

VALUING SOCIAL DATA

Amanda Parsons* & Salomé Viljoen**

Social data production—accumulating, processing, and using large volumes of data about people—is a unique form of value creation that characterizes the digital economy. Social data production also presents critical challenges for the legal regimes that encounter it. This Article provides scholars and policymakers with the tools to comprehend this new form of value creation through two descriptive contributions. First, it presents a theoretical account of social data, a mode of production that is cultivated and exploited for two distinct (albeit related) forms of value: prediction value and exchange value. Second, it creates and defends a taxonomy of three “scripts” that companies follow to build up and leverage prediction value and explains their normative and legal ramifications.

Through the examples of tax and data privacy law, the Article applies these descriptive contributions to demonstrate how legal regimes fail to effectively regulate social data value creation. Tax law demonstrates how legal regimes historically tasked with regulating value creation struggle with this new form of value creation. Data privacy law shows how legal regimes that have historically regulated social data struggle with regulating data’s role in value creation.

The Article argues that separately analyzing data’s prediction value and its exchange value is helpful to understanding the challenges the law faces in governing social data production and its surrounding political economy. This improved understanding will equip legal scholars to better confront the harms of law’s failures in the digital economy, reduce legal arbitrage by powerful actors, and facilitate opportunities to maximize the beneficial potential of social data value.

* Associate Professor, University of Colorado Law School.

** Assistant Professor of Law, University of Michigan Law School. The authors extend their sincere thanks to Ashraf Ahmed, José Argueta Funes, Oren Bracha, William E. Forbath, Don Herzog, Sari Mazzurco, Katharina Pistor, Angie Raymond, Neil Richards, Felix Wu, Tim Wu, and participants at the 2024 University of Michigan Law School Legal Theory Workshop, the 2023 University of Texas School of Law Faculty Colloquium, the 2023 Schwartz Reisman Institute at the University of Toronto Seminar Series, the 2023 LPE of Tech Workshop, the 2023 Ostrom Workshop, the 2023 Privacy Law Scholars Conference, the 2023 NYU Privacy Research Group, the 2022 Joel R. Reidenberg Northeast Privacy Scholars Workshop, and the 2022 Columbia Academic Fellows Workshop for helpful and thought-provoking conversations and comments. We are grateful to the editorial staff at the *Columbia Law Review* for their helpful suggestions and edits, especially Zakiya Williams Wells. Finally, we would like to thank our new co-authors, Henry (born December 2023) and Aviva (born March 2024), who did not write a word but were everything nonetheless.

INTRODUCTION	995
I. PREDICTION VALUE AND THE DATA POLITICAL ECONOMY.....	1005
A. What Is “Value”? A Quick Background.....	1005
B. What Is Prediction Value?.....	1009
1. Defining Social Data.....	1010
2. Defining Prediction Value.....	1013
3. Prediction Value <i>Versus</i> Exchange Value.....	1016
4. Prediction Value <i>Into</i> Exchange Value?.....	1018
C. Why Datafy? And Why Take Data Value Seriously?	1019
II. THE BUSINESS OF SOCIAL DATA.....	1022
A. The Three Scripts	1023
1. Script One: Directly Converting Prediction Value Into Exchange Value	1023
2. Script Two: Indirectly Converting Prediction Value Into Exchange Value	1026
a. Lowering Costs	1026
b. Increasing and Stabilizing Revenues.....	1029
c. Expanding Business Operations.....	1032
3. Script Three: Converting Prediction Value Into Economic and Political Power	1035
B. The Business Models and Practices of the Digital Economy.....	1045
1. The Business of “Free”	1045
2. Building Ecosystems	1048
3. Aggressive Acquisitions.....	1050
III. LEGAL COLLISIONS	1054
A. Two Camps of Legal Collisions	1054
B. Taxing Prediction Value	1054
1. A Conceptual Collision	1055
2. Colliding With the Digital Economy’s Scripts.....	1059
a. Income as a Tax Base.....	1059
b. Tax Deferral Opportunities.....	1060
c. International Tax Implications	1061
C. Governing Social Data Value	1064
1. Privacy Law Background	1065
2. Data Privacy as a Value Regulation Regime	1068
a. Market Power and Political Power: Knitting the Two Accounts Together.....	1070
b. Demand-Side Versus Supply-Side Data Regulation	1071

3. Promising Horizons for Regulating Social Data Value in Data Privacy (and Beyond?).....	1074
CONCLUSION.....	1076

INTRODUCTION

Anna is sixteen weeks pregnant.¹ Anna uses Pineapple, a popular fertility and pregnancy tracking and social media app. In the app, Anna inputs information about her ovulation cycle, pregnancy symptoms, sleep patterns, eating habits, exercise, and moods. Anna also consumes Pineapple’s content on pregnancy and fetal development and engages with other Pineapple users in their forum, swapping questions on pregnancy and preparing for a new baby.

Pineapple collects, aggregates, and synthesizes data that Anna and other users share. This data includes not only the information that Anna inputs in the app but also data about the content she consumes and her interactions with other users. Pineapple does not sell this data directly—in fact, their privacy policy explicitly states that they will never sell or license individual user data. But Pineapple does sell insights about their user base as a whole to clients like advertisers, employment agencies, and consumer credit agencies. This is how Pineapple makes its money—the app is free to Anna. Pineapple’s clients then combine the data they receive from Pineapple with data from other companies to build out a more complete picture of the behavior of pregnant people. This could include data on TV viewing patterns from video streaming platforms, movement and sleep patterns from wearable fitness devices, or online purchasing behaviors.²

1. This example is adapted from various presentations made by Salomé Viljoen in 2022 and 2023, including at the 2022 Annual Meeting of the American Society of International Law, Seventh Annual Detlev Vagts Roundtable. Salomé Viljoen, Remarks by Salomé Viljoen, 116 Proc. Am. Soc’y Int’l L. Ann. Meeting 122, 122–25 (2022).

2. While this is a hypothetical example, it draws on real uses of social data. See, e.g., Hooman Mohajeri Moghaddam, Gunes Acar, Ben Burgess, Arunesh Mathur, Danny Yuxing Huang, Nick Feamster, Edward W. Felten, Prateek Mittal & Arvind Narayanan, *Watching You Watch: The Tracking Ecosystem of Over-the-Top TV Streaming Devices*, 2019 Proc. Ass’n for Computing Mach. Conf. on Comput. & Commc’ns Sec. 131, 142 (studying two thousand “over-the-top” streaming channels, finding widespread user tracking and data collection with little recourse for consumers to disable tracking through countermeasures); Tong Yan, Yachao Lu & Nan Zhang, *Privacy Disclosure From Wearable Devices*, 2015 Proc. Ass’n for Computing Mach. Workshop on Priv.-Aware Mobile Computing 13, 18 (studying how aggregated data from fitness trackers can be used to infer a user’s behavioral patterns, such as when they will go grocery shopping, get coffee, or work out); Jonah Engel Bromwich & Jessica Testa, *They See You When You’re Shopping*, N.Y. Times (Nov. 26, 2019), <https://www.nytimes.com/2019/11/26/style/powerfront-software-e-commerce-cartoons.html> (on file with the *Columbia Law Review*) (describing how e-commerce customer service representatives can now visualize emotional profiles of customers visiting their sites or using support chats); Aljoscha Dietrich, Kurunandan Jain, Georg Gutjahr, Bianca Steffes

Becca has never used Pineapple. But she does watch streaming services, owns a wearable fitness device, and shops online. And Becca's behavioral patterns on these platforms have shifted in similar ways to Anna's and other Pineapple users' behaviors. Pineapple's clients can, therefore, infer that Becca is also likely pregnant and treat her accordingly.

Why do companies care about Anna's and Becca's pregnancy status? Because early pregnancy data is incredibly valuable. Pregnancy signals that a consumer is about to undergo a significant change in their daily habits and their buying activities; the birth of a child is a time when someone's buying habits, brand loyalties, and daily routines are in flux. Getting to such consumers early is a valuable opportunity to shape their future purchasing behaviors. Diaper companies can advertise to Becca or Anna before competitors and get them locked into their brand. Grocery stores, online subscription services, car manufacturers, and others can also reach out, offering deals favorable to new parents that entice them to switch entrenched behaviors and brand loyalties. Aggregate pregnancy data also provides an opportunity to understand the nature of consumer change more generally—how and why do consumption patterns change? When are such changes most robust, and why? How can you predict (and modify) those behavioral changes?

Data about Anna's and Becca's pregnancies is what this Article calls *social data*.³ Social data refers jointly to two interrelated types of data about people. The first is data that directly materializes and stores traces of human activity.⁴ This includes, for example, information on Anna's or Becca's TV viewing patterns, ovulation, or movement. This type of social data is directly collected from data subjects, like the data Pineapple collects and uses about Anna. The second is data that is used to apprehend, infer, or predict human activity.⁵ For example, Pineapple collects data about Anna (and other users) to aggregate and analyze for insights about pregnant people as a group, which it sells to third parties. Those third parties may use this data in turn to gain insight about, and drive decisions regarding, Becca. Thus, data about Anna and her pregnancy is *also* data

& Christoph Sorge, I Recognize You by Your Steps: Privacy Impact of Pedometer Data, *Computs. & Sec.*, Jan 2023, no. 102994, at 1, 7 (demonstrating how fifteen minutes of pedometer data alone—data collected on nearly every smartphone—can identify users).

3. Part I provides a more detailed definition of this concept.

4. This first category is adapted from Julie E. Cohen's concept of the "data refinery"—the data-processing practice of "refin[ing] and massag[ing] flows of personal data to produce virtual representations . . . optimized for modulating human behavior systematically"—as a "centrally important means of economic production." Julie E. Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* 66–68 (2019) [hereinafter Cohen, *Between Truth and Power*].

5. Part I describes the difference between the two below. But for now, it is important to note that data used to infer or predict human activity need not always derive from data directly about human activity. For further discussion of the significance of the second category, see Alicia Solow-Niederman, *Information Privacy and the Inference Economy*, 117 *Nw. U. L. Rev.* 357, 400–03 (2022).

about Becca's pregnancy, even though it was not directly collected from Becca.⁶ Indeed, data that can be used to infer or predict human activity need not be collected from people at all—for example, data about weather can be used to predict and infer commuting behavior. In contrast with the more commonplace term “personal data,” “social data” nicely expresses the view (and a central focus of this Article) that data is socially useful and economically valuable—not only for what it can tell the world about any one person, like Anna, but also, and especially, for what it can tell the world about people.⁷

The value of Anna's and Becca's social data is what this Article refers to as *prediction value*.⁸ Prediction value is a particular form of use value that lies in social data's capacity to infer or predict things about people—in this case, pregnancy status—and to act on that knowledge. For example a firm with access to Becca's social data may send Becca a diaper coupon or free prenatal vitamins, a hospital where Anna will give birth may use it to inform labor and delivery staffing plans, or an employer's hiring algorithm may flag Becca as a potentially risky and expensive hire and exclude her from a pool of prospective employees. Social data stores the value of being able to apprehend behavior, to infer, predict, and direct the future actions of people (who are not always the data subject), and to develop informed strategies to obtain some objective. It provides the valuable capacity to exert some measure of insight into and control over future behavior.⁹ The capacity of social data to store insight into human behavior, guide predictions about that behavior, and optimize strategies to guide and change human behavior is (much of) what drives companies to collect the data they collect and use the data in the way that they do.¹⁰

6. See Salomé Viljoen, *A Relational Theory of Data Governance*, 131 *Yale L.J.* 573, 606–08 (2021) [hereinafter Viljoen, *Relational Theory*] (describing a hypothetical scenario in which a tattoo AI company creates an algorithm based on the dataset of the social media company it acquires that allows data collected from one person to be used to infer information about a second person).

7. *Id.* at 609–11. For examples of the ubiquity of the term “personal data” in common references to data privacy, see, e.g., Colleen McClain, Michelle Faverio, Monica Anderson & Eugenie Park, *Pew Rsch. Ctr., How Americans View Data Privacy* 12, 18 (2023), <https://www.pewresearch.org/internet/2023/10/18/views-of-data-privacy-risks-personal-data-and-digital-privacy-laws/> [https://perma.cc/M82M-BATC] (using the term “personal data” frequently to refer to aggregated, predictive insights); Hossein Rahnama & Alex “Sandy” Pentland, *The New Rules of Data Privacy*, *Harv. Bus. Rev.* (Feb. 25, 2022) <https://hbr.org/2022/02/the-new-rules-of-data-privacy> (on file with the *Columbia Law Review*) (describing “personal data” as individual data points aggregated to generate broader consumer insights).

8. Part I provides a more detailed definition of this concept.

9. See David Graeber & David Wengrow, *The Dawn of Everything* 364–65 (2021) (arguing that “control of information” is one of three bases of social power).

10. Part II covers this at length. For a specific example relevant to the scenario here, see Eric Siegel, *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die* 6–12 (rev. ed. 2016) (detailing the variety of predictive models organizations employ across various sectors, including healthcare, retail, and the mortgage industry, to attract, retain,

Social data cultivation is key to the business strategies of some of the wealthiest and most powerful companies operating today. Companies face generalized market pressure to engage in the accumulation and cultivation of social data and its prediction value to stay competitive.¹¹ Indeed, the widespread practice of treating social data as a key input to production is part of what it means to refer to contemporary capitalism as an informational capitalism.¹² Recent technological transformations, like improved chip processing power, ubiquitous connected devices like smartphones, and improvements in machine-learning techniques, have all contributed to the feasibility and utility of entities cultivating, refining, and extracting social data value.¹³ These technological changes have allowed entities to exploit for economic gain what has long been true: People are social beings, deeply knowable and materially influenced by relations to one another. Thus, the stakes of understanding social data's particular form of value, and the social and economic effects that its widespread cultivation produces, have grown more salient.

