PRIVACY REGULATION AND INNOVATION POLICY

Yafit Lev-Aretz[*] *& Katherine J. Strandburg*[†]

22 YALE J.L. & TECH. 256 (2020)

Industry players and opponents of privacy regulation claim broadly that privacy regulation will "stifle" innovation. This Article responds by bringing together traditional theories of regulation and innovation policy, and applying them in the context of markets involving personal information. Dire predictions about regulation's impact on innovation are common in many arenas, but seem to hold particularly great policy sway with regard to information privacy regulation. Here, we seek to bring analytical clarity to the debate about information privacy regulation, by showing how the interplay between misaligned demand signals in personal information markets and incentive distortions associated with variation in the extent to which suppliers can appropriate returns from innovative activities jointly determine whether and how the unregulated market's innovation portfolio deviates from the portfolio of innovative activity that would be most socially desirable.

Our analysis suggests that the characteristics of personal data do entangle some sorts of privacy regulation with appropriability in ways that can affect innovation incentives. Privacy regulation's possible effects on innovation do not justify blanket opposition, however, because they depend on details of regulatory design. Moreover, some sorts of privacy regulation designed to address misaligned market demand signals can potentially mitigate failures of appropriability and provide a more socially beneficial portfolio

[*] Assistant Professor of Law, Zicklin School of Business, City University of New York. This Article was partly co-authored during Professor Lev-Aretz's Post-Doctoral fellowship at the Information Law Institute at New York University School of Law. ILI Fellowships are funded in part by a generous grant from Microsoft Corporation.

[†] Alfred B. Engelberg Professor of Law and Faculty Director, Information Law Institute, New York University School of Law. Professor Strandburg acknowledges the generous support of the Filomen D. Agostino and Max E. Greenberg Research Fund.

The authors are also grateful for terrific research assistance from Grace Ha, Melissa Arseniuk, Christopher Bettwy, Gabriel Ferrante, Melodi Dincer, and Sara Spaur, and for helpful comments from members of the NYU Privacy Research Group and Information Law Institute, attendees at the Privacy Law Scholars Conference 2017, NYU faculty workshop attendees, and New York Law School faculty workshop attendees.

of innovation incentives. Proposals for information privacy regulations should thus be judged on their individual merits, taking both misaligned market demand signals and failures of appropriability into account.

TABLE OF CONTENTS

I. INTRODUCTION

The amount of personal information accumulated by companies has mushroomed in recent years, giving rise to calls for more stringent information privacy regulation.[1] In the EU, such calls led to the enactment of the General Data Protection Regulation (GDPR), which came into effect last year.[2] US lawmakers have tended to be more skeptical about regulation than their European counterparts, at least at the federal level. As a result, the question of whether and how to regulate the commercial collection and use of personal information continues to be hotly debated.[3] Nonetheless, while federal proposals remain stalled, some states and even cities are moving ahead with privacy regulation.[4]

Proposals for heightened privacy protections are routinely countered with general claims that privacy regulation will stifle

---

[1] *See, e.g.*, Marc Rotenberg, *America Needs a Privacy Law*, N.Y. TIMES (Dec. 25, 2018), https://www.nytimes.com/2018/12/25/opinion/letters/data-privacy-united-states.html.

[2] *See* Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) [*hereinafter* GDPR]; *id.* at art. 99 ("[This Regulation] shall apply from 25 May 2018."). The EU had approved pan-European data protection rules the previous year. *See* Mark Scott, *Europe Approves Tough New Data Protection Rules*, N.Y. TIMES (Dec. 15, 2015), https://www.nytimes.com/2015/12/16/technology/eu-data-privacy.html.

[3] *See, e.g.*, Mike Masnick, *One Year Into the GDPR: Can We Declare It a Total Failure Yet?*, TECHDIRT (May 24, 2019), https://www.techdirt.com/articles/20190521/17425842255/one-year-into-gdpr-can-we-declare-it-total-failure-yet.shtml; Julie Brill, *GDPR's First Anniversary: A Year of Progress in Privacy Protection*, MICROSOFT ON THE ISSUES (May 20, 2019), https://blogs.microsoft.com/on-the-issues/2019/05/20/gdprs-first-anniversary-a-year-of-progress-in-privacy-protection (arguing for the adoption of regulations similar to GDPR); Cameron F. Kerry, *Why Protecting Privacy Is a Losing Game Today—And How to Change the Game*, THE BROOKINGS INSTITUTION (July 12, 2018), https://www.brookings.edu/research/why-protecting-privacy-is-a-losing-game-today-and-how-to-change-the-game.

[4] *See, e.g.*, 2018 Cal. Stat. ch. 55 (enacting California Consumer Privacy Act of 2018); 201 MASS. CODE REGS. § 17.03 (requiring that any entity that collects personal information of Massachusetts residents maintains comprehensive data security plans); 23 N.Y. COMP. CODES R. & REGS. tit. 23, § 500 (2017) (requiring financial institutions active in New York state to maintain comprehensive plans addressing cyber security risks); Data Collection and Protection Ordinance, CHICAGO, ILL., MUN. CODE §4-402 (2018) (providing consumers with opportunities to control personal data via informed consent to disclosure, information on use, and redress for misuse).

socially valuable innovation.[5] This rhetoric is powerful and superficially convincing. It goes something like this:

The information economy is the lifeblood of US economic growth. Increasingly, it runs on personal information collected and aggregated by companies as they provide us with services. The use of this information has brought us many benefits and conveniences and is the mainstay of our most successful companies. Sure, each of us might, in principle, prefer not to have our own activities tracked, but do we really want to risk stalling out the engine of our innovative economy by imposing privacy regulations?

Sweeping claims about the dire ramifications of regulation for innovation, jobs, and economic competitiveness are certainly not new.[6] Ideological and political battles over regulation continue

---

[5] For instance, Adam Thierer & Ryan Hagemann have argued that regulations on the collection and use of personal data in the field of driverless vehicles will lead to higher costs for start-ups and small operators, and prevent consumers from enjoying the potential benefits of innovations. *See* Adam Thierer & Ryan Hagemann, *Removing Roadblocks to Intelligent Vehicles and Driverless Cars*, 5 WAKE FOREST J.L. & POL'Y 339 (2015); *see also* THE INTERNET ASSOCIATION, *Re: Request for Comments Concerning Big Data and the Consumer Privacy Bill of Rights* (Docket No. 140514424-4424-01) (Aug. 5, 2014) 34 ("At this time, any legislative proposal, to address 'big data' may result in a 'precautionary principle problem' that hinders the advancement of technologies and innovative services before they even develop."); Bob Quinn, *Privacy Regulation: Symmetry or Asymmetry?*, AT&T: PUB. POL'Y BLOG (Mar. 9, 2016, 1:51 PM), https://www.attpublicpolicy.com/privacy/privacy-regulationsymmetry-or-asymmetry (asserting that the FTC's framework and the Obama Administration's 2012 Consumer Privacy Bill of Rights "[are] familiar to consumers, [have] worked well for them for many years, and contributed to today's thriving, innovative, and free Internet").

[6] *See generally* Wesley A. Magat, *The Effects of Environmental Regulation on Innovation*, 43 DUKE J.L. & CONTEMP. PROBS. 4 (1979) (claiming five main types of environmental regulation reduce technology innovation when a firm invests in compliance instead of pure research and development); Henry G. Grabowski, *Estimating the Effects of Regulation on Innovation: An International Comparative Analysis of the Pharmaceutical Industry*, 21 U. CHI. J.L. & ECON. 133 (1978) (comparing U.S. to U.K. firms to analyze the relation between increased FDA regulation and pharmaceutical research and development investment). *But see* Nathan Goldschlag & Alex Tabarrok, *Is Regulation to Blame for the Decline in American Entrepreneurship?*, 33 ECON. POL'Y 5 (2018) (finding increased federal regulation is not directly responsible for economic trends in the U.S. such as a decline in business startups and increase in job reallocation); Adam B. Jaffe & Karen Palmer, *Environmental Regulation and Innovation: A Panel Data Study*, NAT'L BUREAU OF ECON. RESEARCH, Working Paper No. 5545 (1996) (finding regulated industries' proportion of successful patent applications were not significantly impacted by compliance costs); Shunsuke Managi, *Environmental Regulations and Technological Change in the Offshore Oil and Gas Industry*, 81 LAND ECONS. 303 (2005) (finding support for a more restrained version of the Porter hypothesis after

to play out, most notably over climate change. In most regulatory arenas, however, a substantial and nuanced discussion about precisely what and how to regulate competes for the floor and influences, even though it does not control, regulatory policy and design.[7] The debate about information privacy regulation, however, seems mostly stuck at the shouting match stage, despite the growing influence of personal data collection, aggregation and use in society and despite growing exposures of misuse such as the infamous Cambridge Analytica debacle.[8] The information privacy regulations now on the books worldwide, including the GDPR, are nearly all based primarily on a set of Fair Information Practices (FIPs) drafted in the late 1970s and early 1980s,[9] despite the introduction of new concepts such as "privacy by design."[10] With

---

testing the relation between environmental regulation and technological innovation of offshore oil and gas industries).

[7] *See generally* Zachary Liscow & Quentin Karpilow, *Innovation Snowballing and Climate Law*, 95 WASH. U. L. REV. 2 (2012) (discussing long-term impacts on technological innovation as a significant consideration for policymakers when drafting climate policy); David Popp, *Innovation and Climate Policy*, 2 ANN. REV. RESOURCE ECON. 275 (2010) (surveying environmental innovation literature in relation to clean energy technologies and discussing its implications on climate policy); Emi Kolawole, *Health Care Innovation: From Regulation to 'Bigger Brains'*, WASH. POST (Feb. 14, 2012), https://www.washingtonpost.com/blogs/innovations/post/how-to-keep-the-us-on-the-cutting-edge-in-health-care-innovation/2012/02/14/gIQARJurDR_blog.html?utm_term=.fda4730333ad (discussing regulatory considerations around increased use of big data in health care technology innovation).

[8] *See generally* Kevin Granville, *Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens*, N.Y. TIMES (March 19, 2018), https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html.

[9] *See* FED. TRADE COMMISSION, *Fair Information Practice Principles* (2007), archived Mar. 31, 2009 at the WAYBACK MACHINE, https://web.archive.org/web/20090331134113/http://www.ftc.gov/reports/privacy3/fairinfo.shtm (2007) (providing five core principles of privacy protection, including notice to users, choice and consent regarding data collection and use, users' access to collected data, data security, and enforcement and redress measures); *see also* Robert Gellman, *Fair Information Practices: A Basic History* (Apr. 11, 2017), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2415020 (discussing the history of Fair Information Practices with a focus on the U.S.)

[10] *See* Gellman, *supra* note 9 at 28-30 (describing privacy by design as a departure from the "classic FIPs principle" of transparency). For further discussion of privacy by design measures, see Ira S. Rubenstein, *Regulating Privacy by Design*, 26 BERKELEY TECH. L.J. 1410 (2011) (analyzing the meaning of "privacy by design" in order to show how regulatory incentives might be balanced against economic costs of compliance with privacy regulations); EUR. UNION AGENCY FOR NETWORK & INFO. SEC., *Privacy and Data Protection by Design Report: From Policy to Engineering* (2014), https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design

some notable exceptions,[11] the academic discourse on privacy also has not focused on regulatory specifics.

There are many possible reasons for this state of affairs, including the fact that the extent and nature of present-day information privacy issues are novel and evolving, and that various and competing values are in play. We believe, however, that at least part of the problem is that the threat of innovation stifling seems to hold particularly strong sway in the debate about information privacy regulation, casting a spell even over constituencies otherwise inclined to a pro-regulatory stance. This Article thus analyzes whether there is any basis to expect that privacy regulations pose a uniquely serious threat to innovation that justifies blanket opposition or requires special treatment.

We ground our analysis in previous work where we developed a framework for understanding the interplay between regulation and innovation.[12] We take as our starting point the classic economic view that regulation should be adopted when it can be reasonably expected to ameliorate market failures at sufficiently low cost, thereby improving social welfare. We combine that view with the standard economic incentive theory of innovation from intellectual property theory, a perspective that has been surprisingly absent from the debate about regulation and innovation.

Our framework is based on the observation that suppliers' incentives to pursue innovations in particular goods and services vary not only based on anticipated market demand for particular innovations, but also in the extent to which suppliers expect to be able to appropriate returns on investment in light of market competition. Thus, market demand and anticipated appropriability jointly influence the market's portfolio of innovative activities. Failures of market demand, such as externalities, collective action problems and information asymmetries, are classic justifications for regulation, which aims to re-align supplier incentives so that the market will produce a more socially desirable portfolio of goods and services. Environmental, consumer protection and

---

(providing an inventory of existing privacy by design strategies with a focus on privacy enhancing technologies).

[11] *E.g.*, Ira Rubinstein & Dennis Hirsch, Better Safe than Sorry: Designing Effective Safe Harbor Programs for Consumer Privacy Legislation, 10 BNA PRIVACY & SECURITY L. REP. 1639 (2011).

[12] Yafit Lev-Aretz & Katherine J. Strandburg, *Regulation and Innovation: Approaching Market Failure from Both Sides*, available at SSRN, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3462522.

health and safety regulations are justified by failures of market demand.

Market failures, however, are also endemic to the supply side, particularly where incentives for innovation are concerned. Goods and services vary in the extent to which suppliers can maintain market exclusivity and thus appropriate supra-competitive returns on investment. Suppliers have incentives to distort the market's portfolio of goods and services away from what the market would otherwise demand by shifting production toward those goods and services that allow them to appropriate higher market returns. These "appropriability failures" can materialize even in the face of demand signals that are perfectly aligned with social preferences. The well-known "free rider problem," which classically justifies intellectual property, is one sort of appropriability failure. Where it occurs, competitor "free riding" precludes innovators from recouping their innovation costs and thus tends to dampen incentives for innovation. Intellectual property law attempts to address this appropriability failure by using market exclusivity to provide innovators with a period of supra-competitive market returns. Competition law addresses different sorts of "appropriability failures" that favor early suppliers of particular goods and services.

The important point here is that, overall, the market's portfolio of goods, services, and innovative activities, is jointly elicited by suppliers' perceptions of market demand, which may be shaped by regulation, and their expectations about appropriability, which may be shaped by intellectual property and competition law. While the regulatory policy literature has addressed market demand failures and the intellectual property and competition law literatures have addressed failures of appropriability, the interplay between market demand failures and appropriability failures has been under-appreciated. As we have argued elsewhere, neglecting these interactions is of little consequence in the many innovation contexts where demand failures and appropriability failures are uncorrelated. In those contexts, designers of substantive regulation need not concern themselves with appropriability failures, which can be addressed independently by IP and competition law.[13] In general, well-designed regulation is likely to shift innovative activity into more socially desirable *directions*, rather than to reduce innovation overall.[14]

[13] *See id.* at 16-21.
[14] *See id.*

Notably, however, in some innovation contexts regulations designed to address demand failures can interact with appropriability failures to jointly affect suppliers' incentives to pursue particular innovative paths. Even in those contexts, however, there is no general stifling effect; depending on regulatory design, the interplay can either enhance or discourage innovation.[15]

This Article applies these insights to information privacy regulation, which is one important arena in which calls for regulation are often met with sweeping assertions that regulation will stifle innovation. We conclude that certain characteristics of personal information-based products and services suggest that some sorts of privacy regulation may interact with appropriability failures to affect innovation. Nonetheless, we find no reason to assume that any such effects on innovation must be "stifling." To the contrary, our analysis suggests that some privacy regulation designs might mitigate failures of appropriability, thereby enhancing innovation. We thus conclude that across-the-board assertions about the stifling effects of information privacy regulation on innovation are simply wrong. Worse, they distract from difficult and important questions of regulatory design. While information privacy regulation can affect the appropriability of innovative activity, those effects can be either dampening or salutary. In other words, well-designed privacy regulation has the potential to improve the extent to which the market produces a socially desirable portfolio of innovations.

