned about the breadth of health care access.

Significant challenges remain in setting bundled payment levels that account for differences in patient populations that may involve higher-than-average costs for treating certain conditions. But by adjusting payment levels over time and determining payment rates based on the condition being treated instead of the individual procedure performed, a bundled payment system can provide incentives for both quality and cost-containment with less direct involvement by CMS, Congress, outside experts, and local Medicare plan administrators in treatment decisions. A bundled payment system thus has the potential to better regulate the use of technology in Medicare while reducing clashes between approaches to governing Medicare by promoting quality care, cost-efficiency, and limiting direct government involvement in treatment options.

VI. CONCLUSION

Analyzing Medicare policy from the perspective of technology regulation affords a useful approach for analyzing Medicare reform proposals. As the CTA debate illustrates, the technology regulation lens calls for increased flexibility in accommodating technological advances, while responding to new data on technological benefits, costs, and risks. It further calls for more comprehensive payment models that better account for the dynamic effects of new medical technologies and treatment strategies—payment models that provide incentives for medically beneficial utilization of new technologies while reducing incentives for overutilization. Lastly, it calls for incentives for smart investment in technology, and greater security in such investments, by affording providers greater control and responsibility for how they use new technologies in their practices. A shift away from the current fee-for-service model in Medicare and towards a bundled payment system can accomplish these objectives by paying providers based on the condition being treated

instead of the particular treatments they administer. In shifting the focus of payment to a patient's condition, the bundled payment system allows providers to determine how best to treat their patients, including when and how to utilize new technologies.

At the same time, a bundled payment system would advance Medicare policy goals that weigh heavily in the national political discourse surrounding Medicare. It would promote cost-effective, quality care while limiting direct government control over patient care. Thus, the technology regulation analysis suggests that a bundled payment system facilitates important Medicare policy goals, while avoiding some of the major obstacles that have prevented other reforms from being effectively implemented.