A primary way the digital economy works is by using prediction value to increase monetary value: to grow profits by raising revenue and by lowering costs, or to grow market share (and, the thinking goes, future profits) by expanding customer bases and entering new markets.¹⁴ As Part II will survey in greater detail, companies deploy a variety of strategies to transform prediction value into *exchange value*—the priced, monetary value of a good, service, or company, typically expressed as a “market price.” Exchange value, as a general theory and form of value, posits that the value of a thing is the value derived from its exchange, expressed via

and win back customers). Specifically, Eric Siegel covers Target's use of big data to identify pregnant customers: Target created a training model based on users who signed up for Target's baby registry and then applied it to customers who had not registered. *Id.* at 48–50.

11. See Cohen, *Between Truth and Power*, *supra* note 4, at 6 (“In a regime of informational capitalism, market actors use knowledge, culture, and networked information technologies as means of extracting and appropriating surplus value, including consumer surplus.”).

12. See Viljoen, *Relational Theory*, *supra* note 6, at 577, 586 (“Data plays a central role in both descriptive and critical accounts that characterize the contemporary digital political economy as informational capitalism.”).

13. See Beth Gutelius & Sanjay Pinto, *Ctr. for Urb. Econ. Dev., Pain Points: Data on Work Intensity, Monitoring, and Health at Amazon Warehouses 20* (2023), https://cued.uic.edu/wp-content/uploads/sites/219/2023/10/Pain-Points_Final_Oct2023.pdf [<https://perma.cc/WQ2K-8V9V>] (describing the impact on workers of the “company's system of technology-enabled workplace monitoring”); Valerio De Stefano & Simon Taes, *Algorithmic Management and Collective Bargaining*, 29 *Transfer* 21, 23–25 (2023) (discussing how employers use data gathered from various technological tools to track workers' physical locations, mental and emotional states, and digital activity to make decisions and predictions regarding workers' conduct and productivity); Matt Burgess, *All the Data Apple Collects About You—And How to Limit It*, *Wired* (Jan. 16, 2023), <https://www.wired.com/story/apple-privacy-data-collection/> (on file with the *Columbia Law Review*) (describing all of the ways that Apple products may track and collect data about how users interact with their products).

14. Sections II.A and II.B provide a detailed analysis of these profitmaking strategies.

price, on a real or imagined market.¹⁵ So, prediction value can be—and is—transformed into exchange value. But it doesn't have to be. Prediction value is distinct from (and not always neatly transformed into) exchange value. Part I provides greater detail defending the descriptive and analytic virtues of cataloging this distinction.

Prediction value confers on its holder the power to apprehend, shape, and thus exert some measure of control over people's behavior. In fact, the central preoccupation of privacy scholars and many other observers of the digital economy is this potential for the control power of social data, cultivated for its conversion into priced value, to be repurposed toward other (potentially disempowering) ends.¹⁶ These purposes can coexist with strategies to grow exchange value, such as the use of prediction value in labor settings to reduce operating costs by eroding workplace protections, or lie outside the commercial realm entirely, such as immigration officials repurposing location data cultivated for commercial ends to detect and detain suspected undocumented immigrants.¹⁷ Indeed, much of privacy law's traditional concern regarding privately cultivated surveillance capacities is how such capacities fall into the hands of state actors and empower state action without sufficient scrutiny.¹⁸

Of course, there is also some amount of speculative behavior around prediction value, as when entities, in order to secure valuations of high exchange value, overclaim or overpromise on the prediction value their products can deliver—a phenomenon the computer scientists Sayash Kapoor and Arvind Narayanan refer to as “AI snake oil.”¹⁹ But this, too, highlights the importance of disentangling assessments of social data value from priced exchange value—to better identify when claims of social data value (and its potential to transform into priced exchange value) are

15. See Dave Elder-Vass, *Inventing Value* 1, 22 (2022) (noting the mainstream notion of value as the equilibrium price of demand and supply and arguing that equilibrium price is “not the same thing as actual price[]” but a “regulative concept: the notional price at which marginalist theory says goods ought to be exchanged”); see also R.H. Coase, *The Nature of the Firm*, 4 *Economica* 386, 388 (1937) (“Outside the firm, price movements direct production, which is co-ordinated through a series of exchange transactions on the market.”).

16. See *infra* section III.C.

17. Paul Blest, *ICE Is Using Location Data From Games and Apps to Track and Arrest Immigrants*, Report Says, *Vice News* (Feb. 7, 2020), <https://www.vice.com/en/article/v7479m/ice-is-using-location-data-from-games-and-apps-to-track-and-arrest-immigrants-report-says> [<https://perma.cc/C24R-GB9R>].

18. On use of data in workplace settings, see *infra* section II.A. On the traditional focus on state actors as a source of power-related privacy harm, see *infra* section III.C.

19. See Sayash Kapoor & Arvind Narayanan, *About the Book and This Substack, AI Snake Oil*, <https://www.aisnakeoil.com/about> [<https://perma.cc/SM2P-J4ET>] (last visited Feb. 19, 2024) (explaining the foundations of the AI snake oil project, which seeks “to dispel hype, remove misconceptions, and clarify the limits of AI”); see also Louise Matsakis, *The Princeton Researchers Calling Out ‘AI Snake Oil’*, *Semafor* (Sept. 15, 2023), <https://www.semafor.com/article/09/15/2023/the-princeton-researchers-calling-out-ai-snake-oil> [<https://perma.cc/XMU9-TTNP>].

overblown.²⁰ As Aaron Shapiro notes in his excellent work on gig platforms, when it comes to understanding the way platforms capitalize on prediction value by turning it into market valuation, there is a considerable “gap between what platforms do and what they say they do.”²¹ Clarifying the two modes of value production (and how they relate to each other) can help regulators and other observers traverse this gap and evaluate when claims are plausible and when they are not.

This Article argues for the importance of understanding how social data value is cultivated and used for regulating the digital economy. Part I provides greater detail on the concepts of social data and prediction value and argues for the distinctive value proposition of cultivating, accumulating, and using social data. It also provides theoretical context to distinguish the concept of exchange value—priced monetary value—from the concept of value more broadly and from prediction value as a particular kind of use value.

Part II offers a taxonomy of the business models and practices developed around cultivating and using social data value. This taxonomy divides the ways in which companies leverage prediction value to produce wealth and power for themselves and their investors into three scripts. The first script is direct and immediate conversion of social data value into exchange value through means such as direct sale of data, or through the premiums charged for targeted, as opposed to untargeted, advertising. The second script is indirect and often delayed conversion of prediction value into exchange value through improving and developing new products and services, lowering costs, increasing and stabilizing revenue, and expanding into new business lines and industries. The third script is leveraging prediction value to accrue power. This script catalogs how social data value can be a source of economic and political power, and thus of value to companies in their longer-term aims to secure market power and favorable regulatory environments. After cataloging and describing these

20. See Inioluwa Deborah Raji, I. Elizabeth Kumar, Aaron Horowitz & Andrew D. Selbst, *The Fallacy of AI Functionality*, 2022 *Proceedings Ass’n for Computing Mach. Conf. on Fairness, Accountability & Transparency* 959, 961–65 (describing various ways in which AI systems fail to produce certain claimed outcomes, whether because the objectives were impossible, the systems were designed in faulty ways, or the systems’ capabilities were falsified, misrepresented, or overstated); Angelina Wang, Sayash Kapoor, Solon Barocas & Arvind Narayanan, *Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms that Optimize Predictive Accuracy*, *Ass’n for Computing Mach. J. Responsible Computing*, Mar. 2024, no. 9, at 1, 8–16 (identifying specific shortcomings in datasets or modeling that interfere with the ability to optimize the predictive value of AI and machine-learning models).

21. Aaron Shapiro, *Platform Sabotage*, 16 *J. Cultural Econ.* 203, 204 (2023) [hereinafter Shapiro, *Platform Sabotage*]. Tim Hwang compares the behavioral advertising market to the subprime mortgage crisis of 2008, arguing that companies’ claims that behavioral advertising is more effective are, like the supposed value of subprime mortgage-backed financial products, empirically dubious. Yet, similar to the 2008 financial crisis, these empirically dubious value propositions nevertheless produce widespread social disruption as companies pursue them. Tim Hwang, *Subprime Attention Crisis* 76–92 (2020).

three scripts, the Article explores some specific business practices associated with following these scripts, each of which focus on growth and expansion. These practices include offering free and low-cost services, creating ecosystems of products and services, and embarking on aggressive merger and acquisition strategies. The Article shows how these strategies differ from traditional ones in ways that carry both legal and normative significance.

In Part III, this Article explores how disambiguating prediction value and exchange value (conceptually and normatively) can illuminate why such a variety of existing legal regimes fail to properly manage the social and economic disruptions that have accompanied capitalism's informational turn. In short, the same legal regimes that structure the transformation of prediction value into exchange value fail to grasp—in its entirety—the messy, imperfect, and socially disruptive process by which this transformation occurs. While the regulatory challenges of the digital economy increasingly place strain on various areas of law, most consider only small portions of this process and lack a systematic understanding of social data value production.²²

The Article identifies two contexts in which the legal regimes that structure this process index only part of its legally relevant features. The first context is legal regimes that have historically been tasked with governing value creation.²³ Such regimes are focused on evaluating and regulating companies' claims of exchange value and thus only apprehend or index prediction value (indeed, they only consider such value normatively and legally relevant) at the point it is transformed into exchange value. As section III.B will show, this can miss many legally salient features of prediction value, such as how it is cultivated and the wider social effects that cultivation creates. This leaves such regimes poorly equipped to properly achieve their normative goals. The Article chronicles these struggles through the example of tax law.

The second context is legal regimes that have not historically understood themselves to be tasked with governing value creation but that are focused on the legal significance of informational power. Such regimes are attentive to the capacity of information about people to create power over them, but they regulate social data along a strict public–private divide. Through the example of privacy and data governance law, the Article shows the conceptual and programmatic challenges of this approach. Privacy and data governance law govern private data collection primarily

22. Julie E. Cohen & Ari Ezra Waldman, Introduction: Framing Regulatory Managerialism as an Object of Study and Strategic Displacement, 86 *Law & Contemp. Probs.*, no. 3, 2023, at i, ii–iii (detailing the “long and growing” list of harms from informational capitalism as managed under current regulatory paradigms and noting that adequate responses to these harms are beyond the capacities of current regulatory approaches).

23. See *infra* Part III.

via individual control and consent rights.²⁴ Privacy law traditionally apprehends or indexes social concerns regarding prediction value, and its capacity to coerce action and remake social relations, only if or when it falls into the hands of public actors. And while the near-exclusive focus on state surveillance in the field is shifting, both popular and doctrinal conceptions of socially coercive privacy harm remain primarily focused on public, rather than private, actors. This ignores many salient concerns regarding informational power that arise as social data is imbricated into the strategies of commercial actors and neglects the role of privacy and data governance law in facilitating this form of value creation. It also overlooks the potential social benefits of prediction value if cultivated in procedurally fair ways and put toward collectively determined ends.

This analysis has broad implications for other areas of the law. For example, other legal fields that, like tax law, have historically been tasked with governing value creation have legal frameworks developed around the concept of exchange value and are not achieving their normative goals when applied to prediction value. Antitrust and financial regulation are prominent examples here, as there is growing evidence to suggest these regimes are struggling to index the profit-seeking behavior of technology companies and thus achieve these legal regimes' regulatory briefs in the digital economy.²⁵ This understanding will also be invaluable to legal fields that, like privacy and data protection law, have not historically been seen as regimes governing value creation and that, as a result, have not

24. This is also referred to as the “notice and choice” regime. See Neil Richards & Woodrow Hartzog, *Privacy’s Trust Gap: A Review*, 126 *Yale L.J.* 1180, 1197–98 (2017) (reviewing Finn Brunton & Helen Nissenbaum, *Obfuscation: A User’s Guide for Privacy and Protest* (2015)); Neil Richards & Woodrow Hartzog, *Taking Trust Seriously in Privacy Law*, 19 *Stan. Tech. L. Rev.* 431, 434 (2016). There are hundreds of privacy laws in the United States, and providing a systematic review of the role consent plays in each one is beyond the scope of this project. Yet, U.S. privacy laws are generally understood to derive from the Fair Information Practice Principles (FIPPs), which lay out principles for fair data processing, including meaningful individual rights of control and consent over data collection and use. Fed. Priv. Council, *Fair Information Practice Principles (FIPPs)*, <https://www.fpc.gov/resources/fipps/> [<https://perma.cc/BYS2-SY4F>] (last visited May 8, 2024). For some canonical examples of privacy laws that operationalize consent, see, e.g., *Children’s Online Privacy Protection Act (COPPA)* of 1998, 15 U.S.C. §§ 6501–6506 (2018); *Fair Credit Reporting Act (FCRA)* of 1970, 15 U.S.C. §§ 1681–1681x; *Family Educational Rights and Privacy Act (FERPA)* of 1974, 20 U.S.C. § 1232g (2018); *Health Insurance Portability and Accountability Act (HIPAA)* of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (codified as amended in scattered sections of 18, 26, 29, and 42 U.S.C. (2018)); *California Consumer Privacy Act (CCPA)* of 2018, Cal. Civ. Code §§ 1798.100–.199.100 (2023); *Biometric Information Privacy Act (BIPA)*, 740 Ill. Comp. Stat. Ann. 14/10, 14/15 (West 2023).