In Part II, we define what we mean by information privacy regulation, distinguishing it from broader possible uses of the term. We next provide a similar discussion of our usage of the term "innovation." We then describe the general framework proposed in our previous work,[16] referring readers to that work for details. Finally, we briefly describe the most relevant academic literature on privacy regulation and innovation, distinguishing our approach.

Part III reviews justifications for regulation relating to the commercial collection, flow and use personal information that have been identified in the literature. It categorizes these justifications as versions of market demand signal failure, broadly construed, including: collective action problems in responding to externalities, distortions of individual preferences related to information problems and transaction costs and misalignment with

[15] *See id.* at 5-8.
[16] *See* generally *id*. at 2-4.

social values. It then explains how aggregation effects exacerbate these failures.

Part IV turns to failures of appropriability in markets involving personal information. Taking intellectual property doctrine as the backdrop, we identify three likely sources of appropriability failures affecting personal information-based innovation: failures of trade secrecy's limiting doctrines in the face of data aggregation and network effects. As a net result of these effects, innovative activities exploiting caches of personal information, far from being deterred by fear of free riding, will tend to be over-compensated.

In Part V we pull everything together. The combination of misaligned demand signals and appropriability failures in personal information markets provides good reason to believe that the unregulated market will produce a portfolio of goods, services and innovation undesirably skewed toward collecting and exploiting personal information. This prevalence of market failures justifies regulatory efforts. Moreover, while privacy regulation can interact with appropriability failures to affect innovation incentives, whether that impact is socially positive or negative depends on regulatory design. Part V uses three hypothetical approaches to privacy regulation design to illustrates this point. Part VI briefly concludes.

II. FRAMING THE ANALYSIS OF PRIVACY REGULATION AND INNOVATION

The understandings of "information privacy regulation" and of "innovation" that we bring to bear here play a significant part in our analysis. Thus, sections A and B of this Part seek to make those definitions explicit. Section C summarizes our general analytical framework for the interaction between regulation and innovation, which we have described and supported in detail in previous work.[17] Section D briefly discusses the primary previous literature on privacy regulation and innovation, distinguishing our approach.

A. Information Privacy Regulation as Regulation of Personal Information Flow

Our working definition of information privacy regulation (which we sometimes term simply, "privacy regulation," in the

[17] *See id.*

interests of brevity), draws on Helen Nissenbaum's contextual integrity theory of privacy.[18] We thus assume that privacy is achieved through maintaining personal information flows that are normatively appropriate for the context.[19] The appropriateness of a given flow of personal information depends on many factors, including the context in which the personal information is being used (medical or employment, for example); the actors that participate in the exchange and their relationships; the subject matter of the information and the identity of the person to whom it pertains; and the transmission principles that constrain the flow (such as whether the consent of either party is required).[20] Though Nissenbaum's treatment focuses on information flows, similar contextual factors determine the appropriateness of other potential regulatory targets, such as personal information collection, retention and use. Nissenbaum's theory evaluates the appropriateness of an information practice in terms of context-specific social norms, as well as the values, goals and ends of the specific context and broad moral and political factors implicated by the practice.[21]

Though we adopt Nissenbaum's contextual and catholic perspective on what "information privacy" might entail, we focus on commercial actors' collection, retention, transfer and use of information in digital form pertaining to individuals, rather than, for example, the social exchange of information between peers, whether on or offline. We thus adopt the economic language of social welfare, market demand and supplier incentives for our analysis. Our basic argument does not rely on any particular definition of social welfare or means of assessing regulation's expected social impact, though these questions would presumably be part of the debate over particular proposals. We merely assume that some such assessment mechanisms, however imperfect or contested, are useful for policy debate, as they are in other regulatory contexts.[22]

---

[18] Helen Nissenbaum, PRIVACY IN CONTEXT (2009).
[19] *Id.* at 127.
[20] *Id*.
[21] *Id.* at 181-82.
[22] *See, e.g.*, Matthew D. Adler & Eric A. Posner, *Rethinking Cost-Benefit Analysis*, 109 YALE L.J. 165 (1999) (criticizing the efficacy of cost-benefit analysis in certain regulatory structures while defending it in others); Eric A. Posner, *Controlling Agencies with Cost-Benefit Analysis: A Positive Political Theory Perspective*, 68 U. CHI. L. REV. 1137 (2001) (providing theoretical justifications for cost-benefit analyses' guiding regulations).

Nissenbaum's framework, unlike traditional accounts of privacy, does not use notions of secrecy or control as its benchmarks for information privacy. Thus, it suggests, and we agree, that information privacy regulation can be used not simply to *restrict* access to personal information but also to *facilitate* access to personal information, to the extent that the increased information flow overcomes failures in market demand. This is, in fact, a classic argument for attorney-client and doctor-patient confidentiality rules. In a more modern debate, some scholars have claimed that individuals have an ethical obligation to share their health information for research purposes.[23] If the sharing of health information reflects a sufficiently strong social value, but voluntary contribution is plagued by collective action problems, then privacy regulation, rather than limiting access to such information, might mandate access to it for research purposes, subject to constraints on its responsible use and flow. Or, rather than mandating disclosure by everyone, privacy regulation might encourage the disclosure of medical information by forbidding researchers from re-purposing it, disclosing it outside of the research context and so forth.[24]

Equally importantly, Nissenbaum's framework does not rely on any contextual categorization of particular types of information as "sensitive" or "private." It thus allows for the possibility, now central to information privacy concerns, that information can be aggregated to make inferences about individuals.

The take-away point is that when we speak of privacy regulation in this Article, we have in mind a very broad menu of mechanisms for constraining and redirecting the collection, retention, flow and use of personal information, rather than assuming that privacy regulation necessarily must restrict collection, provide notice and an opportunity for consent or

---

[23] John Harris, *Scientific Research is a Moral Duty*, 31 J. MED. ETHICS 242 (2005); Rosamond Rhodes, *In Defense of the Duty to Participate in Biomedical Research*, 8 AM. J. BIOETHICS 37 (2008); G. Owen Schaefer, Ezekiel J. Emanuel & Alan Wertheimer, *The Obligation to Participate in Biomedical Research*, 302 JAMA 67 (2009); Joanna Stjernschantz Forsberg, Mats G. Hansson & Stefan Eriksson, *Why Participating in (Certain) Scientific Research Is a Moral Duty*, 40 J. MED. ETHICS 325 (2014); Angela Ballantyne & G. Owen Schaefer, *Consent and the Ethical Duty to Participate in Health Data Research*, 44 J. MED. ETHICS 6 (2018).

[24] *Cf.* 45 C.F.R. § 164.502 (2013) (forbidding the disclosure of certain types of information relating to personal health except in enumerated extraordinary circumstances).

implement some form of the Fair Information Practice principles.[25] All of these mechanisms, and others, should be on the table when privacy regulations are designed.

B. How Businesses Collect Personal Information

We set the stage by describing just some of the ways in which businesses collect personal information. The sort of data collection that concerns us here most often occurs as a by-product of providing some sort of service.[26] Sometimes the service is a means for communication (such as, phone service or a social media platform) with other individuals. The business gains access to any personal information that is communicated as a side effect. Sometimes information is disclosed intentionally to a service provider for one purpose (e.g. "This is what I want to buy; here's my credit card information; please send it to me at this address."), but after collecting it, the business repurposes it for additional uses. Sometimes the information is in some sense created by the service provider by observing and analyzing the individual's behavior (e.g. webpages visited, times of heavy energy use, channels watched on the smart TV). More and more data is now collected by the Internet of Things and smart devices, including TVs, cellphones, electric meters, Alexas, Roombas, security systems, and even toys. The information is often merged from different sources and can be used to infer information that individuals never intended to share, such as sexual orientation or political views.[27]

Uses of the collected data vary widely and can include: (i) targeting advertising or otherwise influencing user preferences or choices; (ii) divining consumer preferences, in the sense of market research or testing; (iii) making decisions affecting consumers in arenas such as employment and insurance; (iv) creating a product or service using the data as a production tool, much as an assembly line is used to produce toys (here, we have in mind, e.g. Google's use of search data to suggest search terms or the use of personal information to train a machine learning model for targeting advertising, predicting credit risk, and so forth); (v) creating a product or service that incorporates the data, much as plastic is incorporated into a toy (e.g., an advertiser is offered placements on

---

[25] INT'L ASS'N OF PRIVACY PROFS, *Fair Information Practice Principles*, https://iapp.org/resources/article/fair-information-practices.

[26] At least, data collection is a by-product from the perspective of the individual. It is quite often the primary purpose from the business's perspective.

[27] Michal Kosinski, David Stillwell & Thore Graepel, *Private Traits and Attributes Are Predictable from Digital Records of Human Behavior*, 110 PROC. NAT'L ACAD. SCI. 5802 (2013).

particular individuals' Facebook pages based on an analysis of consumer data); or (vi) aggregating the data and selling it as a product to others. Some might argue that consumer data can also be used as a tool for innovation, much like a traditional research tool. A company might, for example, use consumer data while experimenting with innovative approaches to particular tasks, such as recommender systems, or while designing a new approach to machine learning. Businesses also share data with the government, sometimes volunteering it, sometimes selling it and sometimes turning it over only in response to legal processes.[28] More and more often, data is aggregated, both across a given individual's activities and across many individuals to be used in some sort of predictive modeling.[29] Data aggregation is the source of privacy concerns like identity theft, embarrassment, fraud, sexual harassment and stalking. Price discrimination is another potential issue,[30] along with other forms of discrimination and bias[31] (intentional and unintentional) based on factors such as race, religion, gender, genetics, health status, political views, arrest history and economic class, one's friends' characteristics or behavior, or particular lapses in judgment or "sins of the past."

C. Market Innovation in Personal Information-Based Products and Services

Like privacy, innovation has been defined in various ways for various purposes. Because our focus is on the effects of regulations that apply to commercial actors, we concentrate on innovation induced by market forces. Here we are concerned with innovation in personal information-based products and services (or PI-based products and services), as explained in further detail later in this section.

---

[28] *See, e.g.*, Wendy Everette, The FBI Has Not Been Here (Watch Very Closely for the Removal of This Sign): Warrant Canaries and First Amendment Protection for Compelled Speech, 23 GEO. MASON L. REV. 377, 384 (2016) (explaining how Apple generally requires the government to initiate legal processes to obtain user data and uses "warrant canaries" to inform the public of when this has happened); Developments in the Law—More Data, More Problems, 131 HARV. L. REV. 1715, 1729-36 (2018) (discussing ways in which corporations may resist or comply with government data requests depending on their incentives).

[29] *See* Andrew Kasabian, Note and Comment, Litigating in the 21st Century: Amending Challenges for Cause in Light of Big Data, 43 PEPP. L. REV. 173, 192-93 (2015).

[30] Douglas M. Kochelek, Note, *Data Mining and Antitrust*, 22 HARV. J. LAW & TECH. 515 (2009).

[31] Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671 (2016).

a. Innovation

While "innovation" always relates to introducing something new, some usages impose some threshold of difference (or improvement) compared to what came before.[32] We impose no such threshold; our usage here encompasses any novel aspect of goods, services or their production that provides a competitive advantage. Usages in the literature also differ by scope, with some narrowly referring only to "technological" innovation and others expansively encompassing new approaches to business, the arts, marketing (or perhaps even regulation).[33] Here, we will generally have in mind the sorts of technological or expressive outputs that are the subject matter of patent and copyright protections, though not only those that qualify for those protections. Usages also vary in the extent to which they require market entry rather than mere "invention." [34] Because we are concerned with the ex-ante assessment of regulatory impact, our focus here is on innovative activity and investment that aims for market entry.

We also assume that society benefits most when many innovators can enter the market and build upon each other's work in competitive fashion. Intellectual property doctrine accords with this view. It is designed to limit exclusive rights to what is necessary to induce invention, primarily for fear of suppressing downstream and follow-on invention. The academic debate over whether monopoly or competition best fosters innovation is longstanding, however, with heavy hitters like Schumpeter and Arrow famously taking different sides. [35] The well-known "prospect theory" of patents also argues that broad and deep

---

[32] *See* Jon-Arild Johannessen, Bjørn Olsen & G.T. Lumpkin, *Innovation as Newness: What Is New, How New, and New to Whom?,* 4 EURO. J. OF INNOVATION MGMT. 20 (2001)

[33] *See, e.g.*, Sulieman Ibraheem Shelash Al-Hawary & Faraj Mazyed Faraj Aldaihani, *Customer Relationship Management and Innovation Capabilities of Kuwait Airways*, 5 INT'L J. OF ACAD. RESEARCH IN ECON. AND MGMT. SERV. 201 (2016); Anna Butenko & Pierre LaRouche, *Regulation for Innovativeness or Regulation of Innovation?*, 7 L., INNOVATION, AND TECH. 52 (2015).

[34] *See, e.g.*, Wendy Seltzer, *Software Patents and/or Software Development*, 78 BROOKLYN L. REV. 929, 971 (2013) ("Economists and management scholars distinguish between invention and innovation. Invention is the spark of an idea, while 'innovation means invention implemented and taken to market.'") (internal citation omitted).

[35] *See* JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM, AND DEMOCRACY 106 (3d ed. 1950); Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, *in* ESSAYS IN THE THEORY OF RISK-BEARING 157 (Julius Margolis ed., 1971)

exclusive rights will foster innovation.[36] Nonetheless, the majority view among intellectual property scholars aligns with Arrow in favor of competition and that is the position we adopt here.

b. Innovation in Personal Information-Based Goods and Services

Because information privacy regulations affect the collection, transfer and use of personal information, in the form of digital data, by commercial actors, we focus on innovation in products and services for which personal information is an important contributor to market value. "Innovation" in PI-based goods and services does not, however, include improvements that result merely from employing "more" personal information in a known way, even if they increase market value. Innovation, in our usage, must involve something new about the way personal information is used or some other aspect of the good or service.[37] We believe this distinction is consistent with popular conceptions of innovation and usage in the intellectual property and innovation policy literatures.

PI-based products and services ordinarily combine algorithms, software code, and user interfaces with personal information in one of several ways. Personal information might be employed as a tool to develop novel PI-based products and services. For example, a company might use its database of personal information to experiment with different ideas about how to provide purchase recommendations to consumers until it comes up with a novel approach. Personal information might also be used as an input to produce a product or service. For example, having come up with a new approach to providing recommendations, the company might use personal information in training a machine-learning-based model implementing the approach. Finally, personal information might be used to deliver the product or service to the customer. For example, the company might deliver a recommendation to a consumer by inputting her personal information into the model and displaying the output

---

[36] *See* Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J L & ECON 265, 265-66, 268 (1977) (positing the "prospect theory" of patents, which holds that a scheme of broad and deep exclusive rights for patents is socially beneficial and functions similarly to property rights over oil, gas, and minerals by incentivizing inventors to maximize the value of the innovation covered by the patent).

[37] We recognize that this line is not always bright. For example, the availability of "more" personal information can spark or facilitate novel uses of that information. We nonetheless think it is a conceptually meaningful distinction.

recommendation. The first sort of use qualifies as innovation in our usage while the third does not. The second—using personal information to train a machine-learning model—is a harder call, but we take the position that routinely applying well-known machine learning methods to new data is not innovation, even though it produces new models. What we mean by innovation would, however, include the development of a novel machine-learning algorithm to produce the model.

D. Regulation, Innovation and Market Incentives

Our framework for analyzing the relationship between regulation and innovation, which we present in detail elsewhere,[38] distinguishes two forces that shape market incentives for innovation: the demand signals perceived by potential innovators and the extent to which those suppliers expect to appropriate market returns from particular innovative avenues. Substantive regulation traditionally addresses the former, attempting to realign demand signals with individual and social preferences. Intellectual property doctrine and, less explicitly, competition law target the latter by attempting to redress appropriability failures that can skew suppliers' incentives away from the portfolio of innovative activity that is most responsive to market demand because they expect to appropriate excess or insufficient returns from particular innovative activities.

a. Misaligned Demand Signals and Appropriability Failures

Misalignment between market demand signals and social welfare may originate in: (i) externalities and associated collective action problems; (ii) failures to accurately express individual preferences because of information asymmetries, non-rational behaviors and transaction costs; and (iii) the market's failure to reflect social values related to distributive concerns, treatment of minorities, and ethical norms.[39] Though we believe all of these sorts of failures sometimes justify regulation, our analytical framework is agnostic about the sources of demand misalignment that are considered. Because regulation aims to realign market demand, it will often shift the market's portfolio of innovative activity. Indeed, driving innovation toward socially promising directions not induced by unregulated market demand—and away

[38] *See generally* Lev-Aretz & Strandburg, *supra* note 12.