25. See Saule T. Omarova & Graham S. Steele, *Banking and Antitrust*, 133 *Yale L.J.* 1162, 1244–55 (2024) (describing the challenges faced by banking regulators striving to manage digital financial markets and financial tech companies); Sanjukta Paul, *Fissuring and the Firm Exemption*, 82 *Law & Contemp. Probs.*, no. 3, 2019, at 65, 72–76 (analyzing why antitrust law has failed to adequately regulate platforms like ride-hailing firms).

developed a positive agenda for regulating prediction value. First Amendment law is a prominent example.²⁶

The idea that information like social data confers power, and is thus a source of value with significant ontological, political-economic, and legal implications, is not new.²⁷ Others, particularly political economists of communication and historians of science, have long identified and analyzed the role of informationalism in contemporary capitalist value formation as it emerged and took on growing importance.²⁸ Previous work

26. Here, the Article refers particularly to discussions of free speech and the First Amendment as a primary legal regime regulating against the discursive and democratic ills of social media platforms. Similar to the developments in privacy and data governance law discussed in Part III, free speech scholars are increasingly attentive to the role free speech law plays in structuring the platform economy and the conceptual limits of that role. See, e.g., Jack M. Balkin, *To Reform Social Media, Reform Informational Capitalism*, in *Social Media, Freedom of Speech and the Future of Our Democracy* 233, 233–34 (Lee C. Bollinger & Geoffrey R. Stone eds., 2022) (arguing that, in order to restore a robust digital public sphere, people should focus on the “industrial organization of digital media and the . . . business models of social media companies” rather than focusing reform on First Amendment doctrines). There is, of course, a more general literature on free speech and the First Amendment that has argued that current free speech doctrine serves as a (Lochnerized) regulatory paradigm for economic activity. See, e.g., Amy Kapczynski, *The Lochnerized First Amendment and the FDA: Toward a More Democratic Political Economy*, 118 *Colum. L. Rev. Online* 179, 179–80 (2018), https://www.columbialawreview.org/wp-content/uploads/2018/11/Kapczynski-THE_LOCHNERIZED_FIRST_AMENDMENT_AND_THE_FDA_TOWARD_A_MORE_DEMOCRATIC_POLITICAL_ECONOMY.pdf [<https://perma.cc/6FZM-YZZ4>]; Robert Post & Amanda Shanor, *Adam Smith’s First Amendment*, 128 *Harv. L. Rev. Forum* 165, 165–67 (2015), https://harvardlawreview.org/wp-content/uploads/2015/03/vol128_PostShanor2.pdf [<https://perma.cc/88XH-LXMG>]; Jedidiah Purdy, *Neoliberal Constitutionalism: Lochnerism for a New Economy*, 77 *Law & Contemp. Probs.*, no. 4, 2014, at 195, 202–03.

27. In 1963, Kenneth E. Boulding remarked that “[t]he very concept of a knowledge industry contains enough dynamite to blast traditional economics into orbit.” Kenneth E. Boulding, *The Knowledge Industry, Challenge*, May 1963, at 36, 38 (reviewing Fritz Machlup, *The Production and Distribution of Knowledge in the United States* (1962)). Fritz Machlup was one of the first to systematically analyze the production and distribution of commoditized knowledge. See Fritz Machlup, *The Production and Distribution of Knowledge in the United States* 13–43 (1962) (reviewing knowledge classifications and proposing his own). In 1994, Robert E. Babe remarked on the curious need for mainstream economics to treat information as a commodity to make sense of it within the economic paradigm, a need that “obscures many essential properties of information, as well as consequences of informational exchange.” Robert E. Babe, *The Place of Information in Economics*, in *Information and Communication in Economics* 41, 41–42 (Robert E. Babe ed., 1994).

28. See generally S.M. Amadae, *Rationalizing Capitalist Democracy* (2003) (describing how informational tools “led to a far-reaching and comprehensive system for defining appropriate beliefs and actions”); Manuel Castells, *The Rise of the Network Society* (2d ed. 2010) (describing the historical transformations leading to the creation of a network society); Dan Schiller, *How to Think About Information* (2007) (“Companies engaged in making and selling entertainment, banking, communications, data processing, engineering, advertising, law, and other information-intensive services have played an increasingly critical role in overall U.S. investment, employment, and international trade.”). Ian Parker lays out the broad contours of this emergence: “In the 1870s, in North America,

has established the centrality of social data as a vital, even paradigmatic, factor of production under informational capitalism.²⁹ Others have identified the importance of behavioral monitoring and prediction to the governance capacities and challenges of the digital economy.³⁰ Legal scholars have also explored the legal facilitations and fallouts of the informational turn.³¹

about 70 percent of Gross Domestic Product (GDP) was based on material commodity production, with commoditized services constituting about 30 percent of GDP.” See Ian Parker, *Commodities as Sign-Systems*, in *Information and Communication in Economics*, supra note 27, at 69, 74. In the early 1990s, when Parker was writing about them, these percentages had reversed. See id. “The rise of the service sector” (and the decline of material commodity production or manufacturing) has been one of “the most fundamental . . . shifts” to happen across advanced capitalist nations in the latter half of the twentieth century. See id. This shift to the service sector encompasses three broad categories: increases in government expenditure associated with capitalist states’ fiscal policies, the commoditization of care work and other services previously associated with social reproduction (in part a result of the integration of women into the wage-labor market), and the rapid growth of the informational economy. See id. Since around the 1970s, however, the average share of government expenditures as a portion of GDP has plateaued. This means that the continued relative growth of the sector is driven by the other two trends: the ongoing transformation of care and of information into commodities. See id.

29. See Cohen, *Between Truth and Power*, supra note 4, at 42 (“[P]latforms represent both horizontal and vertical strategies for extracting the surplus value of user data. Because that project requires large numbers of users generating large amounts of data, the platform[’s] . . . goal is to become and remain the indispensable point of intermediation . . . in its target markets.”); Nick Couldry & Ulises A. Mejias, *Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject*, 20 *Television & New Media* 336, 337 (2018) (“Just as historical colonialism over the long-run provided the essential preconditions for the emergence of industrial capitalism, . . . we can expect that data colonialism will provide the preconditions for a new stage of capitalism . . . for which the appropriation of human life through data will be central.”); Jathan Sadowski, *When Data Is Capital: Datafication, Accumulation, and Extraction*, *Big Data & Soc’y*, Jan.–June 2019, at 1, 1, <https://journals.sagepub.com/doi/epdf/10.1177/2053951718820549> [<https://perma.cc/RBB3-8AAJ>] [hereinafter Sadowski, *When Data Is Capital*] (“Industries focused on technology, infrastructure, finance, manufacturing, insurance, and energy are now treating data as a form of capital.”).

30. See Katharina Pistor, *The Code of Capital* 213 (2019) (“By constantly contesting the existing boundaries of legal rules in general, and by expanding the remit of the code’s modules to make them fit for ever newer asset classes, lawyers turn any of their clients’ assets into capital.”). See generally Shoshana Zuboff, *The Age of Surveillance Capitalism* (2019) [hereinafter Zuboff, *Surveillance Capitalism*] (“We celebrate the networked world for the many ways in which it enriches our capabilities and prospects, but it has birthed whole new territories of anxiety, danger, and violence as the sense of a predictable future slips away.”).

31. See Cohen, *Between Truth and Power*, supra note 4, at 1 (“Networked information technologies inevitably will alter, and are already altering, the future of law . . .”). See generally Kiel Brennan-Marquez & Daniel Susser, *Knight First Amend. Inst., Privacy, Autonomy, and the Dissolution of Markets*, (Aug. 11, 2022), <https://s3.amazonaws.com/kfai-documents/documents/854ed7a7b7/Brennan-Marquez-Sussner-Privacy-Autonomy-and-the-Dissolution-of-Markets-08.11.22.pdf> [<https://perma.cc/SM73-SKQR>] (describing how, with the rise of information companies, “the capacity of individuals to self-determine—and the capacity of polities to self-govern—is under threat”); Amy Kapczynski, *The Law of Informational Capitalism*, 129 *Yale L.J.* 1460 (2020) [hereinafter Kapczynski, *Informational*].

This Article builds on that earlier work with two goals in mind. First, the Article's primary goal is to provide a granular and reasonably systematic accounting of the various ways data is used (or can be used) by platforms and other firms to produce value (and power). It takes up this goal in Part II. The Article's second goal is to provide a theoretical account of social data as a value form whose cultivation is a primary aim of digital firms—indeed, it is part of what marks the digital economy as “digital.” This theoretical contribution, laid out in Part I, is in service of the primary goal: to illuminate the distinctive value proposition of data and to help explain the conceptual and normative significance of social data as a value form. Taken together, Parts I and II describe the current structure of social data production and suggest why legal scholars and regulators have had trouble grasping the implications and effects of data production under the particular conditions of the contemporary digital economy. Part III explores these legal implications directly. In addition to its two substantive goals, the Article makes a modest methodological contribution to how legal scholarship engages with law's constitution and regulation of production. Its theoretical account supplies a way of analyzing and evaluating productive activity that does not, at the conceptual level, presuppose market ordering of that activity. In doing so, the Article provides a model for similar analysis, when appropriate, for other kinds of productive activity.

I. PREDICTION VALUE AND THE DATA POLITICAL ECONOMY

This Part lays out an account of data as a material store of prediction value. The aim is to sketch out what this Article means when it uses the word “value” and why it's useful to think about value in relation to social data production. To do so, this Part explores three questions. First, *what* is prediction value? What makes it “valuable” and what makes it different from other kinds of value? Second, *why* do companies and governments cultivate and accumulate it? In other words, what is it good for, and how does it fit into current market behaviors and competitive practices? Third, if social data is so valuable, then where has it been all this time? Why are people only talking about it *now*?

A. *What Is “Value”? A Quick Background*

Before the Article turns to social data and prediction value, it is perhaps worth saying a few words about the concept of value more generally. Nowadays, talking about the “value” of something refers to that

Capitalism] (reviewing Cohen, *Between Truth and Power*, supra note 4, and Zuboff, *Surveillance Capitalism*, supra note 30) (discussing the inequitable impacts of informational power); Omri Marian, *Taxing Data*, 47 B.Y.U. L. Rev. 511 (2022) (arguing that in data-rich markets, data should form the tax base); Katharina Pistor, *Rule by Data: The End of Markets?*, 83 *Law & Contemp. Probs.*, no. 2, 2020, at 101 [hereinafter Pistor, *Rule by Data*] (analyzing the possibility of using data to organize markets).

thing's *exchange value*: the priced, monetary value at which it can be (or could theoretically be) bought, sold, or exchanged for something else on a market.³² This is the classic economic sense of "value": the "market value" of a house, or a painting, or a corporate merger, or a bushel of corn. Something's value in a market is a subjective, relative, and contingent measure of that item's desirability. It captures a specific or "average" buyer's "willingness to pay" (WTP) for an additional (that is, marginal) unit of said item. WTP in turn is determined by a combination of external, changing conditions (supply and demand) and internal, stable ones (the buyer's particular reasons for wanting the item). Expressed as a priced market value, WTP also captures relative desirability between goods, since buyers (or most buyers, anyway) have finite wealth and must prioritize among their desires.

This is taken to be quite distinct from "values" in an ethical or sociological sense: beliefs regarding the importance of certain things or actions that motivate individual or societal behavior. It is in this sense that one can speak of the "core values" of an institution or that someone places a "high value on honesty." Indeed, much of what makes exchange value such a useful concept in economics (and beyond) is its stated neutrality on matters of ethical or sociological value: Value is simply the price someone is willing to pay.³³ One need not inquire into why people want what they do and whether those reasons are good or bad (for example, as Adam Smith endeavored to show, "the market rewards us justly for our labours"³⁴). Such questions, to the extent they are answerable at all, lie outside the purview of economic theory.

But economics began as an exploration of these exact questions. How do the individuals in a particular society put their limited time and resources toward productive activity? What does that say about what they value? What is the origin and nature of such value? Reflecting on these questions, observers distinguished exchange value from what they called

32. In classical political economy, exchange value (or *Tauschwert*) refers to only one attribute of a commodity: the proportion at which one commodity can be traded on the market for other commodities. See Jörg Guido Hülsmann, *The Value of Money*, Mises Inst. (Mar. 13, 2013), <https://mises.org/library/value-money-0> [<https://perma.cc/62ZC-TNLL>] (referring to the Mengerian concept of "inner exchange value" and "outer exchange value" as *innerer Tauschwert* and *äusserer Tauschwert*). In this understanding of the term, exchange value isn't necessarily money price, although a market price will generally bear at least a rough correspondence to a commodity's exchange value. See Karl Marx, *1 Capital* 138–39 (Ben Fowkes trans., Penguin Books 1990) (1867). But the general adoption of marginalism in economic thought around the 1930s eliminated such distinctions: Discussions of different value-attributes fell off, and value was taken to measure how much a hypothetical buyer would desire an additional unit of said item, expressed in the form of a price. See Elder-Vass, *supra* note 15, at 18; Herbert Hovenkamp, Coase, Institutionalism, and the Origins of Law and Economics, 86 *Ind. L.J.* 499, 503 (2011).

33. David Graeber, *Value: Anthropological Theories of Value*, in *A Handbook of Economic Anthropology* 439, 443 (James G. Carrier ed., 2005).

34. *Id.*

“use” value or “natural” value: the value one gets from, say, *wearing* one’s favorite coat (that is, using it as a coat), as opposed to the value that one or another would pay to *acquire* the coat.³⁵ These were not mere differences in the degree to which people valued things, which continues to be widely observed today. For example, any economist will point out that the decision *not* to sell something—a home, a corporate asset, or one’s favorite coat—is simply a signal that one values that thing over its (current) market value. Behavioral economists have added to this observation the concept of “endowment effects”—that people systematically tend to overestimate the market value of something that they already have when compared with what they would pay to obtain it (a testament to the notion that people grow attached to things they think of as their own).³⁶ Both of these accounts, however, ground their explanations of such behavior in numerical difference in value—quantitative assessments of the priced value of a particular good.³⁷

In contrast, early economists used the concept of use value to express a distinct *way* that a good could be valued, not simply differing degrees to which a good could be valued along the same dimension. So for example, when one enjoys the use of one’s favorite coat, how one enjoys it is distinct from its price. One feels fondness for it from its familiarity, aesthetic satisfaction from its cut and drape on one’s body, and the pleasure of being warm on a cool day. This all *can* be, in an abstract sense, translated into a price if someone were to offer to buy the coat. But it would be a translation:

35. The concepts of “use value” and “exchange value” are quite old. Marx quotes Aristotle on the subject:

For twofold is the use of every object The one is peculiar to the object as such, the other is not, as a sandal which may be worn is also exchangeable. Both are uses of the sandal, for even he who exchanges the sandal for the money or food he is in need of, makes use of the sandal as a sandal. But not in its natural way. For it has not been made for the sake of being exchanged.

Marx, *supra* note 32, at 179 n.3 (alteration in original) (internal quotation marks omitted) (quoting Aristotle, Republic bk. I, ch. 9) (misquotation). This quotation is properly attributed to Aristotle’s *Politics*. See, e.g., Aristotle, *Politics* bk. I, at 41 (H.W.C. Davis, ed., Benjamin Jowett, trans., Oxford Clarendon Press 1908) (c. 350 B.C.E.). As David Graeber details, Adam Smith’s famous “paradox of value” is also much older than the eighteenth century: St. Augustine argued that “according [to] their own merits,” “plants are clearly superior to stones, animals to plants, humans to animals,” but because of humans’ fallen nature, and thus “endless physical needs and desires,” people value things like bread and gold over animals like mice. Graeber, *supra* note 33, at 441–42 (quoting St. Augustine, *The City of God*, bk. IX, ch. 16). To St. Augustine, this characterizes how people “come to see things through [their] own needs (use value) rather than their absolute worth” or “their position along the Great Chain of Being.” *Id.*

36. See Daniel Kahneman, Jack L. Knetsch & Richard H. Thaler, Experimental Tests of the Endowment Effect and the Coase Theorem, 98 *J. Pol. Econ.* 1325, 1326–27 (1990) (overviewing empirical research displaying endowment effects).

37. Put another way, these are both instances when people are making an assessment of the current market price of such goods—deeming that price lower than what they consider the goods’ exchange value.

It would leave out dimensions of what one derives from the use of one's favorite coat. Use value and exchange value thus captured distinct aspects of a good. The two concepts indexed different ways that people relate to and derive value from things, and different motivations they might have for making or obtaining things. The disaggregated notions of value thus corresponded to, or described, different aspects of "productive" or economic activity. The two were related, of course, and much of early economic thought was devoted to developing and debating various accounts for the origins of value and the formal causal or mathematical relation between use value and exchange value.³⁸

There is widespread agreement that classical political economists' preoccupation with developing some "true," scientific, or systematic relationship between the different notions of value was destined for failure. The history of economic thought is marked by nonexchange conceptions of value coming into conflict with and ultimately being (albeit imperfectly and partially) transfigured into exchange value.³⁹ Over time, "the value of an object became increasingly indistinguishable from its price: how much potential buyers were willing to give up to acquire some product on the market."⁴⁰ And to be clear, this Article does not aim to revive these old debates, nor to argue for or defend a formal and systematic relationship between the two concepts of value.

Yet, in abandoning the study of distinct concepts of value, economics also lost its ability to fully express—and thus take seriously on their own terms—how the uses of things can motivate and explain economic activity. This was particularly true for things that were of high social or personal utility but that nonetheless had low exchange value. Or more accurately, economic thought moved on from puzzling deeply over such enduring "paradoxes" of commercial activity to providing a ready answer for resolving them. The standard line is that such things suffer from having high total utility but low marginal utility. Thus, "undervaluation" problems have, in theory, a simple solution. To properly value such things requires

38. Early economic thought was divided into three schools: mercantilists, physiocrats, and political economists, all of whom took a different position on how value was created. Graeber, *supra* note 33, at 440. Mercantilists believed wealth originated from precious metals (gold, silver); physiocrats believed it originated from nature, and hence, agriculture. The classical political economists believed value was a product of human labor, that it "emerged . . . at exactly the point where our minds became a physical force in nature." *Id.*

39. And in turn, as economic thought permeated social thought more generally, the process of subordinating other ways of thinking about value to exchange value spread beyond economic theory. For one treatment of the normative and ethical shortcomings of this trend, see generally Elizabeth Anderson, *Value in Ethics and Economics* (1993). For a sociological account of how thinking of all political and social goals in terms of exchange value came to dominate policymaking, see generally Elizabeth Popp Berman, *Thinking Like an Economist: How Efficiency Replaced Equality in U.S. Public Policy* (2022). On the conceptual and descriptive shortcomings of contemporary theories of value, see Elder-Vass, *supra* note 15, at 18–24.

40. Graeber, *supra* note 33, at 440.

that people take the steps necessary to express their value *as* exchange value—in other words, to *create* a market (or if that isn't possible, to simulate a market) in such goods.⁴¹

* * *

When thinking about value these days, it can be hard to escape the exchange value-sea everyone swims in. Saying something has “economic value” is necessarily taken to mean that this value *also* takes a particular form: exchange value. But this Part endeavors to show that it can be worth returning to the old tradition of disambiguating forms of value—not to revive the old views of value entirely⁴² but to take seriously in the exploration of economic activity the distinctions between different ways of cultivating and deriving productive value, and the imperfect and messy task of transformation that occurs between forms of value.

B. *What Is Prediction Value?*

Social data production produces value from its capacity to materialize and store traces of human activity. This allows entities to catalog, analyze,

41. Here a reader could ask, “Aren't you just talking about commodification?” And the answer is, in some sense, yes. But commodification better describes the social process that *follows* from, or accompanies, the intellectual shift being described here. This shift is conceptual: how and why economics came to think of essentially all social value, and human behavior, in terms of prices and price theory to begin with. Indeed, much criticism lodged against neoclassical economics is due to its tendency to engage in “economic imperialism,” casting any behavior involving scarce resources within the conceptual precepts of neoclassical price theory. For a critical account of this trend, see generally Berman, *supra* note 39; Don Herzog, *Externalities and Other Parasites*, 67 U. Chi. L. Rev. 895 (2000) (book review). For example, family planning, racial discrimination, crime, marriage, divorce, drug addiction, politics, immigration, and suicide have all been cast as problems that can be thought of—and solved—via pricing and markets. See, e.g., Richard A. Posner, *A Theory of Primitive Society, With Special Reference to Law*, in *Chicago Studies in Political Economy* 149, 191–205 (George J. Stigler ed., 1988) (analyzing elements of family, tort, and criminal law as systems for selling and buying legal entitlements). On specific policy suggestions, see, e.g., Eric A. Posner & E. Glen Weyl, *Radical Markets: Uprooting Capitalism and Democracy for a Just Society* 127–67 (2018) (proposing a market in immigration in which private citizens could bid to host immigrants in exchange for a share of the income they earn); Richard A. Posner, *Sex and Reason* 241–429 (1992) (developing an economic theory of sexuality and applying it to the AIDS epidemic, abortion, gay rights, surrogacy, motherhood, marital rape, date rape, sexual harassment, sexual abuse, and pornography); Elisabeth M. Landes & Richard A. Posner, *The Economics of the Baby Shortage*, 7 J. Legal Stud. 323, 323–48 (1978) (proposing the introduction of a market for adoption in which people may buy and sell babies). Economists Gary Becker and Julio Jorge Elías have argued for the application of markets to organ donation. See generally Gary Becker & Julio Jorge Elías, *Introducing Incentives in the Market for Live and Cadaveric Organ Donations*, J. Econ. Persps., Summer 2007, at 3. Becker has also written on topics like birth rates, family size, marriage, divorce, family dynamics, and child raising. See generally Gary Becker, *A Treatise on the Family* (1981) (applying economic theory and decisionmaking to the creation and maintenance of family units).

42. Again, this Article has no interest in wading into centuries-old fights about value theory.

aggregate, and mine such traces for insight and, in turn, to use such insights to apprehend, predict, and modify behavior. In other words, the value of social data lies in its capacity to apprehend and predict—and based on such apprehension and prediction, to manage—social behavior.⁴³ Both apprehension and prediction *and* the behavioral management capacities they endow have value. Apprehension and prediction can aid in planning how companies allocate resources, streamline logistics, manage risks, and enter new sectors. Companies can use data systems to affect human behavior (either directly or indirectly). For instance, they can aid workforce efficiency and other goals and help manage consumer interactions. These strategies, which companies can pursue independently or at the same time, are covered in detail in Part II below.

1. *Defining Social Data.* — Definitions of “data” and its use have a fraught and thorny history.⁴⁴ Given considerable debate regarding the accuracy or usefulness of terms like “data” or “privacy” or “personal data,” it is worth saying a bit more about this Article’s use of the term “social data,” as well as its limits. At its simplest, this Article uses the term “social data” to mean data about people. This can include data collected directly from people—such as their movement around a city on foot or in a car, their breathing patterns and heart rate, or their cursor or eye movements across a screen.⁴⁵ It also includes data applied to make inferences about

43. The value proposition here is that better prediction informs better strategies for action. The more an entity knows about a behavior or attribute and the effect of a proposed intervention on that behavior (or attribute), the more effectively it can convert prediction into a desired outcome (e.g., watching more Netflix, buying more products of the right kind on Amazon, driving longer hours on Lyft, being charged the highest price one is willing to pay in an ad exchange). Increased efficacy can refer to a variety of comparative advantages in modulating behavior toward a desired goal. Greater prediction value can produce interventions that are more accurate, more subtle, more widely deployed, more quickly deployed, etc. The particular usage of prediction value will depend on the setting in which it is being used and the desired goal.

44. See, e.g., Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. Rev. 1814, 1828, 1832, 1841–45 (2011) (detailing competing definitions of personally identifiable information (PII), including the limitation that PII is understood as solely information about one specific person).

45. See Ifeoma Ajunwa, *The Quantified Worker* 139 (2023) (describing a hiring company that uses AI to analyze vocal and facial expressions such as “brow furrowing, brow raising, [and] the amount eyes widen” (internal quotation marks omitted) (quoting Ivan Manokha, *How Using Facial Analysis in Job Interviews Could Reinforce Inequality*, PBS News Hour: Making Sen\$e (Oct. 7, 2019), <https://www.pbs.org/newshour/economy/making-sense/how-using-facial-recognition-in-job-interviews-could-reinforce-inequality> (on file with the *Columbia Law Review*))). See generally Annette Bernhardt, Lisa Kresge & Reem Suleiman, U.C. Berkeley Labor Ctr., *Data and Algorithms at Work, The Case for Worker Technology Rights* (2021), <https://laborcenter.berkeley.edu/wp-content/uploads/2021/11/Data-and-Algorithms-at-Work.pdf> [<https://perma.cc/VX7S-QULV>] (describing how employers use social data like keystroke tracking and real-time location monitoring to surveil workers).

people—such as changes in light patterns from pixels of a smart television screen or local weather data. This Article uses “social data” because it better expresses the view that data is (almost) always produced for what it can reveal about people, not individual persons.⁴⁶ It thus readily expresses the notion that data is used and is useful in the aggregate and is valued for its repurposable, aggregative, and relational properties.⁴⁷ “Social data” also avoids some semantic traps of the term “personal data,” which can be defined expansively but is commonly defined far more narrowly in the many privacy laws that use it. Thus, “social data” is both more accurate and invokes the correct common intuitions about the concept.

One limitation of the term “social data” is that it may indeed be too expansive, covering many forms of data that are not obviously about people.⁴⁸ Given the pervasive issue of “personal data” definitions failing to capture relevant forms of data use, proceeding from a more expansive view may be a welcome corrective. Moreover, if “social data” covers many forms of data that are not obviously about people (until that data is used to infer or act on human behavior in a particular application), this is not a *prima facie* disqualifier for the term. Rather, it is a relevant feature of social data that in turn implies conceptual and normative criteria for its appropriate regulation.

Another limitation may cut the other way: The concept leaves out forms of data value that do not derive from understanding, predicting, or intervening on human behavior. Other data may be valuable for its capacity to predict things that have nothing whatsoever to do with human action. This Article takes no issue with this claim. Although such data may indeed produce value, it is at least plausible that such data production produces less social disruption and transformation, and thus its cultivation poses less salient or novel legal issues. Given that digital technology companies themselves focus on human-derived or human-applied data, and that critical, scholarly, and policy-oriented communities similarly focus

46. See Viljoen, *Relational Theory*, *supra* note 6, at 609–12; see also Natasha Singer, *Just Don't Call It Privacy*, *N.Y. Times* (Sept. 22, 2018), <https://www.nytimes.com/2018/09/22/sunday-review/privacy-hearing-amazon-google.html> (on file with the *Columbia Law Review*) (arguing that congressional hearings on large corporations' privacy policies should focus on how the data is used, not individual privacy concerns).