[39] These categories are not entirely distinct and we do not claim they are comprehensive. They are simply illustrative.

from socially deleterious or wasteful paths—is often a primary aim of regulatory design.[40]

Even if demand signals are perfectly aligned, markets will fail to produce socially optimal portfolios of goods and services if there are failures of appropriability, because suppliers can appropriate greater market returns from some goods and services than from others. When suppliers shift production toward goods and services with higher appropriability and away from those with lower appropriability, they distort the market's portfolio of goods and services. As a general rule, markets rely on competition to keep the appropriability landscape level, but innovators can run into free rider problems, a form of appropriability failure, even in competitive markets. Innovators generally make upfront investments, but competitors can often free ride by cheaply copying and then undersell the innovator. If innovators anticipate this problem, they may be deterred from investing in innovative activities.

IP aims to level out the appropriability landscape by granting innovators exclusive rights, which allow them to charge supra-competitive prices and recoup their upfront investments during the term of protection.[41] The IP solution, however, has two important weaknesses. First, IP compensates only roughly for free rider issues, inevitably leaving the appropriability landscape scattered with hillocks of over-compensation and hollows of under-compensation. Second, IP doctrine primarily mitigates appropriability failures associated with competitor free riding, leaving other sources of appropriability failures unaddressed.[42] Barriers to entry, which favor early entrants by imposing higher upfront costs on later market entrants, are one such source of appropriability failure. As we will discuss in detail later, markets

---

[40] Suzanne Scotchmer, *Cap-and-Trade, Emissions Taxes, and Innovation*, 11 INNOVATION POL'Y & ECON. 29, 49 (2011) (concluding that "[a]ny regulatory policy that imposes financial burdens for emitting carbon also creates an incentive to invest in carbon-reducing technologies").

[41] Obviously, IP exclusivity also generates social costs if the length and breadth of the exclusive rights are not tailored to innovators' upfront investment, and to the extent the exclusivity overly restricts follow-on innovation. Various limiting doctrines, such as patent law's non-obviousness requirement and copyright's fair use are aimed at minimizing instances of unnecessary exclusivity. *See, e.g.*, *id.* at 45.

[42] In some circumstances, antitrust or competition law is used to reduce barriers to entry. When that is the case, these appropriability distortions will be modified and (one hopes) reduced.

involving personal information are particularly likely to exhibit certain naturally arising barriers to entry.[43]

b. Implications of Correlations between Misaligned Demand Signals and Appropriability Failures

Regulation that realigns demand signals necessarily affects the way demand interacts with the appropriability landscape to determine the resulting portfolio of innovation. IP doctrine and competition laws are intended to balance the costs and benefits of smoothing out the appropriability terrain. In most contexts, these doctrines will be similarly effective (or flawed) in remedying appropriability failures for the portfolios of innovative activities induced by regulated and unregulated demand. When that is the case, regulatory design can proceed without taking failures of appropriability into account.[44] In some contexts, however, regulatory demand shifts may turn out to be correlated significantly and systematically with appropriability failures, undercutting this separability assumption. At the extreme, even if a regulation perfectly realigns demand signals with social preferences, it could combine with an unfavorable appropriability terrain to induce a *worse* portfolio of innovation from a social perspective.

For example, traditional network effects, such as those associated with telephone networks, simultaneously create value for consumers and barriers to entry that deter competitive or follow-on innovation and can lock consumers into less preferable technologies.[45] Breaking up networks, which could work to mitigate appropriability failures in such contexts, would simultaneously reduce the network's value to consumers. Such negative correlation resulting from the tradeoffs between

---

[43] *See infra* Part IV.

[44] *See id.*

[45] *See* SEAN HOWELL, BIG DATA AND MONOPOLIZATION 4 (2018) ("[D]ata-driven markets tend to feature strong network effects and economies of scale, which create barriers to entry that other firms may have a hard time overcoming."); JENS PRUFER & CHRISTOPH SCHOTTMULLER, COMPETING WITH BIG DATA 2, 15 (2017); Joseph Farrell, *Coordination and Lock-In: Competition with Switching Costs and Network Effects*, 3 HANDBOOK INDUS. ORG. 1967, 2034 (2007) ("[E]arly choices are powerful, able either to help coordination or to wield disproportionate influence. Thus any early lead in adoptions (whether strategic or accidental) will tend to expand rather than to dissipate. Network markets are "tippy": early instability and later lock-in.").

appropriability and demand should be accounted for in regulatory design.

Correlation between demand and appropriability can also produce socially desirable side effects, however. For example, a likely result of an environmental regulation encouraging the use of alternative energy sources is increasing market demand for solar panels and decreasing demand for electricity produced by coal plants. If the power plant market is susceptible to barriers to entry, but the solar panels market is not, electricity markets might become somewhat less infected with appropriability failures once the regulation is in place.

Another example from the broadcast media context illustrates how correlations between demand and appropriability can produce side effects. The advertising-supported business model was created at least partly in response to gaps in intellectual property law's response to customer free rider problems. While the free rider problem that intellectual property is traditionally equipped to mitigate is that of free riding *by competitors*, in some contexts, innovators face the risk of free riding *by consumers*. Consumers free ride when they can take advantage of an innovation without paying for it because practicalities make it difficult or impossible for producers to identify users and demand payment. Television and radio confronted just this problem, since it was essentially impossible to monitor and demand payment for consumption of content once it was broadcast into the ether. The advertising-supported business model solved some of those free rider problems, but simultaneously created misalignment between market demand for programming content and social preferences. A regulation that realigns broadcast programming content more closely with individual and social preferences by creating barriers to advertiser influence on content would result in both the intended demand re-alignment and appropriability side effects. For example, if the unregulated advertising-based business model tended to over-compensate content creators for their upfront costs, the regulation might simultaneously improve the market's satisfaction of consumer and social preferences and mitigate appropriability failures—a win-win result. But a regulation that reduces advertising revenue to the point of under-compensating investments in socially desirable content creation would be an overall social loss.

c. Regulation and Innovation "Stifling"

In our previous article, we rebutted the contention that regulation will generally "stifle" innovation in some detail.[46] Thus, we reprise that discussion only briefly here. The complaint that regulation will stifle innovation sometimes means only that it will effectively shift demand by raising costs along some particular innovative path (often the path that the complainant is currently following). This sort of shift is an expected—and intended—effect that does not stifle innovation in any socially relevant sense.

The contention that regulation will stifle innovation might also mean that regulation will decrease the total "amount" of innovative activity. Quantifying the amount of existing innovation, not to mention isolating the amount of innovation that is adversely affected by regulation, is notoriously challenging, both conceptually and empirically.[47] But even if a regulation makes innovative activity in a regulated sector less attractive, a sweeping reduction in innovation *overall* seems unlikely, as investments can and will be shifted from one sector to another. Moreover, the extent to which regulation reduces investment in innovation in the regulated sector is a matter of regulatory design. A well-designed regulation would minimize regulatory compliance and other transaction costs as much as possible and take them into account in deciding whether the regulation is justified. Compliance costs could even be spread to taxpayers, rather than borne by the regulated sector.[48] As we elaborate elsewhere,[49] even compliance costs borne by the regulated sector are only likely to depress innovative investment when they are i) imposed on innovators in proportion to their innovative activity, ii) not avoidable by compliance innovation; and iii) not recoupable via first mover advantages or intellectual property.

---

[46] *See* Lev-Aretz & Strandburg, *supra* note 12.

[47] Scholarly attempts to empirically test the relationship between regulation and innovation include Avi Goldfarb & Catherine Tucker, *Privacy and Innovation*, 12 INNOVATION POL'Y & ECON. 65, 84-86 (2012); Nathan Goldschlag & Alex Tabarrok, *Is Regulation to Blame for the Decline in American Entrepreneurship?*, 33 ECON. POL'Y 5 (2018); Joseph M. Crabb & Daniel K.N. Johnson, *Fueling Innovation: The Impact of Oil Prices and CAFÉ Standards on Energy-Efficient Automotive Technology*, 31 ENERGY J. 199 (2010); and Shunsuke Managi et al., *Environmental Regulations and Technological Change in the Offshore Oil and Gas Industry: Rethinking the Porter Hypothesis*, 81 LAND ECON. 303 (2005). None, to date, has conclusively showed that regulation has generated innovation-stifling effects.

[48] Lev-Aretz & Strandburg, *supra* note 12.

[49] *Id.* at 21-25.

Finally, claims that shifting society's innovation portfolio through regulation will stifle innovation may reflect concerns about innovation's path-dependence, unpredictability and cumulative nature. The risk that regulation will unwittingly suppress high social value innovation in the long run must always be weighed against the long-term social costs of the misalignment of unregulated demand with true individual and social preferences. [50] Moreover, the long term unpredictability of innovative activity is a feature of both the regulated and unregulated portfolios, so who is to say which will turn out best? Indeed, Liskow & Karpilov have argued that the risk of innovation failure is itself path dependent, increasing the long-term benefit of regulation that realigns innovative activity with social preferences.[51]

E. Prior Literature

The academic literature on the interplay between regulation and innovation mostly centers around a few contexts, with a special focus on environmental regulation. Studies are mostly empirical, and outcomes are affected by the chosen methodologies and metrics for innovation. Richard Stewart's seminal 1981 article offers one of the few theoretical treatments.[52] Stewart distinguished "market innovation," which "create[s] benefits that firms can capture through the sale of goods and services in the market"[53] from "social innovation,"[54] which "create[s] social benefits, such as cleaner air, that firms cannot directly capture through market sales."[55] Stewart argued, in brief that regulations may adversely affect "market innovation" in various ways,[56] while "government, rather than the market, ordinarily must provide incentives for regulated firms to undertake investment necessary to generate social innovation."[57] He therefore argued that regulation should focus on incentivizing social innovation and critiqued command-and-control approaches to regulation for failing to succeed in doing

---

[50] *Id.*
[51] *Id*. at 18.
[52] Richard B. Stewart, *Regulation, Innovation, and Administrative Law: A Conceptual Framework*, 69 CALIF. L. REV. 1256 (1981).
[53] *Id*. at 1279.
[54] *Id*.
[55] *Id*.
[56] *Id.* at 1279-80.
[57] *Id.* at 1281.

that.[58] The market/social innovation framework has been adopted by a number of later scholars.[59]

Tal Zarsky's 2015 paper,[60] which is the most relevant previous work addressing privacy regulation and innovation, adopts Stewart's distinction between market and social innovation in assessing privacy's implications for "social" and "market."[61] While some of his discussion is irrelevant to our focus on privacy *regulation*, the Article raises a number of important issues that we pick up and discuss here, most notably the possible links between privacy regulation and lower market barriers.[62] Like us, Zarsky emphasizes the need for "in-depth policy discussions, examining what form of regulations must be introduced to enhance privacy, and what forms of innovation society is striving to achieve."[63] Zarsky's bottom-line assessment of privacy regulation is more skeptical than ours, however. To some extent, that greater skepticism appears to reflect implicit assumptions about the form of privacy regulation, and its similarity to consent-based and FIPs models, which we do not share. We also disagree at times with Zarsky's reading of the limited empirical evidence about the impact of privacy regulations on innovation. Most importantly, however, our treatment differs because we do not adopt Stewart's distinction between market and social innovation.

As we explain in detail elsewhere, attempting to distinguish between market and social innovations is a doomed enterprise because, as Stewart explicitly recognized, "[a] given innovation may confer both market and social benefits."[64] Indeed, the assumption that innovations produced by firms in response to market demand will create some benefits captured by firms *and* some benefits that spill over to society underlies all efforts to incentivize market innovation. Furthermore, the classification of innovations as social because they create "social benefits, such as cleaner air, that firms cannot directly capture through market sales"[65] is inadequate to analyze the impact of regulation on innovation.

---

[58] *Id*. at 1281-82.
[59] Tal Zarsky, *The Privacy-Innovation Conundrum*, 19 LEWIS & CLARK L. REV. 115 (2015).
[60] *Id*.
[61] *Id*. at 126
[62] *Id.* at 135-6.
[63] *Id.* at 166.
[64] Lev-Aretz & Stradburg, *supra* note 12, at 6, citing Stewart, *supra* note 53, at 1279.
[65] *Id.*

Regulations aim to correct many types of market failures, involving not only various sorts of externalities, but also information asymmetries and other problems that misalign market demand with individual and social preferences. Indeed, the goal of some regulation is to re-direct effort away from certain types of socially negative forms of innovation, rather than to create "social" innovation. Finally, the definition of social innovation in terms of benefits that firms cannot capture through market sales blurs the distinction that we think is most useful for the analysis—between market failures that result from faulty demand signals and those that result from failures of appropriability related to competition among suppliers.[66]

III. PERSONAL INFORMATION AND MISALIGNED MARKET DEMAND SIGNALS

Some who argue that privacy regulation will "stifle" innovation may simply be skeptical that there really are significant market failures regarding the flow and use of personal information. If these skeptics are correct, there is no justification for regulation in the first instance.[67] In this view, the so-called "privacy paradox" between consumers' self-reported positions valuing privacy and their real-world acceptance of the information disclosure practices of the products and services they use reflects the weaknesses of

[66] This distinction might seem artificial, but it is not because markets solve a number of different informational and coordination problems, including not only "What do consumers want?" but also "Who will supply it?" The second question is particularly important for innovation, since it is difficult to predict in advance who will do the best job. There are, of course, non-market mechanisms for handling this issue (peer-reviewed grant funding, for example), but one reason for relying on intellectual property is that inventors identify themselves through their activities and are rewarded after the fact. Thus, while overcoming the demand-side collective action problem to collect the funds needed to pay for the clean air technology allows consumers to signal their demand, it does not tell them who should get the money. Relying on the competitive market to answer that question brings appropriability questions into play.

[67] Some early economic accounts of privacy took a decidedly hostile view of privacy, based on the idea that efficient markets depend on the free flow of information and that privacy was essentially a mechanism by which individuals could engage in rent-seeking based on asymmetric information. *See, e.g.*, Richard Posner, *The Right of Privacy*, 12 GA. L. REV. 393 (1978) (applying the Coase theorem); Ronald Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 16 (1960). Note, however, that the Coase theorem applies only in a world devoid of transaction costs, an assumption that Coase himself did not take to be true.

survey evidence,[68] rather than any real market failure.[69] Perhaps, when the rubber hits the road, consumers are simply unwilling to pay the costs of exercising a preference for non-disclosure.[70] Opponents of privacy regulation also argue that, when consumers care enough about privacy to take action, companies respond by changing their privacy policies and opting in to other self-regulation.[71] Businesses may also choose to limit their collection practices and keep consumer data secure in order to avoid negative publicity or to distinguish themselves from their competitors by their strict data policies.[72]

Clearly, we disagree with this sanguine view of the market's treatment of personal information. We do not attempt a fulsome exposition of this debate here, but simply summarize some of the sources of misaligned demand signals regarding personal information flow that have been identified in previous literature.[73] One need not accept all of these arguments, or agree with our broad understanding of market failure, to conclude that there is a prima facie basis for some kind of information privacy regulation.

Reasons to anticipate significant misalignment between market demand signals and socially preferable personal information flow fall into three general categories: i) externalities and associated collective action problems; ii) failures to accurately express individual preferences because of information asymmetries, non-rational behaviors and transaction costs; and iii) misalignment with social values related to distributive concerns, treatment of minorities and ethical norms.[74] In addition, personal information markets gives rise to a fourth category of what we call

---

[68] *See, e.g.*, Caleb Fuller, *How Consumers Value Digital Privacy: New Survey Evidence*, GEO. MASON U.: PROGRAM ON ECON. & PRIVACY (Feb. 20, 2018), https://pep.gmu.edu/wp-content/uploads/sites/28/2018/02/Fuller_How-Consumers-Value-Digital-Privacy.pdf.

[69] *See id*.; *see also* Robert W. Hahn & Anne Layne-Farrar, *Is More Government Regulation Needed to Promote E-Commerce*?, 35 CONN. L. REV. 195, 213 (2002) (noting in the context of online shopping that "it is not even clear that any e-commerce has been deterred. Absent evidence of a significant market failure, the case for further government intervention is weak at best").

[70] *See* Fuller*, supra* note 68; *see also* Solveig Singleton*, Privacy Versus the First Amendment: A Skeptical Approach*, 11 FORDHAM INTELL. PROP. MEDIA & ENT. 97, 147 (2000).

[71] *See* Fuller*, supra* note 68.

[72] James P. Nehf, *Recognizing the Societal Value in Information Privacy*, 78 WASH L. REV. 1, 59 (2003).

[73] These categories are used for expositive clarity, though we are aware that they often overlap.

[74] These categories are not entirely distinct and we do not claim they are comprehensive. They are simply illustrative.

"aggregation failures" that combines aspects of the others. As our discussion below shows, a well-functioning market in which consumers pay for goods and services with personal information is unlikely.[75]

A. Collective Action Problems in Responding to Externalities

Information collection and use practices have been long blamed for imposing negative externalities on the subjects of the information.[76] In this section, we confine our discussion of externalities to effects on individuals that arise from *other people's* use of personal information-collecting goods and services.[77] Such externalities are likely because personal information, by its nature, often pertains to more than one individual. Today, in addition, predictive analytics models that have important consequences for individuals might be significantly derived from other people's data.[78] Sometimes, these predictive models amount explicitly to characteristic—if not guilt—by association, making inferences about individuals based on their family members or friends. The Internet of Things introduces another source of externalities when information is collected about the behavior of individuals who do not own the smart thing, but nevertheless interact with it (perhaps

---

[75] *See* Katherine J. Strandburg, *Free Fall: The Online Market's Consumer Preference Disconnect*, 2013 U. CHI. LEGAL F. 95, 107; *see also* Kenneth C. Laudon, Markets and Privacy, 39 COMM. ACM 92, 97 (1996) ("In a perfect world characterized by perfect information . . . [i]n this most felicitous world of 19th-century economic thought, symmetry of information among market participants—capitalists, laborers, and consumers—is the lubricant of social and economic progress.").

[76] *Laudon, supra* note 75, at 93 (arguing that the massive scope of information collection at the time had significant costs and the businesses' use of the information was wasteful and inefficient); Robert W. Hahn & Anne Layne-Farrar, *The Benefits and Costs of Online Privacy Legislation*, AEI-BROOKINGS JOINT CTR. FOR REG. STUD., Working Paper No. 01-14 (2001), https://papers.ssrn. com/sol3/papers.cfm?abstract_id=292649; Nehf, *supra* note 72, at 79-80; Pamela Samuelson, *Privacy as Intellectual Property*?, 52 STAN. L. REV. 1125, 1132-33 (2000); PETER SWIRE & ROBERT LITAN, NONE OF YOUR BUSINESS 7-8 (Brookings Institution Press 1998).

[77] A looser understanding of externalities might also encompass effects that arise from flows of personal information that are not taken account of or anticipated by individuals in their own transactions with commercial entities. We discuss those issues in the next section relating to the market's failure to accurately reflect individual preferences.

[78] *See, e.g.*, Mikella Hurley & Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE J. L. & TECH. 148, 151 (2016) (examining harms resulting from "creditworthiness-by-association," in which loan applicants are scored at least partially by their social connections' data).

even unwittingly).[79] Businesses that rely on collecting and using personal information benefit from these externalities and are insufficiently incentivized to account for them in their behavior: "just as factories have no reason to refrain from filling the air with pollutants, these companies will not hesitate to collect, use, and flood the market with detailed, personal information."[80]

The emerging signaling economy creates another source of externalities. When consumers are given a menu of options about whether and what personal information to disclose, their choices can come to act as signals, especially in information asymmetry ecosystems.[81] If most individuals choose to disclose information except when they have "something to hide," businesses may begin to assume the worst about those consumers who simply have a taste for privacy or have other justifications for preferring nondisclosure (to avoid discrimination, perhaps, or to avoid leaking information to potential abusers or stalkers).[82] As a result, individuals who refuse to disclose might face new forms of economic discrimination.[83] In this sort of scenario, consumers would face a Hobson's choice between the risks of disclosure and the economic penalties of non-disclosure.

Perhaps the most important source of negative externalities for consumers is the prevalence of the advertising-based business model among personal information-based companies, now being partly supplemented by a predictive analytics-based business model. In a traditional two-sided market, intermediaries (oftentimes referred to as "two-sided platforms") concurrently respond to the preferences of two groups, lowering the cost of

---

[79] *See, e.g.*, Allison S. Bohm et al., *Privacy and Liberty in an Always-On, Always-Listening World*, 19 COLUM. SCI. & TECH. L. REV. 1, 7 (2017), http://www.stlr.org/cite.cgi?volume=19&article=Bohm.

[80] Dennis D. Hirsch, *Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law*, 61 GA. L. REV. 1, 28-29 (2006). Hirsch continues to describe the failure in terms of tragedy of the commons, where collection-driven businesses "receive all the benefits of their use of personal information but share the cost (in terms of the erosion of trust) with all others who depend on individuals to provide personal information on the Web." *See also* A. Michael Froomkin, *Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements*, 2015 U. ILL. L. REV. 1713, 1729 (2015) [*hereinafter* "*Regulating Mass Surveillance*"].

[81] Scott R. Peppet, *Unraveling Privacy: The Personal Prospectus and the Threat of a Full-Disclosure Future*, 105 NW. U. L. REV. 1153, 1156 (2011).

[82] *Id.* at 1176.

[83] *Id.* For a real-life example of such effect, see Nizan Geslevich-Packin & Yafit Lev-Aretz, *On Social Credit and The Right to Be Unnetworked*, 2016 COLUM. BUS. L. REV. 339, 372-82.

transactions between them.[84] As one of us has demonstrated elsewhere, while the dominant advertising-supported business model shares some features with a typical two-sided market intermediary, the majority of these businesses tie advertising from a group of sellers with an "associated good" such as relevant content, search results, or social networking services.[85] And because, more often than not, the consumer is interested in the associated good, and not in the advertising, the ad-supported business model does not act simply to reduce transaction costs for the two parties, but also leads to a distorted demand signal that reflects only preferences for the "bundle", rather than disaggregated preferences for the advertising and the associated good.[86]

Essentially, consumer preferences regarding the features of these goods and services—including their personal information practices—are filtered through the preferences of advertisers. The advertising-based business model distorts demand signals because there is an unavoidable mismatch between unfiltered consumer preferences and advertisers' goals.[87] Certainly, advertisers seek to attract eyeballs, and thus are somewhat sensitive to consumer preferences. Ultimately, however, advertisers strive to bolster sales of their own products and will support only those creative efforts that further that goal. Perhaps because advertising-supported products have been viewed as essentially free to consumers, critiques of advertising-supported media have tended to focus on (undoubtedly important) concerns with fairness and bias, rather than on the more mundane failure of advertising-supported products to accurately reflect consumer preferences *for those products*. When companies offer targeted advertising based on personal information, demand signals are likely to be skewed toward products and services that are optimized to collect personal information.

---

[84] *See* Strandburg, *supra* note 75, at 113.
[85] *See id*.
[86] *See id.* at 113-116. For a different example of two-sided market in the context of information privacy, see Hal R. Varian, *Economic Aspects of Personal Privacy* (Dec. 6, 1996) (unpublished manuscript), http://people.ischool.berkeley.edu/~hal/Papers/privacy (discussing a sale of a mailing list as a form of externality, noting that "[e]ven though the first two parties in the transaction—the individual who may want to buy something, and the seller who may want to sell him something—have incentives that are more or less aligned, the transaction between the original owner of the mailing list and those to whom it is sold do not have such well-aligned incentives").
[87] Strandburg, *supra* note 75.

These negative externalities are associated with collective action problems. Collective action problems arise for the usual sorts of reasons that have been well explored in other contexts.[88] These reasons include the high transaction costs of coordination between strangers, who are not directly involved in a transaction, and particularly high information costs experienced by such outsiders. A critical mass problem exists whenever consumers mobilize for the sake of greater privacy protection.[89] In the privacy arena, collective action problems are exacerbated by the distortions of individual preferences discussed in the next section.

Collective action problems associated with network effects can also prevent consumers from expressing their preferences regarding platform technology and features. Network effects arise from positive externalities, in which each consumers' utility from using a certain good or service increases as additional consumers use the same good or service.[90] Goods exhibiting network effects depart from rules of supply and demand where it is the shortage of a good that increases its value.[91] The more common the use of a good with network effects, the higher its value.

Network effects are common in communications industries. Many personal information-based businesses, including social networks, such as Facebook and Twitter, demonstrate clear network effects, given that the appeal of a social network depends on the number of existing users.[92] While network effects are due to positive externalities between users, they can generate market failures in which consumers become "stuck" in sub-optimal equilibria regarding the technology or features of goods and services. In principle, users can transfer the positive externalities of network effects from one supplier to another simply by moving their business. Consumers' ability to do this in practice, however, is undermined by a classic collective action problem: to maintain the positive externalities produced by network effects when they

---

[88] Mancur Olson argued that large groups sometimes induce a powerful incentive for members to free ride on the efforts of the others, because each member can only make a small contribution to the whole. MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS 16 (2d prtg. 1971).

[89] Paul M. Schwartz, *Property, Privacy, and Personal Data*, 117 HARV. L. REV. 2055, 2079 (2004).

[90] Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CALIF. L. REV. 479, 483 (1998)

[91] DAVID EASLEY & JON KLEINBERG, NETWORKS, CROWDS, AND MARKETS: REASONING ABOUT A HIGHLY CONNECTED WORLD 455 (2010)

[92] Maurice E. Stucke & Ariel Ezrachi, *When Competition Fails to Optimize Quality: A Look at Search Engines*, 18 YALE J. L. & TECH. 70, 82 (2016).

move to a different product or service, consumers would have to switch en masse.

In addition, the nature of online activity means that privacy policies generally are stipulated in adhesion terms, which do not leave room for individual negotiation.[93] Since offering individually negotiated terms for treatment of personal information would be impractical and costly, changes in a company's privacy policies generally apply to everyone. This situation effectively turns contract negotiation into a public good, creating a classic collective action dilemma. As long as no single individual has sufficient incentives to bear the high costs of renegotiating privacy policies, cooperation is essential. Collective action problems are heightened in the online environment, where users tend to be strangers to one another. Privacy advocacy organizations can help to overcome these collective action problems, but certainly do not eliminate them.

B. Distortions of Individual Preferences

Here we describe some of the ways in which markets may fail to reflect consumers' actual individual preferences regarding the treatment of their personal data, independently (for the most part) of the externality and collective action problems. Some of the issues we discuss here arise from strategic behavior on the part of businesses seeking to collect and use the information, while others are simply inherent to the nature of information, its flow and its aggregation as well as human cognitive weaknesses. The taxonomy is not critical to our point, which is simply that there are many likely sources of such failures in markets involving personal information.

a. Lack of Information, Information Asymmetries and Myopia

The efficiency and presumptive social benefit of market transactions in personal information depends on the assumption that consumers' decisions to share such information are made in fully informed conditions and thus reflect their true preferences.[94] When consumers lack significant relevant information or cannot meaningfully process the information they have, this basis for

[93] Neil Weinstock Netanel, *Cyberspace Self-Governance: A Skeptical View from Liberal Democratic Theory*, 88 CALIF. L. REV. 395, 437-38 (2000).
[94] Shaun B. Spencer, *Reasonable Expectations and the Erosion of Privacy*, 39 SAN DIEGO L. REV. 843, 892 (2002)

relying on the market loses much of its power.[95] Information imbalances may stem from misrepresentation, concealment of information, or arise when information is too costly to uncover.[96] Not all information asymmetries lead to market failure—markets fail mostly when the information disproportion skews parties' negotiations and affects the transactions that actually occur.[97]

Here, however, access to relevant information is highly asymmetric between collectors of personal information and the subjects of that information. Companies' practices of data collection and, even more, of data use are largely opaque to consumers. Notwithstanding the existence of privacy policies purporting to inform users of information practices, consent to disclosure is largely meaningless.[98] Even to the extent information is provided in privacy policies, users face extremely high transaction costs of obtaining, reading, and understanding those notices.[99] Privacy policies are often vague, too complicated to be understood by an average user, and liable to be changed at any time, sometimes without notice.[100]

Companies have incentives to design privacy policies so as to discourage consumers from reading them and to obfuscate practices that consumers are likely to dislike. These incentives arise not simply for the sorts of collateral reasons common in consumer markets, e.g. companies wanting to protect themselves from liability or to obscure some undesirable features of the product. The collection and use of personal information is often the

---

[95] *See generally* Peter Swire, *Markets, Self-Regulation, and Government Enforcement in the Protection of Personal Information*, in PRIVACY AND SELF-REGULATION IN THE INFORMATION AGE, U.S. DEP'T OF COM. (1997), available at SSRN https://papers.ssrn.com/sol3/papers.cfm?abstract_id=11472&download=yes.

[96] Robin Paul Malloy, *Law in a Market Context: An Introduction to Market Concepts in Legal Reasoning* 171-72 (2004)

[97] Paul M. Schwartz, *Privacy and the Economics of Personal Health Care Information*, 76 TEX. L. REV. 1, 24 (1997).

[98] James Grimmelmann, *Saving Facebook*, 94 IOWA L. REV. 1137, 1181-82 (2009); Julie E. Cohen, *Privacy, Ideology, and Technology: A Response to Jeffrey Rosen*, 89 GEO. L.J. 2029, 2041 (2001) (showing how current consent models harm users' privacy rights without offering sufficient information or true control).

[99] Strandburg, *supra* note 75, at 145. Also, except in the rare instance when a major breach is widely reported, cybersecurity concerns are also not entirely obvious to users. *Id.* at 146 ("Information about data security comes to consumers only episodically, when breaches make news. Moreover, data breach notification tells users little or nothing about the potential for bad acts by rogue company insiders.").

[100] *Id.* at 157.

major, if not the only, ultimate business objective. Companies simply do not want to respond to consumer preferences that would lessen the flow.