47. See Viljoen, *Relational Theory*, *supra* note 6, at 609–12.

48. One strong form of this claim is that it is unclear if any kind of modifier on “data” is needed at all. If we define “social data” as “any physical world observation rendered in datafied form that goes on to be used to derive insight regarding human action and behavior,” then it becomes hard to imagine what data being produced *doesn't* fall within this definition. Or, if such data exists, that it is (economically, normatively, legally) trivial. However, showing the (plausible and interesting) proposition that the subset “social data” also contains the set “data” is not the project of this Article. For now, this Article assumes there is some subset of “data” composed of “social data,” as defined above, and that it is this subset that is the Article's subject of inquiry. The authors thank Thomas Streinz for first drawing our attention to this intriguing point.

on the effects of datafication on human action and social life, this Article constrains its inquiry to human-related data.

It is also important to note that although the term is useful as an analytic category, if directly carried over into a legislative context, “social data” could easily suffer from some of the same issues that plague “personal data.” This is because whether one focuses on data that is “personal” or “social,” data about people is not a quality or category that is inherent to data. It is instead a determination that must be made *about* data, one that requires more information about the context, purpose, and processes that guide its collection and use.⁴⁹ In this way, determining that data is “social” is akin to legal rules whose application require fact-specific inquiries. So, whether social data qualifies as such (let alone whether it is collected and used in appropriate ways) cannot be a predetermined definitional exercise. One cannot observe different categories of information—such as TV streaming data, biometric data, locational data, household energy usage data, city water flow data, or weather data—and a priori determine which types of data, as a category, are personal or social, and which are not (although some may have greater propensity to be used as personal or social data based on what value-seeking strategies motivate their collection and use). Thus, “social data” is not an assessment of the type of data being collected alone. Social data also depends on the contexts in which it is collected and used. The reasons and the purposes of the data’s collection structure how it is cultivated and stored and are informed by and constrain its use. Indeed, social data may be better understood as a category of action than as a category of object—an act that endows the thing being used (materialized stores of information) with the relevant set of properties (social insight into human behavior).

So, “social data” is a definitional shorthand for the concept that human-relevant data value (and risk) is not about persons, but about people. But there are other terms out there.⁵⁰ Ultimately, this Article is more interested in the underlying concept than any one semantic signifier

49. This feature of “social data” is heavily influenced by Helen Nissenbaum’s theory of privacy as “contextual integrity.” See generally Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (2010). But while contextual integrity calls upon social context to engage in normative inquiry, we are merely interested in descriptive inquiry regarding data about people—namely, how entities cultivate it and use it to produce value. This inquiry prefigures any normative evaluation of ways data has been put toward the project of apprehending people and intervening in their actions and whether such data sharing is wrongful. “Social data” in our usage is thus a descriptive category of information, not a normative account of what makes information flows appropriate.

50. See Cohen, *Between Truth and Power*, *supra* note 4, at 48–50 (describing the data flows about people as the “biopolitical public domain”); Neil Richards, *Why Privacy Matters* 24–25 (2021) (favoring the term “human information” because it “refers us back to the human beings whose information is being used in some of the technologies that are the hallmark of the digital age”).

for it. If the reader prefers another term for this underlying concept, then it can be substituted in.

2. *Defining Prediction Value.* — This Article uses the term *prediction value* for the specific form of use value entities derive from the systematic cultivation of social data. It indexes the value of being able to infer or more accurately predict future actions or effects. The prediction value of social data lies in the capacity of data to materialize apprehension and convert it into control: It confers upon its holder the capacity to exert a desired effect on future behavior. Capacity is not the same thing as using that capacity. Similar to power in the physical sense, prediction value, stored via social data, is a *potential* to convert into an effect (the exploitation of data's prediction value via use).

Social data is thus the material store (or medium) of prediction value. Its production materializes the latent predictive value of human activity for other human action so that it can be stored, used, reused, aggregated, and recombined with other data (other media of prediction value) across time and context, and mined for different kinds of insights and interventions.⁵¹ Merging data with other data is a way to combine and compound prediction value. Analyzing data is a way to tap into and exploit prediction value. Artificial intelligence (AI) and machine-learning models trained on data demonstrate one significant way that companies distill prediction value into a defined (and generalizable) product.

Prediction value is derived from data's relational character. Because people are social beings who are like one another, who are the products of social formation, and who construct their self-identity in dialogue with others, datafied traces of observed actions and behaviors of the many have something meaningfully predictive to say about anyone.⁵² Data stores represent a stem-cell-like proto-asset that can be reapplied and specialized to predict behavior across a variety of settings.⁵³

The general observation that information produces social control, and that this in turn can be exploited for commercial gain, is not novel. What this Article calls "prediction value" builds on related concepts in prior works. Shapiro identifies how platforms engage in worker surveillance to cultivate "calculative rationalities," which platforms exploit to manage worker behavior and maximize their own gain.⁵⁴ Kean Birch,

51. Cultivating data is less of a "harvesting" process than it is a "manufacturing" process. The "complete" picture of human activity is not replicated like some digital twin or mirror. The process is more one of an engineered, synthetic distillation of human activity that reflects the goals of data production. See, e.g., Cohen, *Between Truth and Power*, supra note 4, at 56 (describing data mining as generating new techniques and capabilities that distill troves of social data into "Big Data").

52. See Viljoen, *Relational Theory*, supra note 6, at 609–10.

53. See id. at 589–90 & n.25.

54. Aaron Shapiro, *Between Autonomy and Control: Strategies of Arbitrage in the "On-Demand" Economy*, 20 *New Media & Soc'y* 2954, 2968 (2018). Shapiro develops the notion of asymmetries in his work on "platform sabotage," a term he uses to describe

DT Cochrane, and Callum Ward detail how platforms express value from personal data indirectly to investors and other market actors through “user metrics.”⁵⁵ Birch and others have written extensively on “assetization” and informational value,⁵⁶ and sociologists Thomas Beauvisage and Kevin Mellet extend this assetization account to personal data.⁵⁷ Economist Cecilia Rikap details the “intellectual monopolization” of power among the world’s largest companies through the systematic concentration of knowledge and “planning capacity” that, she argues, extends these companies’ power over the economy beyond their legally-owned access.⁵⁸ Niels Van Doorn and Adam Badger discuss the “speculative value” of gig platforms that configure data as a financial asset based on the expectation that data-driven analytics will later realize as efficiency gains.⁵⁹

Mainstream economic theory has taken a renewed interest in prediction value, though economists are decidedly mixed on whether prediction value is a good thing. Jean Tirole details how “managing the flow of information about individuals’ behavior” allows entities to “achieve social control,” which he argues can both promote prosocial behavior and

platforms’ use of data and computation to derive value through strategically inserted inefficiencies in the market encounters they facilitate. Shapiro, *Platform Sabotage*, *supra* note 21, at 204.

55. See Kean Birch, DT Cochrane & Callum Ward, *Data as Asset? The Measurement, Governance, and Valuation of Digital Personal Data by Big Tech*, *Big Data & Soc’y*, Jan.–June 2021, at 1, 11, <https://journals.sagepub.com/doi/epdf/10.1177/205395172111017308> [<https://perma.cc/FKA8-ZWJW>] [hereinafter Birch et al., *Data as Asset?*].

56. See, e.g., Ctr. for Int’l Governance Innovation, *Rethinking Canada’s Competition Policy in the Digital Economy* (2023), https://www.cigionline.org/static/documents/2023_Rethinking_Canadas_Competition_Policy.pdf [<https://perma.cc/6UZU-R2NV>]; Kean Birch & Callum Ward, *Assetization and the ‘New Asset Geographies’*, 14 *Dialogues in Hum. Geography* 9, 9–23 (2024); Kean Birch & D.T. Cochrane, *Big Tech: Four Emerging Forms of Digital Rentiership*, 31 *Sci. as Culture* 44, 48 (2022); Kean Birch, *There Are No Markets Anymore: From Neoliberalism to Big Tech*, *Transnat’l Inst.* (Feb. 3, 2023), <https://www.tni.org/en/article/there-are-no-markets-anymore> [<https://perma.cc/W7C3-2QWF>]. Birch and others distinguish the process of assetization from commodification: Although they can be exchanged, assets are not produced for their exchange value (i.e., to be bought and sold) but instead to secure durable economic rents (investment and return) through the ownership and control of an asset. Kean Birch & Fabian Muniesa, *Introduction: Assetization and Technoscientific Capitalism*, *in* *Assetization: Turning Things Into Assets in Technoscientific Capitalism* 1, 2–3 (Kean Birch & Fabian Muniesa eds., 2020).

57. Thomas Beauvisage & Kevin Mellet, *Datasets: Assetizing and Marketizing Personal Data*, *in* *Assetization*, *supra* note 56, at 75, 77 (“We argue that the ability to capitalize personal data in the present is the result of a versatile and uncertain process of assetization.”).

58. Cecilia Rikap, *From Global Value Chains to Corporate Production and Innovation Systems: Exploring the Rise of Intellectual Monopoly Capitalism*, 7 *Area Dev. & Pol’y* 147, 155 (2022) [hereinafter Rikap, *Intellectual Monopoly Capitalism*]; see also Cecilia Rikap, *Capitalism as Usual?*, 139 *New Left Rev.* 145, 159 (2023) [hereinafter Rikap, *Capitalism as Usual?*].

59. Niels van Doorn & Adam Badger, *Platform Capitalism’s Hidden Abode: Producing Data Assets in the Gig Economy*, 52 *Antipode* 1475, 1476–77 (2020). Part II below discusses the transformation of prediction value into exchange value in greater detail.

“destroy the social fabric.”⁶⁰ Glen Weyl and others argue that “free” online services lead us to systematically under-incentivize and maldistribute data value; to correct these issues, they argue for treating the market for data like a labor market.⁶¹ Importantly, Alessandro Bonatti explores how data’s relational character influences data acquisition and monetization strategies among platforms, including how they strategically use data not only to improve matches but to bolster market power.⁶²

Indeed, astute observers have been developing accounts of social data’s capacity to cultivate power and value for some time.⁶³ In her 1988 ethnography of computerizing workforces, Shoshanna Zuboff describes how these workplaces were deriving value not from automating but from “informating”—datafying worker actions rather than simply automating workers away.⁶⁴ Oscar Gandy’s seminal 1993 work *The Panoptic Sort* details a “system of power” developed from social data and used to “coordinate and control [people’s] access to the goods and services that define life in the modern capitalist economy.”⁶⁵ Gandy detailed how tracking institutions also enacted a distinct view of social life: They are concerned not with cataloging the self-conceptions of groups or individuals but instead with identification (categorization for institutional utility),

60. Jean Tirole, *Digital Dystopia*, 111 *Am. Econ. Rev.* 2007, 2007 (2021) (emphasis added).

61. See Posner & Weyl, *supra* note 41, at 234–39, 246–49; Imanol Arrieta-Ibarra, Leonard Goff, Diego Jiménez-Hernández, Jaron Lanier & E. Glen Weyl, *Should We Treat Data as Labor? Moving Beyond “Free”*, 108 *Papers & Proc. Am. Econ. Ass’n* 38, 40–41 (2018) (“[Large tech companies] benefit from the free or extremely cheap availability of data. . . . [U]sers aware of the value of their data would likely demand compensation in a range of settings, dramatically reducing the share of value that could be captured. . . .”).

62. See Alessandro Bonatti, *The Platform Dimension of Digital Privacy* 22 (May 23, 2023), <https://www.nber.org/system/files/chapters/c14782/c14782.pdf> [<https://perma.cc/W77G-KWS9>] (unpublished manuscript).

63. For several older works, see *supra* notes 27–28. James Beniger’s seminal work *The Control Revolution* came out in 1986. James R. Beniger, *The Control Revolution: Technological and Economic Origins of the Information Society* 89–91 (1986) (describing the evolution of social control and social production of meaning).

64. The authors thank Dan Greene, Nathan Beard, Tamara Clegg, and Erianne Weight for pointing to this example. Daniel Greene, Nathan Beard, Tamara Clegg & Erianne Weight, *The Visible Body and the Invisible Organization: Information Asymmetry and College Athletics Data*, *Big Data & Soc’y*, Jan.–June 2023, at 1, 3, <https://journals-sagepub-com.ezproxy.cul.columbia.edu/doi/epdf/10.1177/20539517231179197> [<https://perma.cc/HFK7-ZXPZ>].