Even when consumers are generally mindful that information about them is being collected, they still have little idea about how much of it is retained and for how long, with whom it is shared, for what purposes it is used, and how the sharing or use affects them.[101] Thus, consumers cannot effectively assess the costs of information collection at the time of collection and are unable to express their preferences when making information-related decisions.[102] With generally little awareness of data-sharing activities, individuals cannot appropriately estimate potential injuries and protect themselves against them.[103] Because it is difficult, and sometimes impossible, to trace harmful personal information flows to particular sources, the market often does not produce usual information about the past behavior of particular companies to inform consumer's decisions.[104]

Consumers' are also likely to have difficulty expressing their true preferences in transactions involving personal data collection because of the absence of a salient exchange transaction or "point of purchase."[105] In common sales transactions, consumers can estimate their disutility from whatever payment they are making by relying on extensive market experience.[106] Data collection ostensibly substitutes for payments for various nominally "free" products and services. But personal information is not money. It is virtually impossible for a consumer to "price" the expected disutilities that will stem from such data collections, because they are not predetermined and almost entirely dependent on future uses or misuses of the data.[107] User data is "credence payment" collected at intervals, and users are lacking knowledge

---

[101] *Id.* at 143 ("[P]rivacy policies often disclose the fact that consumer information collected by one online entity is shared with other entities, without providing specifics about to whom disclosure is made, from whom information is obtained, and for what purposes information from different sources is combined.").
[102] Strandburg, *supra* note 75, at 144.
[103] James P. Nehf, *Recognizing the Societal Value in Information Privacy*, 78 WASH L. REV. 1, 20 (2003).
[104] *Id.* at 28; *see also* Strandburg, *supra* note 75, at 148 ("[I]t is often difficult, if not impossible, even in retrospect, to trace any particular disutility caused by data access to any particular data disclosure.").
[105] Strandburg, *supra* note 75, at 150-52.
[106] *Id*.
[107] *Id*.

as to most aspects of it, including, particularly, the cost over time.[108]

The effects of inadequate information and information processing capacity are exacerbated by commonly observed cognitive limitations that have been identified in the field of behavioral economics. These cognitive limitations push against the very basic assumption of market theories - that humans are rational actors making rational decisions to maximize utility in the face of uncertainty.[109] These effects can cause consumers' purchasing activity to fail to reflect their true long-term preferences. Behavioral economics suggests, for example, that human beings systematically suffer from difficulties in assessing risk.[110] Specifically, humans fail to accurately estimate the expected costs of low probability, high cost harms—often precisely the kind of disutility that data collection produces.[111] Individuals' bounded capacity for information processing and bounded rationality seem particularly likely to be at play in their assessments of transactions in which personal information is disclosed. Indeed, studies show that at the zero-price point, where information-services exchanges often take place, individuals react irrationally, overly zealous to purchase goods at zero price while ignoring potential dis-utilities.[112]

b. High Transaction Costs and Lock-in Effects

Transaction costs are pervasive, and unavoidable to some extent, in personal information (and other) markets. To the extent that transaction costs are unavoidable, and not imposed strategically by one party to a transaction, they do not result in market failures. An individual's true preferences sensibly incorporate such unavoidable costs of doing business. In some situations, however, transaction costs distort demand signals in ways that could potentially be remedied by information privacy regulation.

---

[108] *Id.* at 131-32 .

[109] Melvin Aron Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47 STAN. L. REV. 211, 213 (1995); Alessandro Acquisti & Jens Grossklags, *Privacy and Rationality: A Survey, in* PRIVACY AND TECHNOLOGIES OF IDENTITY: A CROSS-DISCIPLINARY CONVERSATION 15, 16 (Katherine J. Strandburg & Daniela Stan Raicu eds., 2006).

[110] Cass R. Sunstein, *The Storrs Lectures: Behavioral Economics and Paternalism*, 122 YALE L.J. 1826, 1842-52 (2013).

[111] Strandburg, *supra* note 75, at 149.

[112] Kristina Shampanier, Nina Mazar & Dan Ariely, *Zero as a Special Price: The True Value of Free Products,* 26 MARKETING SCI. 742, 743 (2007).

Currently, for example, the transaction costs incurred in understanding and agreeing to the terms of service and privacy policies of most information-collecting businesses are prohibitive, especially in light of the frequent changes to privacy policies.[113] High transaction costs also prevent consumers from negotiating for better terms or looking for other market solutions and further exacerbate information processing problems. [114] In these circumstances, high transaction costs benefit businesses, which in turn lack market incentives to figure out ways to lower those costs and better inform consumers of risks associated with information collection and use.[115] Privacy regulation has the potential to shift the default so as to lower transaction costs, shift them onto the businesses that are best able to determine how to minimize them and thus facilitate better expression of consumer preferences in the market.

Transaction costs also contribute to lock-in effects that can inhibit consumers from expressing their preferences by shifting from one supplier, product or service to another. Some switching costs are unavoidable for products and services that involve a learning curve, but commentators have argued that transaction costs are sometimes used strategically by personal information-based businesses to exacerbate these unavoidable switching costs.[116] Information-intensive businesses often lure people into a lock-in either by showcasing robust privacy practices[117] or by offering a "free" product/service.[118] In what some have called a

---

[113] Aleecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, 4 I/S: J. L. & POL'Y FOR INFO. SOC'Y 543, 565-68 (2008); Daniel J. Solove, *Privacy Self-Management and the Consent Dilemma*, 126 HARV. L. REV. 1879, 1880 (2013); Jonathan A. Obar & Anne Oeldorf-Hirsch, *The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services* (Aug. 24, 2016) (unpublished manuscript), http://ssrn.com/abstract=2757465 (saying 74% of surveyed participants failed to read agreement in which the terms stipulated the handing out of their first-born child to the company); Victoria L. Schwartz, *Corporate Privacy Failures Start at the Top*, 57 B.C. L. REV. 1693, 1709-10 (2016) ("[M]any corporations change privacy policies frequently, making it harder for even the most diligent consumers to keep up with all the changes to these contracts of adhesion.").

[114] *Regulating Mass Surveillance*, *supra* note 80, at 1733, 1735.

[115] Jared S. Livingston, *Invasion Contracts: The Privacy Implications of Terms of Use Agreements in the Online Social Media Setting*, 21 ALB. L.J. SCI. & TECH. 591, 633 (2011).

[116] Chris Jay Hoofnagle & Jan Whittington, *Free: Accounting for the Costs of the Internet's Most Popular Price,* 61 UCLA L. REV. 606, 640 (2014).

[117] Paul Ohm, *Branding Privacy*, 97 MINN. L. REV. 907, 909, 922 (2013).

[118] Hoofnagle & Whittington, *supra* note 116, at 643-44.

"privacy lurch," services later trade privacy protection for profit making, shifting towards weaker privacy safeguards after users have already invested their time, energy, and social capital in the service,[119] making it costly for various reasons for them to switch. Alternatively, and more commonly, companies offer a nominally "free" product or service under deficient privacy terms that users do not take adequately into account for reasons already discussed. Over time, users become locked in to the product for various reasons, including the costs of recovering or recreating data that would be lost if they attempted to move to an alternative provider.[120] Some such businesses, most notably Facebook, have used their access to personal information, along with their ability to experiment with and manipulate their users' experiences, to conduct research arguably used to design their platforms to create psychological lock-in effects akin to (or perhaps equivalent to) psychological addiction.[121]

Lock-in effects distort demand signals by obscuring consumers' true preferences. In addition to preventing suppliers from getting an accurate "read" on consumer preferences, lock-in effects create barriers to entry that can cause appropriability failures, as discussed further below.[122]

---

[119] Ohm, *supra* note 117, at 909.

[120] Hoofnagle & Whittington, *supra* note 116, at 643-44; *see also* Gabriela Zanfir, *The Right to Data Portability in the Context of the EU Data Protection Reform*, 2 INT'L DATA PRIVACY L. 149, 152 (2012).

[121] *See, e.g.*, Hoofnagle & Whittington, *supra* note 116, at 610 ("[I]nefficiencies rise as advertising and marketing activities become increasingly intrusive, gradually changing the value exchanged by the consumer for the service. Such costs include lock-in."). The lock-in effects created by transaction costs are distinct from network effects, though both can be present in the same situation. Consider, for example, the choice between using a Mac computer and using a PC. At one time, documents created on PCs could not be easily read by Macs (and vice versa). That situation created a network effect: the advantage of using a PC grew in proportion to the number of one's friends or colleagues who were PC users (and vice versa). For example, when one of us switched careers from physics to law, she went from a context dominated by Mac users to one dominated by PC users. As a Mac user in the legal world, she felt the loss of network benefits quite keenly, but did not immediately switch to a PC. Why? Because of the transaction costs of switching from one system to the other, including buying a new computer, learning a new operating system, etc. Those transaction costs created lock-in effects that competed with network benefits of switching. (Eventually, the network benefits won out.) Nowadays, transferring files between PCs and Macs is pretty seamless; that network effect has nearly disappeared. Switching costs have also gone down as the two operating systems have become more similar, but they are still non-trivial.

[122] *Id.*; *see also* James Grimmelmann, *Saving Facebook*, 94 IOWA L. REV. 1137, 1192 (2009) ("When users can't easily carry their digital identities with them

As discussed in Part III, transaction costs can create collective action problems even without externalities. This is true when reducing the transaction costs enough to make switching attractive requires nonrivalrous measures, such as the discovery or production of information, and these measures are too expensive for each individual to undertake independently. In the privacy arena, much of the information about company practices that individual consumers cannot afford to ferret out or process adequately is nonrivalrous, thus adding to the collective action problems associated with externalities.

c. Misalignment with Social Values

From a traditional law and economics perspective, the question of how much information consumers are willing to disclose and use in exchange for various goods and services can, and should, be answered by the market. Privacy advocates have countered this argument not only by pointing to the market failures discussed in the previous two sections, but also with broader arguments against the use of economic markets to allocate privacy's individual and societal benefits.[123] From this perspective, markets can also fail by redistributing goods in ways that disagree with social principles of equity,[124] by incentivizing behavior that does not conform to people's long-term interests, [125] or by interfering with equal access to fundamental rights. [126] Many scholars have discussed and documented these sorts of failures in the privacy context. Craig Konnoth, for example, showed how the current regulatory framework leads to medical information disclosure practices that favor more socially and economically privileged patients. [127] Researchers have also identified discriminatory uses of digital personal information. For example, Google's reliance on data about which gender is likely to click on certain sorts of ads, resulted in the preferential targeting of ads for

---

from one site to another, it's much harder for new entrants to compete with an entrenched incumbent.").

[123] Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject As Object*, 52 STAN. L. REV. 1373, 1395 (2000).

[124] Joe Wallis & Brian Dollery, *Market Failure, Government Failure, Leadership And Public Policy*, 22 (1999).

[125] *Id*.

[126] *Id.* at 23.

[127] Craig Konnoth, *Health Information Equity,* 165 U. PA. L. REV. 1317 (2017).

high-paying jobs ads to men and those offering low paying jobs to women.[128]

Economic markets unavoidably treat personal information as a commodity, ignoring societal values such as dignity and autonomy that are affected by inappropriate flows of personal information. [129] These values arguably are prerequisites for a functioning polity, [130] and independently worth preserving regardless of whether they can be justified by economic cost-benefit trade-offs. [131] The collection, aggregation and use of personal information is also likely to have major systemic effects on society, particularly when such information is aggregated and used in predictive fashion as discussed in the next section or when it can be used to facilitate government surveillance. These effects lead to a broader spectrum of failures relating to non-utilitarian distributive and ethical values that many, including us, believe can provide convincing justifications for regulation.[132]

These systemic social effects are not only hard to assess at the level of individual transactions, but, as public goods, are especially susceptible to collective action problems. Privacy also invokes the sorts of ethical dilemmas that many believe should be resolved democratically, as well as issues of constitutional rights, especially for disfavored minorities.[133]

---

[128] *See* Amit Datta, Michael Carl Tschantz & Anupam Datta, *Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination*, 2015:1 PROC. ON PRIVACY ENHANCING TECHS. 92.

[129] James P. Nehf, *Incomparability and the Passive Virtues of Ad Hoc Privacy Policy*, 76 U. Colo. L. Rev. 1, 30 (2005).

[130] Scott R. Peppet, *Unraveling Privacy: The Personal Prospectus and the Threat of A Full-Disclosure Future*, 105 NW. U. L. REV. 1153, 1188 (2011).

[131] *Nehf, supra* note 129, at 30.

[132] *See* Paul M. Schwartz, *Internet Privacy and the State*, 32 CONN. L. REV. 815, 834 (2000) (arguing that privacy is essential to the creation and maintenance of both individuals and society, and without it individuals cannot adequately participate in a democratic collective); Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1424 (2000) (arguing that to construct a self, individuals need personal autonomy with "insulation from outside scrutiny" and this autonomy holds benefits not only to the individual but also to society); DANIEL J. SOLOVE, UNDERSTANDING PRIVACY 92 (2008) ("Privacy protects aspects of individuality that have a high social value; it protects individuals not merely for their sake but for the sake of society.").

[133] Cohen, *supra* note 123, at 1395.

C. Aggregation Failures in Personal Information Markets

The way personal information is aggregated to infer additional personal information give rise to particularly profound implications for an additional type of market failure. Information aggregation combines and exacerbates the effects of many of the market failure mechanisms discussed in the previous three sections. The aggregative quality of information makes the market failures associated with personal information particularly resistant to consumer self-help efforts and simple fixes; taking it into account is critical to effective regulatory design.

The basic observation is that personal data, when aggregated—across individuals, across sources and across time—becomes more than the sum of its parts. When personal data is aggregated, it can be synthesized to make inferences, create generalizations, and draw conclusions.[134] More and more, this synthesis is performed computationally and its inputs, methodologies and outputs are kept secret by personal information-based businesses.[135] Many businesses also supplement the information they collect from their customers or users by aggregating it with data purchasing additional data from online markets and data brokers.[136] The aggregative qualities of personal information make it harder to fix the market failures discussed in the previous three sections.

For example, one common regulatory response to information asymmetry is information disclosure. Disclosure of company information practices, combined with an opportunity for informed consent, is a primary requirement of most privacy regulations in the U.S. and abroad.[137] Such transparency is desirable, but ultimately insufficient because of the effects of data aggregation. Even if a company informs consumers what information it collects directly and what other sources of information it uses (far more than is disclosed in the typical privacy policy), it remains extremely difficult, if not impossible,

---

[134] Shubha Ghosh, *Informing and Reforming the Marketplace of Ideas: The Public-Private Model for Data Production and the First Amendment*, 2012 UTAH L. REV. 653, 674 (2012).

[135] CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY (2016)

[136] This practice is known as "appending" or "enhancing" data. Hoofnagle & Whittington, *supra* note 116, at 646-47.

[137] *See* DLA PIPER, *Data Protection Laws of the World*, https://www.dlapiperdataprotection.com (last visited Sept. 9, 2018 6:08 PM).

for users of personal information-based products and services to assess the marginal disutility of any given instance of personal data collection and account for it in their market behavior.[138] When information is aggregated (and, often, cross-referenced and "enhanced" with information obtained from other sources, such as data brokers) companies can infer personal details that were not directly disclosed, often with a high level of accuracy.[139] Without access to the algorithms that companies use to make inferences about aggregated personal data, consumers cannot predict what a business might infer about them when any given "piece" of personal information is aggregated with other data that is currently in the business's possession. Certainly, they cannot assess the inferences that a company might make by aggregating it with data that might come into the company's hands in the future. The likely consequences of agreeing, at a given time, to a particular sort of data collection are thus, at a minimum, unreasonably expensive to game out and likely unpredictable.[140] As a result, users' market behavior is highly unlikely to reflect meaningful privacy choices.[141]

If consumers evaluate information collection at its independent marginal disutility in terms of privacy loss, without accounting for aggregation effects, they are guaranteed to assign a lower value to the data than collectors, who value the information when aggregated with other data.[142] Under this valuation distortion, individuals will always agree to sell personal information at a price collectors always agree to pay, effectively generating substantial over-disclosure.[143] But taking aggregation into account in assessing marginal disutility is essentially impossible:

---

[138] *See* James P. Nehf, *Shopping for Privacy Online: Consumer Decision-Making Strategies and the Emerging Market for Information Privacy*, 2005 U. ILL. J.L. TECH. & POL'Y 1, 29-30 (2005) (explaining that many factors make it difficult for consumers to rationally evaluate and compare privacy decisions, resulting in a lack of attention paid to privacy).

[139] Michal Kosinski, David Stillwell & Thore Graepel, *Private Traits and Attributes Are Predictable From Digital Records of Human Behavior*, 110 PROC. OF THE NAT'L ACAD. OF SCI. 5802 (2013).

[140] *Regulating Mass Surveillance*, *supra* note 80, at 1732.

[141] Jan Whittington & Chris Jay Hoofnagle, *Unpacking Privacy's Price*, 90 N.C. L. REV. 1327, 1359-61 (2012).

[142] A. Michael Froomkin, *The Death of Privacy?*, 52 STAN L. REV. 1461, 1503-04 (2000). That is true even if they do not know the specifics of how the information might be aggregated with other data and what will be the exact value—they still know that the value of the data is likely to increase.

[143] *Id*.

> To determine marginal disutility, an Internet user must have information about how the incremental data collected in association with the particular activity changes the overall availability of information about her in the online ecosystem. Not only that, she must be able to connect that increment in available information to an increment in expected disutility. This is essentially an impossible task.[144]

Moreover, aggregation compounds the significant difficulty that consumers face in tracing privacy harms to particular businesses or particular information disclosures. Aggregation thus exacerbates the effects of cognitive myopia, making it even more unlikely that consumers will take sufficient account of potential future harms.[145]

Data aggregation also exacerbates the externalities associated with disclosing personal information, since it makes easier for a company to draw inferences and make predictions about individuals even when they have chosen not to disclose certain information to that company. [146] For example, one individual's disclosure of information about herself may be used to infer information about others in her network, who chose not to disclose that information.[147] That indirectly revealed information can then be added to profiles, either enhancing or initiating them.[148]

While businesses also confront uncertainty about the magnitude of the value they will be able to extract by collecting and aggregating personal information, they can ordinarily expect that aggregation will be worth their while, particularly because of the low cost of data collection and storage. This asymmetry of expectations gives businesses strong incentives to engage in more collection and use of personal information than consumers would prefer.[149] Indeed, many companies are incentivized to collect and retain as much personal data as possible simply because it might

---

[144] Strandburg, *supra* note 75, at 147-48.
[145] *See* Froomkin, *supra* note 142, at 1502-04.
[146] Mark MacCarthy, *New Directions in Privacy: Disclosure, Unfairness, and Externalities*, 6 I/S: J. L. & POL'Y FOR INFO. SOC'Y 425 (2011).
[147] *Id.* at 449; *see also* Nizan Geslevich-Packin & Yafit Lev-Aretz, *On Social Credit and The Right to Be Unnetworked*, 2016 COLUM. BUS. L. REV. 339.
[148] *Id.* at 449-50.
[149] Strandburg, *supra* note 75, at 150.

come in handy someday.[150] The likelihood that personal information-based markets will fail to align with individual and social preferences is increased because business models in this sector commonly aim to use aggregated personal information for purposes that do not focus on satisfying consumer preferences. Such purposes include ad targeting, selling personal data, selling predictive models based on personal data, and manipulating consumers.[151] As the trope goes, "you're not the customer, you're the product."

Perhaps most importantly, the interconnectedness of personal information makes it essentially pointless for consumers to try to express their privacy preferences by picking and choosing among businesses based on their use of personal information. The nature of personal information collection today belies the notion that consumers can make meaningful market trade-offs based on the benefits and potential privacy costs of particular goods and services.[152] Market failures essentially leave consumers with only three real choices (and the efficacy of the third is questionable):

1. Go more or less "all in" for the online experience
2. Withdraw significantly or completely from online activities in order to protect their privacy, or
3. Attempt to deploy drastic and time-consuming technical measures, such as encryption and Tor.[153]

Finally, aggregation also heightens the likelihood that markets will fail to account for important social and ethical values. The aggregation and synthesis of personal information across many sources and for a wide variety of purposes also creates just the sorts of systemic effects, implications for economic distribution and potential for deleterious effects on minority groups that market transactions fail to take into account. The systemic adoption of the "free", advertising-based business model for some sorts of

---

[150] Companies adopt some risks by doing this, since large stores of personal information are honeypots for hackers and data breaches can have some reputational cost. The ubiquity of these practices, along with the very effects discussed in the Section, make it difficult for consumers to credibly threaten market punishments for such breaches however.

[151] Daniel Susser, Beate Roessler & Helen Nissenbaum, *Online Manipulation: Hidden Influences in a Digital World*, 4 GEO. L. TECH. REV. 1 (2019).

[152] An additional implication of trying to opt out of data collection would be the possibility of being profiled as a criminal. *See* Janet Vertesi, *My Experiment Opting Out of Big Data Made Me Look Like a Criminal*, TIME (May 1, 2014), http://time.com/83200/privacy-internet-big-data-opt-out.

[153] Strandburg, *supra* note 75, at 164-65.

products and services has made it difficult, if not impossible, for either individual users or individual companies to opt for another approach.[154] Another systemic effect arises from the influence of ubiquitous personal information tracking on the technological design of the Internet and other infrastructure.[155]

In sum, there are numerous reasons to anticipate significant and widespread misalignment between expressed market demand and true individual and societal preferences regarding the collection, flow, and use of personal information. The resulting market failures are likely to lead to a depressed personal information "price," to over-investment in the supply of surveillance and collection technologies, and to under-investment in privacy enhancing technologies ('PET') and technologies that improve expression of privacy preferences.[156] Most importantly for present purposes, we believe that the cumulative effect of the sources of misaligned demand signals discussed in this Part provides a strong prima facie justification for information privacy regulation. Moreover, because there are a variety of reasons to anticipate failures in markets involving personal information, one need not be persuaded by all of them to agree that such a prima facie justification exists.

IV. FAILURES OF APPROPRIABILITY IN PERSONAL INFORMATION-BASED MARKETS

Having reviewed some of the ways in which misaligned demand can create failures in the market for personal information-based products and services, we now turn to failures of appropriability. We describe the sorts of failures of appropriability that are likely to arise in personal information-based markets,

---

[154] *Id.* at 644; *see also* Strandburg, *supra* note 75, at 124 n.98.

[155] Chris Jay Hoofnagle & Jan Whittington, *Free: Accounting for the Costs of the Internet's Most Popular Price,* 61 UCLA L. REV. 606, 643 (2014) (discussing Google's treatment of referrer headers as an example for such broader lock-in effect); *see also* Oren Bracha & Frank Pasquale, *Federal Search Commission? Access, Fairness, and Accountability in the Law of Search*, 93 CORNELL L. REV. 1149, 1182 (2008) (discussing lock-in effects in the market of personalized search).

[157] Schwartz, *supra* note 97; *see also* Lital Helman, *Curated Innovation*, 49 AKRON L. REV. 695, 705 (2016) ("These technologies potentially have enormous societal values in preventing privacy harms. Yet, due to various failures in privacy-related markets, the adoption rate of these technologies is probably lower than the actual value they provide. Because the market cannot reflect the full value that these technologies generate, innovators are less likely to invest in creating such solutions, despite the societal value such products can yield.").

taking account of the corrections provided by standard intellectual property doctrines. Specifically, we argue that trade secrecy's standard limiting doctrines tend to be ineffective in personal information-based markets, creating systematic tendencies toward excessive exclusivity and thus over-compensation for some personal information-based innovations. In addition, data aggregation and, where present, network effects create barriers to entry for competitive alternatives or follow-on innovations. These effects result in an appropriability landscape that is systematically distorted and thus tend to induce a portfolio of market innovation that diverges from market demand signals.

A. Intellectual Property-Related Failures of Appropriability in Personal Information-Based Markets

As discussed earlier, intellectual property doctrine uses various limiting doctrines to smooth out the appropriability landscape by balancing society's interest in addressing free rider problems against its interest in encouraging competitive follow-on innovation. While personal information-based companies also avail themselves of patent and copyright protections, they tend to rely heavily on trade secrecy regarding their collections of personal information. [157] Certain features of personal information-based products and services combine to make it highly likely that trade secrecy's primary limiting doctrines—independent invention and reverse engineering—tend to be ineffective in these markets.

a. Patent and Copyright Protection for PI-based Products and Services

Technical innovations implemented in personal information-driven products and services are eligible for patent protection to the same extent as other software and business method inventions. Copyright is also available for the expressive aspects of these companies' software and user interfaces. [158] Patent protection in these areas has been highly controversial, with years of debate as to whether patent protection should be available at all

[157] *See* Laura Palk & Krishnamurty Muralidhar, *A Free Ride: Data Brokers' Rent-Seeking Behavior and the Future of Data Inequality*, 20 VAND. J. ENT. & TECH. L. 779, 783 (2018).

[158] *See, e.g.*, Oracle Am., Inc. v. Google Inc., 750 F.3d 1339, 1367 (2014) (saying a computer interface "is entitled to copyright protection as long as the author had multiple ways to express the underlying idea").

and, if so, what it should cover.[159] Many scholars have argued, for a variety of reasons, that patents are unnecessary for business methods and rarely necessary for software innovations.[160] The extent to which copyright should protect software, given that ideas and methods of operation are uncopyrightable, is also a recurring subject of controversy.[161] The upshot, at least at this juncture, is that patents remain available for software and business method inventions, but the scope of patentable inventions has been narrowed significantly by recent Supreme Court decisions.[162] Some groups have proposed legislation aimed at over-turning some of these decisions.[163]

These are fascinating debates, but we do not engage them here. For present purposes, we simply assume that patent and

---

[159] *See, e.g.*, Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575 (2003) (arguing that patent policy should take account of the needs of different industries); Leo J. Raskind, *The* State Street Bank *Decision: The Bad Business of Unlimited Patent Protection for Methods of Doing Business*, 10 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 61 (1999); Julia Angwin, *'Business Method' Patents, Key to Priceline, Draw Growing Protest*, WALL ST. J., Oct. 3, 2000, at B1.

[160] *See, e.g.*, Michael J. Meurer, *Business Method Patents and Patent Floods*, 8 WASH. U. J.L. & POL'Y 309 (2002) (describing economic harm caused by patent floods, in particular a current flood in business patents); Joseph H. Sommer, *Against Cyberlaw*, 15 BERKELEY TECH. L.J. 1145, 1220 (2000) (stating that business method patents are unnecessary because they have nothing to do with technology); Pamela Samuelson, Benson *Revisited: The Case Against Patent Protection for Algorithms and Other Computer Program-Related Inventions*, 39 EMORY L.J. 1025 (1990).

[161] *See* Richard H. Stern, *Symposium: The Future of Software Protection: The Bundle of Rights Suited to New Technology*, 47 U.PITT. L. REV. 1229, 1259 (1986) (presciently noting that patent and copyright may "refuse to protect" algorithms because they are "mere ideas").

[162] According to Paul R. Gugliuzza, the Supreme Court "has decided a remarkable thirty-three patent cases since 2006." Paul R. Gugliuzza, *How Much Has the Supreme Court Changed Patent Law?*, 16 CHI.-KENT J. INTELL. PROP. 330, 338 (2017). *See, e.g.*, Mayo Collaborative Servs. v. Promethus Labs., Inc., 566 U.S. 66, 72-73 (2012) (prohibiting patents directed to laws of nature, natural phenomena, or abstract ideas, unless they also contain an "inventive concept"); KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398, 417-22 (2007) (replacing the "teaching, suggestion, or motivation requirement with a flexible analysis that makes it easier to invalidate a patent based on obviousness). Also, notably, the Court denied the opportunity to review *Ariosa Diagnostics, Inc. v. Sequenom, Inc.*, a case where the Federal Circuit invalidated a patent on a non-invasive prenatal genetic test, viewed by many in the scientific community as a major breakthrough, because the test involved a "natural law." 788 F.3d 1371, 1378 (Fed. Cir. 2015), *cert. denied*, 136 S. Ct. 2511 (2016).

[163] *See, e.g.*, Press Release, INTELL. PROP. OWNERS ASS'N, *IPO Supports Legislation To Amend U.S. Patent Act Section 101* (Jan. 31, 2017) (on file with author).

copyright doctrines, though perennially contested and sometimes evolving, reflect the way in which society ordinarily trades off the competing values of incentivizing innovation, minimizing deadweight losses, and avoiding undue burdens on follow-on innovators. The software and potentially patentable inventions associated with personal information-based business do not appear to pose unique problems for patent or copyright doctrine. We therefore assume that questions of patent and copyright doctrine can be reasonably separated from the design of information privacy regulation.

b. Trade Secrecy is Over-Protective for PI-Based Products and Services

Providers of software-based products and services have long resorted to trade secrecy protection to secure exclusivity in their offerings.[164] Trade secrecy doctrine evolved mostly in state law, but similar rules and applications developed across jurisdictions.[165] Adopted by most states, the Uniform Trade Secrets Act (UTSA) expensively defines "information" that is (i) valuable, and (ii) reasonably protected as trade secret.[166] Recently, Congress enacted the first federal trade secrecy statute: The provisions of the Defend Trade Secrets Act are similar, for present purposes, to pre-existing state trade secrecy laws.[167]

Trade secrecy has both functional and legal aspects. Functionally, secrecy protects any information that is actually kept secret from competitors.[168] Trade secrecy law applies to a wide

---

[164] *See* Peter S. Menell, *The Challenges of Reforming Intellectual Property Protection for Computer Software*, 94 COLUM. L. REV. 2644, 2652 (1994) ("The [software] industry had developed principally through trade secret protection."); Mark A. Lemley & David W. O'Brien, *Encouraging Software Reuse*, 49 STAN. L. REV. 255, 258 (1997) ("Trade secret law remained the dominant form of legal protection of software through the mid-1970s.").

[165] W. Nicholson Price II, *Regulating Secrecy*, 91 WASH. L. REV. 1769, 1776 (2016).

[166] UNIF. TRADE SECRETS ACT § 1(4), 14 U.L.A. 538 (2005). Another oft-cited definition is offered by the Restatement (Third) of Unfair Competition § 39 (1995) ("A trade secret is any information that can be used in the operation of a business or other enterprise and that is sufficiently valuable and secret to afford an actual or potential economic advantage over others."). On the federal front the Economic Espionage Act, 18 U.S.C. § 1831 (2012), and the recently enacted Defend Trade Secrets Act of 2016, Pub. L. No. 114-153, 130 Stat. 376 (codified as amended in scattered sections of 18, 28 U.S.C.) also offer trade secrecy protection.

[167] Defend Trade Secrets Act, 18 U.S.C. § 1836 (2016).

[168] David S. Levine, *Secrecy and Unaccountability: Trade Secrets in Our Public Infrastructure*, 59 FLA. L. REV. 135, 145 (2007) ("At its core, trade secret law

variety of technical and non-technical information that is economically valuable, including methods, facts, and ideas that are excluded from other intellectual property rights.[169] It provides remedies against the misappropriation of information that was subject to reasonable protections against disclosure.[170] However, trade secrecy protection evaporates once information becomes widely known in an industry. Moreover, competitors are free to obtain trade secret information through independent invention or reverse engineering.

Assuming trade secrecy survives long enough, it, like patent and copyright protections, can avert failures of appropriability arising from the free rider problem.[171] Trade secrecy protection is both broader and narrower in scope than patent protection. It is broader in its subject matter and offers a nominally unrestricted term of legal protection. Trade secrecy is narrower than patent protection, however, because competitors are permitted, both functionally and legally, to reverse engineer or independently derive the information needed to create and market a competing product or service.[172] Thus, reverse engineering and independent invention provide important limitations on trade secrecy exclusivity, helping to avoid overcompensating innovators.[173] Especially because trade secrecy is so broadly applicable and the potential term of legal protection is uncapped,

---

envisions a fundamental scenario: competition between private actors whose primary objective is pecuniary gain.").

[169] Michael Mattioli, *Disclosing Big Data*, 99 MINN. L. REV. 535, 550 (2014).

[170] *Id*.

[171] J. Jonas Anderson, *Secret Inventions*, 26 BERKELEY TECH. L.J. 917, 962 (2011) (arguing that the most efficient solution to the free rider problem lays with inventors' choice of protection scheme, whether it is patent or trade secrecy).

[172] UNIF. TRADE SECRETS ACT § 1 Cmts. 1-2 (UNIF. LAW COMM'N, amended 1985). This was also confirmed by the Supreme Court in *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470, 490 (1974) (pointing to independent creation and reverse engineering as the key factors in finding that trade secret was not preempted by patent law).