65. Oscar H. Gandy, Jr., *The Panoptic Sort: A Political Economy of Personal Information* 29 (2d ed. 2001) [hereinafter Gandy, *Panoptic Sort*]; see also Oscar H. Gandy, Jr., *Coming to Terms With the Panoptic Sort*, in *Computers, Surveillance, and Privacy* 132, 133 (David Lyon & Elia Zureik eds., 1996) (“The panoptic sort is a complex discriminatory technology. It is panoptic in that it considers *all* information about individual status and behavior to be potentially useful in the production of intelligence about a person’s economic value.”).

classification (“the panoptic sort is a difference machine”), and assessment.⁶⁶

3. *Prediction Value Versus Exchange Value.* — Social data production materializes and stores value (and risk) in ways that are distinct from and not always neatly transformed into exchange value.⁶⁷ Robert Babe provided an impressively prescient distillation of the problem: Money price states how much exchange value an informational commodity “contains” but is silent regarding what he calls its “quantity of information” (a concept loosely akin to what this Article calls its prediction value).⁶⁸ Data, in other words, is different.⁶⁹

As Babe notes, for neoclassical economic theory (although not, importantly, older traditions) to make sense of informational value, it needs to be reduced to exchange value to be indexed and to “count” as value in economic terms. But social data “does not fulfil[1] the definitional or conceptual requirements of commodit[ies]” as understood by economists of the time.⁷⁰ Moreover, attempts to translate or reduce information’s prediction value into such terms “obscures many essential properties of information, as well as consequences of informational exchange.”⁷¹

These distinct properties also suggest diverse applications in which, as a general matter, one kind of value enjoys a relative advantage over the other. Prediction value would be well spent, for example, in allocating goods for which priced market allocation might be independently wrongful.⁷² For example, one may think it wrong to affix a price to hearts for transplant, but this does not mean one would not value informational mechanisms to accurately identify which operating theaters need hearts and which may supply them. The same might be true in the reverse. For example, one may think the singular quality of artistic creativity and idiosyncrasy in aesthetic taste are intrinsically valuable, so one ought to resist allocating productive resources to artistic production based on

66. Gandy, *Panoptic Sort*, *supra* note 65, at 29.

67. Babe, *supra* note 27, at 41, 45.

68. *Id.* at 45.

69. The authors thank Thomas Streinz for this succinct distillation of this section’s argument.

70. Babe, *supra* note 27, at 42.

71. *Id.*

72. For a discussion of a variety of policy areas in which critics have argued that the application of pricing and market allocation to guide social policy would be wrongful, see *supra* note 41; see also Rahel Jaeggi, *What (If Anything) Is Wrong With Capitalism?*, 54 *S.J. Phil. (Spindel Supp.)* 44, 55 (2016) (distinguishing quantitative and qualitative accounts of what makes the market allocation and organization of certain social functions and goods normatively wrongful); Salomé Viljoen, *Informationalism Beyond Managerialism* 126–28 (Feb. 15, 2024) (unpublished manuscript) (on file with the authors) (“The rich data inputs and infrastructures that sustain market machines if transposed into settings with different productive logics and more democratically determined goals may offer one way around, past, or beyond managerialism as a prevailing regulatory paradigm.”).

prediction. In other words, one may not want to overfit how one produces future artistic work to predictions based on what people enjoyed in the past, even if such predictions are accurate and regularly updated for changing taste.⁷³ Instead, one may want to preserve the societal capacity to invest in and support artistic work that breaks new ground. Investments in artistic work via exchange value are better equipped—in theory at least—with the capacity to support and reward true novelty and creative inspiration.

In some ways, prediction value is *more* general than priced exchange value.⁷⁴ Prediction value can be converted into exchange value, but it need not convert into wealth directly to exert power or control, or to drive decisions.⁷⁵ Prediction value can exert social discipline on others and enact material outcomes or moral judgment that we do not currently express as market discipline (and perhaps could not so express even if we wanted to).

In other ways, prediction value is decidedly *less* general than exchange value. Prediction value of a given data point or dataset is contingent and unpredictable in ways that are unlike a more typical asset or commodity. Prediction value is not stable across contexts⁷⁶ and—as data ages, or is combined with other data, or new analytic techniques or technical applications are developed—prediction value may shrink or grow. Prediction value is (in some general sense at least) nonrivalrous, but it's also not straightforwardly fungible and thus not readily subject to traditional forms of transfer and redistribution. Marc Porat, an early information economist, said that information was, by nature, a “heterogenous commodity” that cannot be collapsed into one sector—like mining. For economists to make sense of it, he argued, they needed to think of the production, processing, and distribution of information goods and services as an *activity* rather than as a product.⁷⁷ While a growing

73. For a discussion of Netflix using social data to guide production of new content, see *infra* notes 154–155 and accompanying text.

74. See Babe, *supra* note 67, at 42 (“[N]eoclassicists’ notions of ‘market,’ ‘price,’ ‘value,’ ‘commodity,’ ‘demand,’ ‘supply,’ and ‘exchange’ are but specialized instances of broader communicatory phenomena.”). As Kenneth Arrow explained: “The meaning of information is precisely a reduction in uncertainty.” *Id.* at 48 (internal quotation marks omitted) (quoting Kenneth J. Arrow, *The Economics of Information*, in *The Computer Age: A Twenty-Year View* 306, 306 (Michael L. Dertouzos & Joel Moses eds., 1979)).

75. See *infra* section II.A.3 for a discussion of using prediction value to gain or exert power.

76. Naturally, problems arise when social data resources have high predictive value for a given task, but the task is one for which using such data is inappropriate. Much like its analytic value, normative analysis of when predictive value may be legitimately spent is also not stable across contexts. See Nissenbaum, *supra* note 49, at 129 (“[T]here is . . . great complexity and variability in the privacy constraints people expect to hold over the flow of information, but these expectations are systematically related to characteristics of the background social situation.”).

77. Marc Uri Porat, U.S. Dep’t of Com., *The Information Economy: Definition and Measurement* 24 (1977). As Babe and others note, focusing on the inputs of information

number of economists are studying prediction value, and recent work offers empirical support for the gap between traditional accounts of data's value and firm behavior, far more empirical work is needed.⁷⁸

4. *Prediction Value Into Exchange Value?* — As life becomes increasingly digitized and datafied, prediction value takes on a greater economic and conceptual significance. Many of the largest companies in the world have, at least in part, gotten that way by accumulating and exploiting prediction value. This in turn puts general market pressure on *other* companies and entities to do the same thing to retain their market position against competitors.⁷⁹

As some of the scholarship reviewed above suggests, companies face pressure to transform prediction value into exchange value, or at the very least translate prediction value into exchange-value terms for investors and, to a lesser extent, regulators. Yet not all prediction value is readily reducible to exchange value, and there are several reasons, both practical and normative, to resist attempts to do so.

Part II explores in greater detail how such transformation works: what is lost and entrenched in the process, what is the effect on the political economy of prediction value, and what are some legal issues that arise along the way. Prediction value confounds traditional approaches to regulating and apprehending value in law. This is a problem insofar as many areas of law (following the “if there’s no price, there’s no value” view) do not register, apprehend, or count as significant forms of value production that do not readily convert into exchange value.⁸⁰ As Part II surveys, many legally and normatively relevant decisions regarding the

production alone (the activity) still only partially captures total prediction value. See Babe, *supra* note 67, at 41–42.

78. For current research into this topic, see generally Alessandro Acquisti & Hal R. Varian, *Conditioning Prices on Purchase History*, 24 *Mktg. Sci.* 367 (2005); Aileen Nielsen, *Measuring Lay Reactions to Personal Data Markets*, 2021 *Proc. Ass’n for Computing Mach. Conf. on A.I. Ethics & Soc’y* 807; Susan Athey, Christian Catalini & Catherine Tucker, *The Digital Privacy Paradox: Small Money, Small Costs*, Small Talk (Nat’l Bureau Econ. Rsch., Working Paper No. 23488, 2017), https://www.nber.org/system/files/working_papers/w23488/w23488.pdf [<https://perma.cc/6F2M-TFY5>]; Alessandro Acquisti, *The Economics of Privacy at a Crossroads* (July 2023), <https://www.nber.org/system/files/chapters/c14785/c14785.pdf> [<https://perma.cc/R3AR-59AS>] (unpublished manuscript); Bonatti, *supra* note 62; Aileen Nielsen, *Whose Data, Whose Value? Simple Exercises in Data and Modeling Evaluation and Implications for Tech Law and Policy* (June 7, 2023), <https://ssrn.com/abstract=4434409> [<https://perma.cc/SA6D-G9LJ>] (unpublished manuscript); Catherine Tucker, *The Economics of Privacy: An Agenda* (June 10, 2023), <https://www.nber.org/system/files/chapters/c14781/c14781.pdf> [<https://perma.cc/L86Z-NMS9>] (unpublished manuscript).

79. See Sadowski, *When Data Is Capital*, *supra* note 29, at 1.

80. Ronald Coase was a very astute observer of how a great deal of firm activity and production refigured pricing. See Sanjukta Paul, *On Firms*, 90 *U. Chi. L. Rev.* 579, 603–06 (2023) (explaining that under Coase’s account, firms arise to decrease the transaction costs inherent to “the price mechanism” (internal quotation marks omitted) (quoting Coase, *supra* note 15, at 390)).

information economy and its effects on the social world occur well before prediction value converts into exchange value, if it ever converts at all.

C. *Why Datafy? And Why Take Data Value Seriously?*

Focusing on data's prediction value helps home in on two questions. First, *why* are companies using time, money, and energy to collect data to begin with? And second, if managing access to information has always been a source of power, why is it worth paying particular attention to prediction value, in the form of social data, *now*? These questions prefigure many of the disruptions and legal issues that result from commercial surveillance, and that have received sustained popular and scholarly attention.

Entities looking to exert influence have long used the controlled management of information to engage in worldmaking and to understand and shape human behavior.⁸¹ But only recently has it been economically feasible to mine, store, and generate prediction value at scale.⁸² In other words, it takes a particular set of conditions for the accumulation of prediction value to become an imperative of commercial competitive success. These conditions include innovations in the cultivation, storage, transfer, aggregation, and combination of social data,⁸³ and improvements in data science and machine-learning techniques to derive insights from large data flows and to apply them via trained models. Together, these improvements all contribute to the feasibility and utility of exploiting social data for prediction value at scale.⁸⁴

These innovations in turn allow entities to exploit for economic value that which has long been true: People are social beings, deeply knowable and materially influenced by our relations to one another. This capacity

81. See Sarah E. Igo, *The Known Citizen: A History of Privacy in Modern America* 71–83 (2018) (charting long trends of informatization and knowledge access as power, such as the development of the Social Security Board and Administration); Chris Wiggins & Matthew L. Jones, *How Data Happened: A History From the Age of Reason to the Age of Algorithms*, at xi–xiv (2023) (providing a historical account of data as an instrument of worldbuilding and a means of allocating power); see also Graeber & Wengrow, *supra* note 9, at 364–65.

82. Consider a simple analogy to energy. In its natural form, as water, lightning, sun, and wind, energy has always existed and has long been known to be a useful source of power. But it was only over the course of the nineteenth century that people developed the techniques to store and transmit energy (and obtained the necessary economic conditions to make the mass adoption of such techniques feasible). This is when energy's capacity or power to do work in the physical sense (i.e., to generate heat and light) could play a transformative role in the political economy.

83. Cohen provides an in-depth review of the rise of these historical and technological conditions. See Cohen, *Between Truth and Power*, *supra* note 4, at 8–9.

84. See David L. Rogers, *The Digital Transformation Playbook: Rethink Your Business for the Digital Age* 99–100 (2016) [hereinafter Rogers, *The Digital Transformation Playbook*] (describing how advances in chip manufacturing and data processing has made it easier to take better advantage of data analytics).

for social data to be stored, mined, and exploited for prediction value, which in turn leads to hegemonic market pressure to datafy social life for exploitation and accumulation, is what demarcates informational capitalism from its predecessors.

Here a reader might raise one plausible objection: If much of what companies claim as “prediction value” is in fact nothing more than speculation and legal maneuvering, then arguments that take the existence of this value seriously may concede too much to claims that value is, indeed, being created from the datafication of social life. The appropriate response, it follows, to the question “why datafy” is to reject the premise that datafication can produce social and economic value, and instead to emphasize and center its corrosive effects on individual autonomy and the epistemic justice issues raised by the platform economy’s distortionary effects on knowledge production.

The relation between knowledge production, profit accumulation, and power preoccupies several observers. For example, Zuboff is clearly concerned with issues of epistemic justice that arise from the walled gardens and skewed paths of platform knowledge formation.⁸⁵ Others raise concerns about the challenges of cultivating self-knowledge and the capacity for self-formation in the shadow of such “surveillance empires.”⁸⁶ Still others (including one of the authors of this Article) consider the impact that the private curation of social knowledge forms has on public scientific inquiry and the future of social scientific work.⁸⁷

85. See Shoshana Zuboff, Opinion, *The Coup We Are Not Talking About*, N.Y. Times (Jan. 29, 2021), <https://www.nytimes.com/2021/01/29/opinion/sunday/facebook-surveillance-society-technology.html> (on file with the *Columbia Law Review*) (“In an information civilization, societies are defined by questions of knowledge—how it is distributed, the authority that governs its distribution and the power that protects that authority. . . . Surveillance capitalists now hold the answers . . . though we never elected them to govern. This is the essence of the epistemic coup.”); see also Lauren Jackson, *Shoshana Zuboff Explains Why You Should Care About Privacy*, N.Y. Times (May 21, 2021), <https://www.nytimes.com/2021/05/21/technology/shoshana-zuboff-apple-google-privacy.html> (on file with the *Columbia Law Review*) (last updated May 24, 2021) (“Instead of this being a golden age of the democratization of knowledge, it’s turned into something very different from what any of us expected. The last 20 years have seen, especially the last decade, the wholesale destruction of privacy.” (internal quotation marks omitted) (quoting Shoshana Zuboff)).