[173] This claim has been empirically tested in Petra Moser's seminal work: Moser used historical data from the Crystal Palace World's Fair to show that inventors rely on trade secrecy protection when secrecy is feasible. Over time, however, the decreased cost of reverse engineering has made trade secrecy less appealing to inventors, who turn to patent protection. Petra Moser, *Innovation Without Patents: Evidence from World's Fairs*, 55 J.L. & ECON. 43 (2012). A similar theory was presented in a work by Keishun Suzuki, finding that "strengthened patent protection can increase economic growth when the risk of leakage of trade secrets is high. Conversely, when the risk is low, stronger patent protection hinders growth." Keishun Suzuki, *Economic Growth Under Two Forms of Intellectual Property Rights Protection: Patents and Trade Secrets*, 115 J. ECON. 49, 50-51 (2015).

the social benefits of trade secrecy protection depend crucially on the extent to which reverse engineering and independent invention are successful in this role. Reverse engineering and independent invention provide a kind of effective term limitation on trade secrecy, under the rough-and-ready theory that more difficult innovations tend to require greater upfront investment and, because independent invention will take longer, will have correspondingly longer periods of market exclusivity.[174]

For personal information-based innovations, however, trade secrecy protection tends to be over-compensatory, creating failures of appropriability inverse to the usual free rider problems. There are three basic reasons for these systematic failures of appropriability. First, and most importantly, personal information databases generally cannot be reverse engineered from the public-facing aspects of personal information-based products and services. As a result, reverse engineering tends to be ineffective in limiting trade secrecy exclusivity.[175] Second, in many cases, trade secrecy tends to be over-compensatory because there is minimal upfront investment to recoup. Upfront costs associated with technological invention or creative expression should be recoupable through standard intellectual property protections. From a free rider perspective, trade secrecy is important primarily for recouping the additional upfront costs of amassing personal information. For many personal information-based companies, however, those costs are extremely low, since they acquire personal information as a cheap by-product of providing other products and services. For this sub-set of companies, at least, there is not much need for trade secrecy to recoup upfront investment. Finally, potential independent inventors often face higher upfront costs than first inventors in these markets, rather than the equivalent (or perhaps slightly lower) upfront investments

[174] The story with reverse engineering is more complicated, but if reverse engineering gets too easy inventors can always opt to apply for patent protection.

[175] As Brenda Simon and Ted Sichelman found this access barrier is exacerbated when the "data-generating invention" is patented. Brenda M. Simon & Ted Sichelman, *Data-Generating Patents*, 111 NW. U. L. REV. 377, 379 (2017) ("Unlike information about the invention itself—which is often disclosed in patented improvements on the original invention—data-generating inventions tend to produce data that can be maintained as a trade secret. Patent holders enjoy an increased ability to aggregate and analyze "big data" obtained through leveraging data-generating patents, and they can protect the results using trade secret protection. This presents unique legal and economic consequences that we contend may be socially problematic under certain conditions.").

ordinarily assumed to be required.[176] This third point results from certain sorts of barriers to entry that are common in these markets, which we describe in the next section.

Overall, then, the combination of patent, copyright and trade secrecy exclusivity is likely to over-compensate personal information-based innovation,[177] though the extent to which this is the case will depend on the particular context. As a result, these sorts of innovations will tend to stick out from the appropriability landscape and stimulate over-investment relative to market demand signals.

B. High Entry Barriers in Personal Information-Based Markets

Markets for PI-based products and services tend to exhibit distinctive natural barriers to entry arising out of the particular qualities of personal information—its unique association with particular individuals, its non-linear aggregation and, often, its collection as a by-product of goods or services exhibiting network effects. From the perspective of the initial innovator, these barriers to entry raise the appropriability of the innovation, relatively more attractive than it would otherwise be. From the perspective of potential follow-on innovators, these barriers raise upfront costs, making follow-on innovation relatively less attractive than it would otherwise be.

Though some have argued that the acquisition and use of big data by online firms does not create significant barriers to entry,[178] others have criticized this position, viewing data as a strategic asset that could lead to market dominance and limit later entry.[179] Daniel Rubinfeld and Michal Gal's extensive analysis of market entry barriers in big data markets, showed that such barriers

---

[176] Patent law's assertion of exclusive rights against independent inventors has been critiqued on this basis. *See, e.g.*, Oskar Liivak, *Rethinking the Concept of Exclusion in Patent Law*, 98 GEO. L.J. 1643, 1657-74 (2010). Independent innovators might be expected to face somewhat lower upfront costs since the first innovator's efforts will have demonstrated that the innovation is technically doable and paved the way for consumer adoption.

[177] *See generally id.*

[178] *See, e.g.*, Darren S. Tucker & Hill B. Wellford, *Big Mistakes Regarding Big Data* 14 ANTITRUST SOURCE 6, 6-7 (2014).

[179] Howard A. Shelanski, *Information, Innovation, and Competition Policy for the Internet*, 161 U. PA. L. REV. 1663, 1679 (2013); Maureen K. Olhausen & Alexander P. Okuliar, *Competition, Consumer Protection, and the Right* [Approach] *to Privacy*, 80 ANTITRUST L.J. 121, 131 (2015); Zarsky, *supra* note 59, at 135.

"can arise in all parts of the data-value chain," though the extent and importance of such barriers is context-dependent.[180] Rubinfeld and Gal rightly disagree with arguments that data collection cannot create barriers to entry because of its non-rivalry. First, data is not fungible, and costs of acquiring certain pieces of personal information can certainly be different for different companies, depending, for example on whether they acquire as a byproduct of providing a service or have to purchase it on the market. Moreover, as Rubinfeld and Gal point out, barriers to entry may also be erected on parts of the data value chain other than data gathering, or as a result of the cumulative effect of a number of low entry barriers in several parts of the data value chain.[181]

Here, moreover, we are not concerned with demonstrating that barriers to entry are large enough or of the right sort to justify action by antitrust or competition authorities. Our question is very different: we look at barriers to entry not to ascertain their effects on competition per se, but to consider whether they undermine our usual reliance on intellectual property law to take care of failures of appropriability, which ordinarily allows us to set aside concerns that a regulation designed to re-align demand might unintentionally magnify appropriability failures.[182] Barriers to entry are especially likely with regard to personal information-based regulation for three reasons: i) the value of aggregated personal information; ii) the cost advantage of acquiring it from users; and iii) the interplay between data aggregation, network effects and lock-in.

C. Data Aggregation and Market Value

While data aggregation may eventually reach a point of diminishing returns, there is often a wide range over which the market value of a PI-based product or service grows non-linearly as more personal information is aggregated to be used as input in creating and delivering the products or services. The value of the product or service to each consumer grows in similar fashion. The quality of search results delivered to each user, for example, may be improved by combining personal information about many individuals with information about previous searches.[183] For ad-supported businesses, the value to individual users may or may not

---

[180] Daniel L. Rubinfeld & Michal S. Gal, *Access Barriers to Big Data*, 59 ARI. L. REV. 339, 369 (2017).
[181] *Id*.
[182] *See* Lev-Aretz & Strandburg, *supra* note 12.
[183] Mike Mathieson, *Using Behavioral Data to Improve Search*, EBAY (Apr. 13, 2011), https://www.ebayinc.com/stories/blogs/tech/using-behavioral-data-to-improve-search.

grow as data as collected, but the value to advertisers (who are the real customers) presumably does. Until they collect enough data to reach a point of diminishing returns (which may or may not exist, depending on the product or service), first entrants who continue to acquire personal information can maintain a persistent advantage against later entrants.

a. Cost Advantage of Acquiring Personal Information from Users

A later entrant might try to overcome the advantage a first entrant acquires by aggregating its users' data by purchasing a database of personal information from a data broker. This tactic will often be ineffective, however, for two reasons. First, if the first entrant collects personal information as a byproduct of some other activity, purchasing data puts a potential competitor at a cost disadvantage, since personal information acquired from users as a side effect of providing a product or service is essentially free. Second, any advantage gained by the purchase of personal data on the open market would be ephemeral, at least up to a point of diminishing returns from aggregation, since the first entrant could leapfrog ahead by purchasing the same data and combining it with data acquired directly from users. Dominant players in personal information markets maintain their advantages by refusing to sell their databases, particularly when the personal information they have collected is distinctive from what is available from data brokers. [184] Instead, they keep the data under trade secrecy protection and offer data-based services themselves. A notable exemplifier of such practice is Facebook, that offers sophisticated ad targeting services, but does not allow paid access to its database.[185]

---

[184] *See* Kurt Wagner, *This is How Facebook Uses Your Data for Ad Targeting*, RECODE (Apr. 11, 2018, 6:00 AM EDT), https://www.vox.com/2018/4/11/17177842/facebook-advertising-ads-explained-mark-zuckerberg ("Selling [its trove of personal] data to advertisers would significantly decrease Facebook's value.").

[185] Facebook allowed app developers to access some of the data it collected about its users until 2014. *See* James Vincent, *Academic Who Collected 50 Million Facebook Profiles: 'We Thought We Were Doing Something Normal'*, THE VERGE (Mar. 21, 2018, 7:39 AM) https://www.theverge.com/2018/3/21/17146342/facebook-data-scandal-cambridge-analytica-aleksandr-kogan-scapegoat.

b. Network Effects

As discussed earlier, many PI-based products and services also exhibit network effects on top of the advantages associated with data aggregation, such that the value of the product or service to each user is directly enhanced by the addition of more users.[186] Network effects are conceptually distinct from the effects of data aggregation. Thus, search engines tend to increase in value as more data is acquired, but the value to each user is not directly enhanced by the fact that others are using the same search engine. Email services and telephone systems, on the other hand, exhibit network effects as users are added even if no personal information is aggregated, simply because each user values the ability to reach more other users. Social media platforms tend to exhibit both network effects and data aggregation effects. While network effects and data aggregation effects are conceptually distinct, they are linked in practice.[187] Where both are present, they may feed back onto one another, with network affects attracting more users, who provide more personal information that can be used to enhance market value and attract yet more users and so on. Another way these two effects can compound one another is illustrated by advertising-supported social media platforms, where network effects may attract users, thus providing more aggregated data that can be used to enhance the value of targeting services offered to advertiser customers.

These effects can create powerful barriers to entry because first entrants begin with more users, and thus can offer a more attractive product or service, which attracts more users whose personal information can be fed back in to further enhance the product or service. And so on. As long as this cycle continues, second comers stand no chance of competing effectively.[188]

---

[186] MAURICE E. STUCKE & ALLEN P. GRUNES, BIG DATA AND COMPETITION POLICY 170 (2016) ("[T]he more people actively or passively contribute data, the more the company can improve the quality of its product, the more attractive the product is to other users, the more data the company has to further improve its product, which becomes more attractive to prospective users.")

[187] Rubinfeld & Gal, *supra* note 180, at 377.

[188] Rubinfeld and Gal point out that there is an ongoing debate about the presence of entry barriers with respect to search: "Microsoft has argued that it faces substantial barriers to entry because it obtains an order of magnitude fewer search queries than does Google. From Microsoft's perspective, its analysis of its own queries puts it at a disadvantage. Google counters by pointing out that efficient scale can be readily achieved through the analysis of queries on Bing, suggesting that if Microsoft is disadvantaged it is due to Google's more successful algorithm or other comparative advantages, not scale. This implies

Anticipated high entry barriers increase the anticipated appropriability of some types of personal information-based innovation, making those innovation paths particularly attractive relative to consumer demand. Existing high entry barriers, on the other hand, have the opposite effect on the attractiveness of competing or follow-on innovation, as the next section explains.

D. Implications for Follow-on Innovation

Barriers to entry discourage competitive and follow-on innovation by raising its expected upfront costs. When there are no network effects, data aggregation effects or similar barriers to entry, relatively small improvements or product differentiations can be enough to attract enough customers to recoup a competing or follow-on innovator's upfront creative investments. This relative ease of follow-on entry facilitates cumulative innovation over time. For PI-based products and services that are affected by these barriers to entry, successful follow-on innovation will be much harder and cumulative innovation may not occur. For example, it is not enough for a later entrant to create an improved search engine algorithm that users would prefer, *ceterus paribus*. To compete with the first entrant's search engine, the improved design must be so much better that users value the improved design, as implemented with little or no personal data, more than they value the first entrant's design "souped up" with all the personal information the first entrant has collected. Moreover, even this sort of leapfrogging innovation may be possible only if a trove of personal information was not needed as a tool for developing the improved design.[189]

Even if a second comer manages to come up with an improvement significant enough to overcome the barriers to later entry, there may be a risk of what we view as hyper free riding. Unless the second comer's follow-on design is patented or otherwise protected by intellectual property, the first entrant may be able to copy the follow-on design (using personal information if necessary) and then enhance its value using the trove of personal information already at hand. Given this situation, a second comer might simply try to sell rights to its follow-on innovation to the first entrant. Because of its market dominance, the first entrant

---

that different data analytical tools can create divergent economies of scale." *Id.* at 354.

[189] *See* Cédric Argenton & Jens Prüfer, *Search Engine Competition with Network Externalities*, 8 J. COMP. L. & ECON. 73 (2012).

might or might not find it profitable to bother purchasing rights to the follow-on innovation, depending on details of its business model. If it does, the business's customers might benefit from the improvement, but, if anything, the purchase will only exacerbate the barriers to further competitive innovation.

To summarize, innovation in PI-based products and services, while unlikely to be plagued by free rider problems that cannot be handled by intellectual property, will often be affected by failures of appropriability related to trade secrecy and associated barriers to entry. These failures will distort the appropriability landscape. As a result, incentives for investments in innovations that can take advantage of the low costs of acquiring personal information as a byproduct or of barriers to entry will be high relative to consumer demand. Conversely, incentives for follow-on innovations or competitive alternatives will be low relative to consumer demand.

V. DESIGNING PRIVACY REGULATION WITH INNOVATION IN MIND

In Part III, we explained why misaligned demand is likely in PI-based markets, thus making a prima facie case for information privacy regulation. Part IV explained why failures of appropriability are also likely to occur. This Part pulls these analyses together to explore the interaction between privacy regulation and personal information-based innovation and discuss its implications for the design of information privacy regulation.

A. Information Privacy Regulation: Is the Game Worth the Candle?

As outlined Part III, there are many reasons to anticipate misalignment between market demand signals and individual and societal preferences in PI-based markets. As a result, the current situation is almost certainly rife with market failures that are directing innovative activity along socially undesirable paths. We should seriously consider regulation precisely because of the importance of innovation—and its path dependence. In previous work, summarized briefly in Section II.C, we argued that there are no general reasons to expect that well-designed regulation aimed at realigning demand with true preference will "stifle" innovation in a socially meaningful way.[190] Any socially problematic stifling of innovation is likely to result from errors in regulatory design that either exacerbate demand misalignment or impose unnecessarily

[190] Lev-Aretz & Strandburg, *supra* note 12.

high compliance costs on innovative activity. Designing regulation to avoid these pitfalls is a contextual task that depends on a particular regulation's goals and mechanisms. Of course, as we discuss in more detail elsewhere, regulatory processes can fail in many well-known ways: by being "captured" by improper influences, by incorrectly identifying failures of perceived demand, by inaccurately assessing and predicting the costs and benefits of various regulatory designs, and so forth. Sometimes the best of even the best regulatory design for addressing a particular failure may be outweighed by its costs. While regulatory failure is a real concern, its mere possibility cannot be an automatic deal breaker because of the deleterious effects of allowing market activity to be governed by demand signals that are significantly misaligned with social value. The lesson, instead, is that regulatory design is a serious matter, to be undertaken with care. Similarly, any fears that regulation will unintentionally and unpredictably suppress high social value innovation in the long run must always be weighed against the long-term social costs of the unregulated demand portfolio's misalignment with social value.

We also see no theoretical reasons to expect market failures associated with information privacy to be uniquely impervious to regulatory intervention or unusually susceptible to regulatory design failures. And it is much too soon to give up on the task. Unlike the design of environmental regulation, which has been the subject of substantial in-depth consideration by academics and policymakers, the detailed study of privacy regulation mechanisms and design possibilities from a social welfare perspective is in its infancy. As yet, there has been relatively little scholarly or policy attention paid to creative regulatory design. For example, the European GDPR is bold in its adoption of more serious penalties and its attempt at uniform applicability.[191] Some of its provisions may turn out to be novel (depending on how they are eventually interpreted). Nevertheless, at its heart the GDPR is founded on regulatory principles and mechanisms developed before the digital age. While the GDPR represents a step forward, when compared with US privacy regulation, which is mostly out of date and overly reliant on consumer consent, there is no reason to view the regulations currently in force anywhere as the be-all and end-all of

[191] Although these large penalties are criticized for harming small players and not effectively deterring bigger players. *See* Mike Gillespie, Sharon Klein & Luke Scanlon, *Q&A: Managing Data Privacy and Cyber Security Risks for Private Equity Funds An interesting,* FINANCIER WORLDWIDE (Sept. 2015), https://www.financierworldwide.com/qa-managing-data-privacy-and-cyber-security-risks-for-private-equity-funds/#.XGWp4c9KjBJ. If these claims are true, it is a clear case of flawed regulatory design.

information privacy regulation design. There is much more work to be done.