86. See, e.g., Brett Frischmann & Evan Selinger, *Re-Engineering Humanity* 30 (2018) (“When outsourcing becomes habitual, we become dependent on a third party for getting stuff done. At the extreme, dependency can result in deskilling. We can forget how to perform a task or become less capable of doing it. Or, we can lose the motivation to increase our knowledge and skills.”).

87. See Christopher J. Morten, Gabriel Nicholas & Salomé Viljoen, *Researcher Access to Social Media Data: Lessons From Clinical Trial Data Sharing*, 39 *Berkeley Tech. L.J.* 109, 112 (2024) (“[W]hen platforms themselves wield absolute control over which researchers get access to data (and how much, and on what terms), platforms can thwart critical research and shape the literature that emerges by selectively providing access to data.”); Jathan Sadowski, Salomé Viljoen & Meredith Whittaker, *Everyone Should Decide How Their Digital Data Are Used—Not Just Tech Companies*, 595 *Nature* 169, 169–70 (2021)

Strategies of social data value accumulation and knowledge economy distortion are related. After all, in the flurry of activity that accompanies the datafication of social life, companies *are* engaged in worldbuilding, creating a peculiar kind of knowledge of the social world that crowds out, destroys, and replaces others.⁸⁸ What's more, entities are acting on and enacting that knowledge back onto the world in ways that are both inscrutable and, perhaps worryingly, quite powerful. These concerns, while important, are not the central focus here. This Article is concerned with the political economic factors *driving* these developments.

Social data does not exist in the ether. It costs money to collect and store and analyze. Data centers must be rented, staffed, and kept cool.⁸⁹ Engineers and designers must be hired to design the technology environments in which data is collected. Privacy managers must be retained to bring such environments into compliance.⁹⁰ Data scientists must be hired to make sense of data and use it.⁹¹ It is undoubtedly true that some amount of these costs invoiced against future prediction value are puffery and speculation.⁹² So why are companies spending so much time, effort, and money on social data? At least part of the answer is *because* social data is indeed valuable to them, in some form or another, for some reason or another. Companies are not, as a rule, engaged in “anti-democratic” knowledge production for no reason, or due to conscious

(describing how private actors gatekeeping data interferes with researchers' ability to conduct their work).

88. For an entertaining and accessible argument that technology platforms destroy systems of meaning, erode expertise and trust, and replace knowledge with informational proxies, see generally Neil Postman, *Technopoly: The Surrender of Culture to Technology* (1992).

89. See Dan Greene, *Landlords of the Internet: Big Data and Big Real Estate*, 52 *Soc. Stud. Sci.* 904, 909 (2022) (“Data center landlords install, manage, maintain, and optimize tenants' servers; taking over tasks that carrier hotels leave to clients. This is the physical business arrangement for the cloud.”); see also Steven Gonzalez Monserrate, *The Cloud Is Material: On the Environmental Impacts of Computation and Data Storage*, MIT Schwarzman Coll. of Computing (Jan. 27, 2022), <https://mit-serc.pubpub.org/pub/the-cloud-is-material/release/1> [<https://perma.cc/C4LV-C59D>] (describing how data centers consume vast amounts of carbon, water, and electricity to cool and remain operational).

90. Ari Ezra Waldman, *Industry Unbound* 15–44, 101 (2021) (detailing the role of privacy compliance offers inside technology companies and the larger role of compliance in enforcing privacy law on the ground).

91. The average data scientist makes over \$100,000 a year. *Average Data Scientist Salary*, PayScale, https://www.payscale.com/research/US/Job=Data_Scientist/Salary [<https://perma.cc/4ZXT-G4VE>] (last visited Feb. 19, 2024). The data analytics company Stitch identified more than 11,400 employees identified as data scientists on LinkedIn in 2015 and found that the number of data scientists in the economy had doubled from four years prior. *The State of Data Science*, Stitch, <https://www.stitchdata.com/resources/the-state-of-data-science/> [<https://perma.cc/P5VG-7UE3>] (last visited Feb. 19, 2024).

92. Cf. *supra* note 21 and accompanying text.

malevolence.⁹³ Indeed, to understand those effects, it's worth thinking deeply about their causes.

* * *

Separating out the study of data's prediction value from its exchange value may prove helpful to understanding some challenges law faces in governing social data production and the political economy organized around such production.

As an initial matter, the primary project of Parts II and III is to distinguish prediction value from exchange value. This distinction helps to diagnose how and why entities go about cultivating, storing, and exploiting social data for gain, and how these activities meet, challenge, and transform legal forms. Clarifying the two modes of value production may also help index the chicanery that is, admittedly, rampant among certain corners of the digital economy, where overblown claims of how these two forms of value relate to one another may be used to obscure and befuddle.⁹⁴ Part II considers how companies translate or transform prediction value into exchange value.

II. THE BUSINESS OF SOCIAL DATA

Social data as a value form has encouraged the growth of business models and corporate behaviors that leverage prediction value to produce both wealth and power for companies and their investors. These business models and behaviors have important normative and legal implications. The capacity of existing law to grapple with those implications varies and is, in many cases, inadequate. This Part begins by cataloging three scripts that companies take when leveraging prediction value. It describes the business models and practices that have emerged as companies pursue these three scripts, which have important normative and legal implications. It then explores some of the business practices that companies use to accumulate social data—in particular, practices that focus on growth and expansion often at the expense of current profits. This Part highlights all of these business models and practices and their important normative and legal implications.

To be clear, data is important to the digital economy, but it is not the sole source of value nor is its acquisition and use the sole cause of contemporary digital firm strategies. While this Article argues that the importance of social data to business strategies has been underappreciated, it is not arguing that all value of digital companies is data value. This Part's primary aim is to provide a granular and reasonably

93. Zuboff, *Surveillance Capitalism*, *supra* note 30, at 513 (“Surveillance capitalism’s antidemocratic and antiegalitarian juggernaut is best described as a market-driven coup from above.”).

94. See *supra* note 20 and accompanying text.

systematic accounting of the various ways platforms and other firms may use data to produce value (and power). Nothing about this account should be read to suggest that data is the only driver of value in the digital economy. Indeed, it may not even be the largest driver of value; that is an empirical question beyond the scope of this Article. What this Part does argue is that data value is an important form of use value, whose cultivation prefigures several strategies to transform that value into profit. Many firms in the digital economy deploy these strategies in their various forms.

A. *The Three Scripts*

This section catalogs three scripts that companies take when attempting to leverage prediction value, as well as the business models and practices associated with those paths. These three scripts are not mutually exclusive. Companies may engage in different scripts at different points in their corporate lives. Some companies may never engage in some of the scripts. Some companies may engage in multiple scripts simultaneously. These companies may either engage in multiple scripts within a single business line or, perhaps more commonly, may operate in multiple business lines that are engaging in different scripts.

The first script is direct conversion of prediction value into exchange value.⁹⁵ This direct conversion transforms data about people into money for companies through means such as targeted advertising. The second script is indirect conversion of prediction value into exchange value by leveraging prediction value to improve products and services, reduce costs, develop new products and services, and expand into different business lines and industries.⁹⁶ The third script is not directly focused on converting prediction value into exchange value. Instead, in the third script, companies focus on transforming data about people into power for companies.⁹⁷

1. *Script One: Directly Converting Prediction Value Into Exchange Value.* — Companies primarily achieve the first script—the direct conversion of prediction value into exchange value—either through the sale and license of social data or through targeted advertising.⁹⁸

The sale and license of social data are the most obvious and legible way that companies can convert social data into money.⁹⁹ The data

95. See *infra* section II.A.1.

96. See *infra* section II.A.2.

97. See *infra* section II.A.3.

98. As discussed in the following section, the ability to predict and modify behavior that undergirds the premium charged for targeted advertising is also essential to pursuing the second script. See *infra* section II.A.2.

99. See Douglas B. Laney, *Infonomics: How to Monetize, Manage, and Measure Information as an Asset for Competitive Advantage* 13 (2018); see also Cohen, *Between Truth and Power*, *supra* note 4, at 48–74 (describing the development of the market for data about people, including data brokerage businesses); Sarah Lamdan, *Data Cartels: The Companies that Control and Monopolize Our Information* 12–15 (2023) (describing how

brokerage industry was valued at \$247.4 billion in 2022.¹⁰⁰ The industry is predicted to grow to \$407.5 billion by 2028.¹⁰¹ The global market for location data alone was estimated to be \$12 billion in 2021.¹⁰²

While sale or license is the most apparent means of directly monetizing social data, it is not the most common means of direct monetization. The transformation of social data into an income-producing asset for companies can take many other forms.¹⁰³ In reality, targeted advertising is the largest source of direct data monetization for companies.¹⁰⁴

publishing companies like Westlaw and Lexis have transformed into data analytics companies “selling raw data, and . . . structured information made from that raw data” allowing them to “multiply[] their existing information troves”); Bruce Schneier, *Data and Goliath* 51–53 (2015) (summarizing the history and practices of the data brokerage industry). In addition to the sale and license of data, companies also use their data to barter with other businesses for goods and services. Laney, *supra*, at 29.

100. Harsha Kiran, 10 Data Broker Statistics You Need to Know, *Techjury: Blog*, <https://techjury.net/blog/data-broker-statistics/> [<https://perma.cc/UQT2-MWXG>] (last updated Jan. 2, 2024).

101. *Id.*

102. Jon Keegan & Alfred Ng, There’s a Multibillion-Dollar Market for Your Phone’s Location Data, *The Markup* (Sept. 30, 2021), <https://themarkup.org/privacy/2021/09/30/theres-a-multibillion-dollar-market-for-your-phones-location-data> [<https://perma.cc/FT3U-GBEP>] (identifying forty-seven companies that are major players in the location data industry).

103. See Laney, *supra* note 99, at 11 (“Let’s dispel the notion right away that information monetization . . . is just about selling your data. It’s much broader than that.”); Birch & Muniesa, *supra* note 56, at 2 (identifying the process of capitalist transformation of data into revenues as deriving “durable economic rent” from data and defining rent as “the extraction of value through the ownership and control of an asset”).

104. See Duncan McCann, *New Econ. Found., I-Spy: The Billion Dollar Business of Surveillance Advertising to Kids* 6 (2021), https://iapp.org/media/pdf/resource_center/i_spy_the_billion_dollar_business_of_surveillance_advertising_to_kids.pdf [<https://perma.cc/7R8W-7BNA>] (discussing targeted advertising as the “primary business model of many digital companies”); Beauvisage & Mellet, *supra* note 57, at 77 (“The online marketing and advertising industries are a striking example of the dynamics of data assetization . . .”); Viljoen, *Relational Theory*, *supra* note 6, at 586–97 (describing the ability to “predict . . . [and] influence behavior” as the biggest source of revenue for tech companies, with advertising comprising the majority of that revenue); Zuboff, *Surveillance Capitalism*, *supra* note 30, at 27–196 (tracing the beginnings of “surveillance capitalism” through Google’s initial development of targeted advertising technology). As is explored in greater detail below, in reality, targeted advertising is a more complex example. See *infra* sections II.A.2–.3. It implicates not only the first script but a range of strategies to exploit prediction value. The intended focus here is the direct markup, or additional amount, platforms can charge advertisers for a targeted ad as opposed to an untargeted one. This markup presents (in theory) a direct exchange value amount companies are willing to pay for the superior prediction value offered by a targeted ad. Script two strategies include using prediction value to improve a platform’s advertising product by nudging users or experimenting with how ads are presented to maximize the likelihood that users will click. See *infra* section II.A.2. And script three strategies include accumulating consumers’ purchasing and other online behaviors to exert desired behavioral changes or to secure positions of market power. See *infra* section II.A.3.

Data about people allows companies to target their advertisements to the individuals who are most likely to be interested in the product.¹⁰⁵ Strollers are advertised to pregnant women but not teenage boys. Dutch ovens are advertised to avid cooks but not people who live on microwave dinners. Nicholas Negroponte presciently foresaw the potential of targeted advertising in the mid-1990s.¹⁰⁶ He presented the example of digital technology facilitating a person in the market for a car receiving nothing but car ads and, additionally, having those ads geographically tailored to include sales from local dealers.¹⁰⁷ Social data collection by companies has enabled Negroponte's predictions to come to fruition. Social media platforms like TikTok collect data not only on their users' activities on the platform but also track their movements across hundreds of thousands of other websites, thus allowing them to gauge potential purchases and offer advertisers access to users with purchase intents that match the advertisers' products.¹⁰⁸

Targeted advertising is a central way in which companies are able to take the prediction value that they draw from social data and turn it into exchange value. Companies can earn money by extracting social data, analyzing that data to divide people into categories based on salient features, and then auctioning ad space at a premium based on the premise that those ads are properly targeted to the most relevant people.¹⁰⁹ For many tech companies, this means of leveraging prediction value constitutes the lion's share of their revenues. For example, more than seventy-five percent of Alphabet's revenues came from online advertising in 2023.¹¹⁰ Even more stark is Meta Platforms, Inc., in which 98.6% of the

105. See Schneier, *supra* note 99, at 53–54 (“If you know exactly who wants to buy a lawn mower . . . you can target your advertising to the right person at the right time, eliminating waste. . . . And if you know details about that potential customer . . . your advertising can be even more effective.”); Joseph Turow, *The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth* 74–76 (2011) (explaining the development of behavioral targeting in advertising).

106. Nicholas Negroponte, *Being Digital* 179–80 (1995).

107. *Id.*

108. Shoshana Wodinsky, *TikTok Will Use Your Data to Fuel Its Multibillion-Dollar Shopping Mall—Whether You Know It or Not*, *MarketWatch* (Oct. 25, 2022), <https://www.marketwatch.com/story/tiktok-will-use-your-data-to-fuel-its-multibillion-dollar-shopping-mall-whether-you-know-it-or-not-11666653414> (on file with the *Columbia Law Review*).