We also do not think that the few attempts so far to study the impact of privacy regulation on innovation empirically provide cause to abandon the enterprise. There are very few studies on the interplay between privacy regulation and innovation. These studies do not reach any uniform consensus that privacy regulation reduces innovation. They all suffer from the usual difficulties of finding metrics for innovation and controlling for external factors.[192] Some have not distinguished *shifts* in innovative activity from *overall* decreases in innovation, which is, in fact, quite difficult to do. A 2012 work, titled "*Privacy and Innovation*," by Avi Goldfarb and Catherine Tucker is an exception, in that it took shifts in innovative direction into account. Goldfarb and Tucker showed that privacy regulations have directly affected usage and efficacy of emerging technologies in the online advertising and health care sectors.[193] As these impacts are heterogeneous across firms and products (meaning that privacy regulation could both advance and deter innovation) Goldfarb and Tucker concluded with a neutral, and in our view accurate, observation that privacy regulations directly influence the direction and rate of data-based innovation. [194] Most important, these empirical studies are unavoidably dependent on the design of currently enacted privacy regulations and cannot directly measure the effects of regulatory

---

[192] *See, e.g.*, Richard G. Newell, *The Role Of Markets And Policies in Delivering Innovation for Climate Change Mitigation*, 26:2 OXFORD REV. ECON. POL'Y, 253, 255-60 (2010) (describing disagreements among researchers as to acceptable measures of innovation and envoiromental regulation impact); Daniel Kammerer, *The Effects of Customer Benefit and Regulation on Environmental Product Innovation—Empirical Evidence from Appliance Manufacturers in Germany*, CTR. FOR COMP. & INT'L STUD. (ETH Zurich and University of Zurich) working paper, at 3 (2008) (noting that "[r]esearchers of business strategy and public policy have analyzed the relationship between regulation and environmental innovation in numerous studies. While qualitative case studies… are based on rather unsystematic analysis of anecdotal evidence, more systematic econometric studies often use indicators that are too simple."); Zarsky, *supra* note 59, at 128 (arguing that "innovation is not an abstract notion but a measureable element," but conceding that it is in "the form of measurement is contestable."). Zarsky also highlights the multifaced success or failure of information technology innovation, which comolicated the difficulty of measuring such innovation even further: "The extent of ICT innovation is obviously measured to examine innovation in the ICT industry . . . . Yet ICT innovation (or lack thereof) is considered indicative of innovation in other, related sectors (such as financial sectors) whose innovation is both reflected and caused by ICT growth." *Id.*, at 128 n.53).

[193] Goldfarb & Tucker, *supra* note 47, at 85.

[194] *Id*.

designs that have not been enacted.[195] It is thus difficult to use the empirical data for comparative analysis of alternative regulatory designs.

In sum, we see no reason to anticipate that attempts to design information privacy regulation are categorically, or even unusually, prone to fail in ways that will stifle innovation across the board. Currently enacted privacy regulations, despite their flaws, have already stimulated innovation aimed at reducing compliance costs, meeting consumer demand for privacy-protective technology and re-aligning innovation in line with the normative preferences expressed in privacy regulation. Moreover, the design of privacy regulation mechanisms is in its infancy. Even if current regulatory approaches raise serious concerns about innovation, it is far too early to fold up and go home. Given the many reasons to believe that unregulated markets will fail to serve individual preferences and social values, the better response is to give more serious attention to designing and evaluating specific privacy regulations.

B. Privacy Regulation and Failures of Appropriability: Regulatory Design Considerations

As discussed in Part V, the personal information-based market is prone to systematic appropriability failures that often are not remedied by intellectual property. When these appropriability failures are correlated with individual and social preferences or with the compliance costs of a particular regulatory proposal, appropriability effects should be considered explicitly in regulatory design. To help in thinking through how these issues apply to personal information-based markets, this section applies this Article's framework of demand-realignment and appropriability failure to three hypothetical information privacy regulations.

a. Hypothetical One: Restricting Retention of Search Engine Data

Our first hypothetical regulation forbids search engine providers from retaining consumer data for more than two weeks. Assume, for purposes of illustration, that this hypothetical

---

[195] *Cf.* Mark Pettigrew et al., *Natural Experiments: An Underused Tool for Public Health?*, 119 PUB. HEALTH 751, 756 (2005) (noting that the "naturalness" of natural experiments, i.e. that their reliance on enacted policies has the potential to introduce bias).

regulation implements true social preferences about the trade-off between the marginal improvement in personalized search results and other advantages to users that can be expected from longer-term data collection and the risks of longer data retention in terms of data breaches, advertiser manipulation, government snooping, and the like.

This regulation would effectively reduce market demand for innovation involving the exploitation of long-term search data, whether for search personalization or for ad targeting, making innovative activity aimed at making better use of short-term search data, improving presentation of search results and combining short-term data with contextual advertising more attractive. As a side effect, this regulation would reduce the appropriability of search engine innovation for dominant providers and decrease upfront costs for competitive and follow-on search engine innovation. This regulation thus appears to tap into a correlation between social preferences and appropriability failure that makes it a winner on both fronts, much like the hypothetical regulation favoring solar panels over coal plants that we discussed in Part III.

Currently, popular search engines, such as Google, are ad-supported businesses that rely on collecting troves of personal information and using them to target advertising. However, we see little reason to fear that this regulation would cripple the search engine business. Unlike broadcast programming, the search engine business does not lack means to collect payment directly from users, for example through subscription or "freemium" models. Moreover, search is a highly valuable service, both socially and individually. There is no reason to expect that consumers will not be willing to pay enough for this valuable service to cover its operating costs. It is true that consumers are used, by now, to getting search services for free, and thus might protest, at least at the transition phase, about paying for search. However, consumers are also very accustomed to—even dependent on—using search services. As a result, if presented with the choice of paying for search engine services and not having them at all, we assume that the vast majority would be willing to pay some non-trivial price.

In the highly unlikely scenario in which aggregate user willingness to pay is insufficient to cover operating costs—potentially because of positive externalities—standard regulatory solutions to externality failures, such as tax credits and subsidies, could be applied. A more serious concern is distributive—some individuals might be unable to afford the market rate for search services. But this problem is a subset of the larger problem of

disparate internet access and could be addressed similarly, for example by providing search engine access at educational institutions, libraries and in other public spaces or subsidizing access based on economic status.

Currently, of course, search engines almost universally adopt an advertising-based business model, rather than collecting payment from consumers. Would this hypothetical regulation sink the economic viability of ad-based search? Search-based advertising appears to be highly valued by advertisers, as reflected in price data. Presumably, current prices are driven partly by the value of ad-targeting services provided by ad-supported search engines using the personal information they collect. Assume for the sake of argument that the regulation's restriction of data retention makes this targeting less precise and thus reduce the prices that advertisers are willing to pay for search-based advertising.[196] Even given this assumption, it seems unlikely, as an intuitive matter, that advertising revenue would drop so low that it could not cover operating costs (which are now, if anything, lower given decreased data storage needs).

If this intuition is wrong and the regulation does undermine the ad-based business model for search engines we should expect one of two scenarios: Alternative, non-targeted advertising-based business models for search engines would emerge, or paid search would become common and acceptable, effectively offering a different revenue source for this market. The main point is that we don't ordinarily worry much about whether the market will manage to provide highly valuable private goods and services for which payment can be collected without high transaction costs. Such goods and services are the bread and butter of the market and search is one of them.

Perhaps the concern is instead that consumers will regret the regulation if the ad-based search business turns out not to be viable because the regulatory process did not adequately account for its possible disappearance. Given well-known behavioral biases that make free an irrationally sticky price point, we might well question whether consumers are likely to experience any such regret once they switch to paying for search on a simple

[196] In fact, there is at least some evidence that very recent search data is what really counts for targeting ads. Jun Yan et al., *How Much Can Behavioral Targeting Help Online Advertising?*, 2009 WWW CONF. PROC. MADRID 261-62, available at http://www.wwwconference.org/www2009/proceedings/pdf/p261.pdf; *see also* Strandburg, *supra* note 75, at 104-05.

subscription model right along with their monthly bills for telephone, electricity, gas, Internet, rent and so forth. But assume consumers really do have preference for ad-supported search. Why would this preference, along with the possibility that the regulation would sink the ad-supported business model for search engines, not have been considered in regulatory design and evaluation? Of course, regulatory design failure is always a possibility, but as we explained at the outset, we assume here that socially beneficial regulation is possible. There seems to be nothing about this particular regulation that raises red flags. If anything, given the bias already mentioned, we might worry that the regulatory design process will over-value free ad-supported search in comparison with paid models. [197]

Should we worry, instead, that this regulation will stifle innovation by swinging the appropriability pendulum so far back that free rider problems re-emerge? This outcome also seems unlikely. Regardless of whether users or advertisers are paying, search engine innovators can take advantage of standard intellectual property mechanisms for recouping upfront investments. The software and data that create search results are protectable by trade secrecy, copyright, and, to some extent, patents. Those protections can be leveraged into higher prices in the usual way.

Compliance costs for this regulation are unlikely to add significantly to the upfront costs of innovative activity. The regulation tamps down the over-compensatory tendencies of trade secrecy in personal information-based markets, but there is no special reason to anticipate that standard IP balancing will fail to deal with free rider issues. To the contrary, some barriers to entry are likely to persist despite this regulation. Search engines are likely to perform better when they draw on data from more users, creating network-like effects. These network-like effects create value for users without the risks associated with long-term collection and retention. On balance, society might prefer to put up with these remaining barriers to entry.

The regulation's net effects on the appropriability side thus depend on whether it lowers barriers to entry enough to induce competitive and follow-on innovation. If so, it is a win for society on both sides of the equation. But even if persistent network effects continue to dampen competitive and follow-on innovation, the

[197] This is not even to mention the lobbying power of companies engaged in advertising-based business models relying on personal information collection.

regulation is likely to be a net social winner because the market's portfolio of goods, services and innovative activity will be better aligned with social value.

b. Hypothetical Two: Collection Restriction

Now let's imagine a different information privacy regulation that is designed to reduce the potential harms from data breaches by imposing restrictions on the transfer of personal information from one entity to another and cybersecurity requirements, but without enacting any limitations on data retention or use by collecting entities. Assume, in particular, that this regulation substantially restricts the sale of personal information by data brokers. Assume that this regulation is correctly designed in that it does reduce the risk of data breaches and does not introduce other sorts of misalignments with individual and social preferences regarding personal information collection, flow and use.[198] This is the sort of restriction that traditional usage of the term "privacy regulation" might bring to mind, in that it simply reduces the flow of information among entities. This sort of regulation shifts the market's demand signals away from products and services that require the purchase or transfer of personal information relative to products and services that either i) do not rely on personal information or ii) rely primarily on using personal information that they collect themselves. We assume, for purposes of this hypothetical, that the resulting market demand for various goods and services is better aligned with individual and social preferences.

Now consider the appropriability implications of this regulation. Businesses that never relied on databases of personal information are in much the same position as before. Dominant players in markets rely on byproduct collection of personal information will now be protected by more impervious barriers to entry, however, even though the direction of their innovative activity may be affected by their inability to purchase additional data. Correspondingly, potential competitive or follow-on innovators will face even greater hurdles than before. Previously, they could have purchased personal information from data brokers and used it in developing and improving their innovative goods and services. Though not a substitute for the troves of information collected by dominant players, this information might have

[198] This could be a fairly strong assumption depending on details of the "restrictions" imposed on transfers. We make the here in order to focus on the implications for innovation.

reduced the degree they would have had to leapfrog existing offerings to enter the market successfully.

Thus, the realignment of demand by this regulation is likely to exacerbate the failures of appropriability already affecting the personal information-based market. The net result may be a portfolio of innovative activity that produces fewer data breaches, but also induces more innovation in goods and services that vacuum up personal information from users than is socially desirable. The negative social implications could be even more pronounced if the innovation induced by these appropriability failures aims primarily at employing personal information for more effective ad targeting and or for uses that are socially problematic for distributional or other normative reasons.

This hypothetical regulation thus exemplifies the sort of information privacy regulation that might look good from a regulatory perspective focused on market failures resulting from misaligned demand. When we account for correlations between consumer preferences and appropriability failures, however, the picture looks very different. The bigger picture raises red flags that should be seriously considered before adopting this hypothetical regulation.

c. Hypothetical Three: Mandated Data Sharing

Finally, we consider a hypothetical regulation that would mandate that dominant players who collect personal information as a byproduct of providing goods and services must share some part of the data they collect with other market players, perhaps at some price. This hypothetical is based on proposals in the literature.[199] These proposals rest on the assumption that data aggregation produces positive network-like effects, somewhat as described in our discussion of the search engine hypothetical. Under this assumption, the primary downside of large aggregations of personal information is that data aggregation also tends to produce barriers to entry. This sort of hypothetical regulation operates in a similar vein to regulations that attempt to preserve the positive benefits of network effects while increasing competition by regulating interconnectivity, imposing standards, and so forth. No

[199] Under our definition, despite the intuitive discordance, such a regulation would qualify as an information privacy regulation because it constraints the flow of personal information. We think that looking at this hypothetical in the same framework as more obviously "privacy"-protecting regulatory designs demonstrates the value of this broad definition.

doubt these proposals do help to level the appropriability landscape.

This approach to leveling the appropriability landscape has serious implications for personal information flow. Proponents of these proposals suggest that "privacy" can be taken care of separately at the back end as a kind of afterthought. Juxtaposing this sort of proposal with the hypothetical "no data transfer" regulation demonstrates the fallacy of this approach. Personal-information-based markets have a wide variety of likely sources of misaligned demand, a tendency toward failures of appropriability. As we have seen, the very nature of personal information-based markets is that these failures of appropriability tend to be correlated, whether for better or for worse, with individual or social preferences regarding information collection, flow, and use. The proposal to correct appropriability failures by sharing personal information of all sorts more widely among market players of all sorts seems almost designed to clash with the goal of realigning market demand with individual and social preferences about information flow and use, which tend to be highly contextual. Because correlations between preference for information flow and appropriability failures are common, it is highly unlikely that these issues will be amenable to separate or sequential regulatory design.

VI. CONCLUSION

The market's portfolio of innovative activity reflects suppliers' perceptions of market demand mixed with their expectations of appropriability. Regulation's traditional goal is to bring market demand into better alignment with individual and social preferences and values can, while intellectual property law (and, at times, competition law) aims to bring suppliers' incentives into line with those preferences by smoothing out the appropriability landscape. In many contexts, these tasks are mostly separable: regulatory design need not pay much attention to appropriability, while intellectual property doctrine assumes that market demand correctly reflects consumer preferences. Our analysis points out that this implicit assumption of separability is not valid in PI-based markets.

Both misaligned demand signals and appropriability failures are common in PI-based markets because of certain characteristics of personal information and common features of personal information-based markets. Moreover, the sources of appropriability failures in these markets mean that appropriability failures are often correlated with individual and social preferences regarding personal information collection, flow and use. As a

result, it is often important to take both into consideration when designing information privacy regulation because regulation can sometimes exacerbate—and sometimes alleviate—appropriability failures.

These interdependencies do not, however, support the sweeping claim against information privacy regulation because of innovation stifling. Indeed, information privacy regulation can sometimes enhance innovation through its collateral effects on appropriability, as we show in the above hypotheticals. The social implications of the regulation depend on the particular regulatory context and on regulatory design specifics. If anything, nuanced regulatory design and evaluation is especially important for information privacy regulation. The goals of information privacy regulation will vary greatly because individual preferences and social values that define them are highly context dependent. The interplay between demand realignment and appropriability in PI-based markets adds to this contextual complexity and makes nuanced and careful regulatory design and analysis more important.