109. See Nick Srnicek, *Platform Capitalism* 56–57 (2017) (describing this process of transforming user data into advertising from targeted revenues); Zuboff, *Surveillance Capitalism*, *supra* note 30, at 63–98 (“[Google] thus created . . . an asset class of vital raw materials At first those raw materials were simply ‘found’ Later those assets were hunted aggressively and procured largely through surveillance. [Google] simultaneously created a new kind of marketplace in which its proprietary ‘prediction products’ . . . could be bought and sold.”); Hal R. Varian, *Online Ad Auctions*, 99 *Am. Econ. Rev.* 430, 430 (2009) (explaining the mechanics of online advertising auctions by search engine companies).

110. Alphabet, Inc., *Annual Report (Form 10-K)* 11 (Jan. 31, 2024).

company's revenues came from advertising in 2023.¹¹¹ Targeted advertising directly translates prediction value into exchange value for companies. They are essentially selling to advertisers their capacity to predict which products will be most salient to specific consumers.¹¹²

2. *Script Two: Indirectly Converting Prediction Value Into Exchange Value.* — Companies transform data about people directly into revenue through the sale and license of data, as well as targeted advertising. But these methods of direct conversion are not the only means by which companies convert data about people into company revenues. In general, companies use social data to expand and improve their existing business operations. Through this expansion and improvement, companies will (eventually) increase revenues and profits, thus indirectly converting prediction value into exchange value. This emphasis on data accumulation and analysis is most strongly associated with Big Tech.¹¹³ As many of the examples in this section will demonstrate, however, using prediction value to improve and expand business operations is not limited to tech companies.

This section divides the means of indirect conversion into three categories and describes some of the business practices associated with each.¹¹⁴ These categories are: (a) lowering costs, (b) increasing and stabilizing revenues, and (c) expanding business operations.

a. *Lowering Costs.* — The predictive capacity of social data allows companies to lower their costs. As this section makes clear, some of these cost-lowering methods are benign while others are more controversial.

Prediction value can allow companies to identify inefficiencies in their production. As Erik Brynjolfsson and Andrew McAfee explained, “The data revolution has turned customers into unwitting business consultants,

111. See Meta Platforms, Inc., Annual Report (Form 10-K) 60 (Feb. 2, 2024). The company acknowledges in its annual report that “[w]e generate substantially all of our revenue from advertising.” *Id.* at 18.

112. Srnicek, *supra* note 109, at 57 (“What is sold to advertisers is therefore not the data themselves . . . but rather the promise that . . . software will adeptly match an advertiser with the correct users when needed.”).

113. See, e.g., Michael Schrage, MIT Initiative on the Digit. Econ., Rethinking Networks: Exploring Strategies for Making Users More Valuable 1 (2016), https://ide.mit.edu/wp-content/uploads/2017/10/Rethinking-Networks_0.pdf?x19853 [<https://perma.cc/S6G2-TWHB>] (quoting “media infopreneur” Tim O’Reilly, who coined the term “Web 2.0,” as stating that “[a] true Web 2.0 application gets better the more people use it” and explaining that “Google gets smarter every time someone clicks on an ad . . . [a]nd it immediately acts on that information to improve the experience for” others).

114. These categories and their accompanying descriptions are not meant to be exhaustive. This section highlights some of the prominent ways that companies in the digital economy use social data value, focusing on practices with particularly important legal and normative justifications. It also aims to acknowledge the existence of both harmful business practices and socially beneficial ones.

as our purchases and searches are tracked to improve everything from websites to delivery routes.”¹¹⁵

By identifying and eliminating these inefficiencies, companies are able to streamline their operations and lower costs. These productivity gains are often broadly beneficial. For example, social data that allows grocery stores to more accurately predict the exact state of ripeness that consumers prefer for their bananas and to adjust their purchase and delivery timing accordingly would increase the grocery stores’ sales, prevent food waste, and provide consumers with optimally delicious bananas.

The healthcare industry is one area where data analysis has been used to identify and streamline inefficiencies and achieve productivity gains.¹¹⁶ Intel recently partnered with a French hospital and used big data analytics to produce fifteen-day predictions of emergency visits and hospital admissions, which then allowed the hospital to plan their staffing to meet the anticipated needs.¹¹⁷ The hospitals should, therefore, be able to lower their costs by not overstaffing during slower periods. And the patient experience should be improved by ensuring adequate staffing when they receive treatment.

Companies’ use of prediction value to lower costs have also led to more controversial techniques. Efforts to streamline production and maximize productivity based on social data can lead to what Zephyr Teachout refers to as “extreme Taylorism.”¹¹⁸ Social data applied to extreme monitoring of workers has allowed longstanding principles of worker management to be applied with a far greater degree of precision and pervasiveness. Employers can use data in real time to increase or decrease pay based on the employee’s efficiency. For example, they may

115. Erik Brynjolfsson & Andrew McAfee, *The Big Data Boom Is the Innovation Story of Our Time*, *The Atlantic* (Nov. 21, 2011), <https://www.theatlantic.com/business/archive/2011/11/the-big-data-boom-is-the-innovation-story-of-our-time/248215/> (on file with the *Columbia Law Review*) (emphasis omitted).

116. See Sabyasachi Dash, Sushil Kumar Shakyawar, Mohit Sharma & Sandeep Kaushik, *Big Data in Healthcare: Management, Analysis and Future Prospects*, 6 *J. Big Data*, no. 54, 2019, at 1, 22–24 (explaining ways in which the healthcare industry is converting the potential of big data to “bolster the existing pipeline of healthcare advances”).

117. Kyle Ambert, Sébastien Beaune, Adel Chaibi, Luc Briard, Amit Bhattacharjee, Venkatesh Bharadwaj, Krishna Sumanth & Kathleen Crowe, *Intel, White Paper: French Hospital Uses Trusted Analytics Platform to Predict Emergency Department Visits and Hospital Admissions 1* (2016), <https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/french-hospital-analytics-predict-admissions-paper.pdf> [<https://perma.cc/35HP-763Y>].

118. Zephyr Teachout, *Algorithmic Personalized Wages*, 51 *Pol. & Soc’y* 436, 442 (2023) [hereinafter Teachout, *Algorithmic Personalized Wages*]. Taylorism, also called “scientific management,” is an approach to labor management aimed at maximizing labor productivity. See Frederick Winslow Taylor, *The Principles of Scientific Management* 12 (1911). Pioneered by Frederick Winslow Taylor at the end of the nineteenth century, it is associated with close monitoring of workers to deter, detect, and correct inefficiencies, and, in some cases, pay based on the worker’s specific output. *Id.*

be able to detect multitasking and decrease pay accordingly—or even reward or penalize employees based on their attitudes, as detected by software tracking facial expressions.¹¹⁹

Extreme Taylorism is just one form of the broader controversial practice of algorithmic wage discrimination.¹²⁰ Gamification is the practice of applying behavioral science to use scoring, competition, and rewards (in place of higher wages) to motivate employees.¹²¹ One example of gamification comes from DoorDash. DoorDash pays drivers extra if they succeed in “challenges,” which typically involve completing a set number of deliveries in a set period of time, along with other possible requirements.¹²² Studies have shown that these techniques induce workers to work longer hours, or attract similar numbers of workers, for lower pay overall.¹²³

Behavioral price discrimination is the practice of adjusting wages based on factors unrelated to employees’ productivity at work. Factors such as employees’ health status, or even their credit scores, gathered outside of the workplace, could be used by companies to differentiate wages.¹²⁴

Dynamic labor pricing is another possible means of algorithmic wage discrimination. While dynamic labor pricing is used in response to supply and demand imbalances, evidence from Veena Dubal and others suggests that companies can respond to those imbalances with real-time wage adjustments based on social data value.¹²⁵ For example, wage experimentation may be used by companies to pinpoint personalized or near-personalized reserve price wages to accomplish the highest levels of productivity for the lowest labor costs under dynamic conditions of shifting supply and demand.¹²⁶ Gamification or other behavioral techniques are

119. Teachout, *Algorithmic Personalized Wages*, supra note 118, at 442.

120. Veena Dubal, *On Algorithmic Wage Discrimination*, 123 *Colum. L. Rev.* 1929, 1952–61 (2023).

121. Teachout, *Algorithmic Personalized Wages*, supra note 118, at 442–43.

122. *How Dasher Pay Works*, DoorDash, https://help.doordash.com/dashers/s/article/How-is-Dasher-pay-calculated?language=en_US (on file with the *Columbia Law Review*) (last visited Feb. 21, 2024).

123. Dubal, supra note 120, at 1952–61; Christopher L. Peterson & Marshall Steinbaum, *Coercive Rideshare Practices: At the Intersection of Antitrust and Consumer Protection Law in the Gig Economy*, 90 *U. Chi. L. Rev.* 623, 633–34 (2023) (detailing the practice of offering drivers disadvantageous rides in the context of bonus challenges, “knowing that drivers will only accept the trips to attain the bonus”).

124. Teachout, *Algorithmic Personalized Wages*, supra note 118, at 443.

125. *Id.* at 443–44; see also Dubal, supra note 120, at 1934 & n.15 (simplifying Teachout’s taxonomy to “wages based on productivity analysis alone,” as is the case with extreme Taylorism, and “wages based on productivity, supply, demand, and other personalized data intended to minimize labor costs”).

126. Teachout, *Algorithmic Personalized Wages*, supra note 118, at 444–45. Uber chief economist Jonathan Hall and coauthors confirmed that, at a certain point, drivers get decreasing returns from working longer hours. Dubal, supra note 120, at 1970; see also Peterson & Steinbaum, supra note 123, at 631 (detailing how in markets in which drivers

then used in concert with personalized wages to meet demand in the most cost-efficient way. By experimenting with dynamic price setting, Uber has found that it can “prod drivers into working longer and harder—and sometimes at hours and locations that are less lucrative for them.”¹²⁷

A common feature of these methods is that they are enabled by accumulating social data and applying it to predict and manage worker behavior and responsiveness. As the examples above show, evidence suggests these social data value techniques may induce labor supply even at inefficient (from the perspective of the worker) price levels. Thus, the role of social data in the context of algorithmic wage discrimination has particularly important normative and legal implications. Brishen Rogers has argued that the practices of worker surveillance that allow for algorithmic wage discrimination and accompanying reductions in overall wages result in class disempowerment.¹²⁸ And Dubal has argued that the practice of algorithmic wage discrimination goes against the norm of equal pay for equal work that serves as the basis for many of the country’s antidiscrimination laws, as well as the general norm of fairness within the wage setting.¹²⁹ While these practices are currently most common in the context of the gig economy, they are moving to more traditional employment settings.¹³⁰

Companies use prediction value to lower their costs in various ways—some beneficial, others more harmful. On the other side of the balance sheet, companies can convert prediction value to exchange value by increasing or, in some instances, stabilizing their revenues.

b. *Increasing and Stabilizing Revenues.* — Social data offers companies insights that they can use to increase or stabilize revenues.¹³¹ Stabilizing revenues refers to avoiding a decline in revenue due to lost customers or

were able to set their own prices in lieu of Uber, “pay for drivers increased substantially while prices charged to customers did not change”).

127. Dubal, *supra* note 120, at 1949 (internal quotation marks omitted) (quoting Noam Scheiber, *How Uber Uses Psychological Tricks to Push Its Drivers’ Buttons*, N.Y. Times (Apr. 2, 2017), <https://www.nytimes.com/interactive/2017/04/02/technology/uber-drivers-psychological-tricks.html> (on file with the *Columbia Law Review*)).

128. Brishen Rogers, *Worker Surveillance and Class Power*, LPE Project Blog (July 11, 2018), <https://lpeproject.org/blog/worker-surveillance-and-class-power/> [<https://perma.cc/WE75-QX3N>] [hereinafter Rogers, *Worker Surveillance*].

129. See Dubal, *supra* note 120, at 1957–61. For further discussions of the legal and normative impacts of algorithmic wage discrimination, see generally Alex Rosenblat, *Uberland: How Algorithms Are Rewriting the Rules of Work* (2018); Peterson & Steinbaum, *supra* note 123.

130. Zephyr Teachout, *Surveillance Wages: A Taxonomy*, LPE Project Blog (Nov. 6, 2023), <https://lpeproject.org/blog/surveillance-wages-a-taxonomy/> [<https://perma.cc/ZQK2-NUBP>].

131. See Viljoen, *Relational Theory*, *supra* note 6, at 589 (noting that techniques designed to predict behavior “point toward new avenues of growth for the data economy”); Rogers, *The Digital Transformation Playbook*, *supra* note 84, at 107–08.

a need to lower prices. Companies can indirectly convert prediction value into exchange value by utilizing both of these strategies.

Companies can apply insight from social data to optimize their business operations in ways that increase revenues. For example, by consolidating and analyzing databases, retailer Dollar General discovered a pattern of customer purchases peaking near closing time.¹³² The company inferred from this that later store hours would better accommodate customer needs and saw a 9.5% increase in sales within a year.¹³³

Prediction value can also result in companies increasing their revenues by accessing new streams of potential customers. Fintech is an important example of this. Fintech platforms leverage a wide array of social data and machine-learning techniques to make consumer-lending decisions.¹³⁴ The social data used by fintech companies moves beyond measures that have traditionally been used by financial institutions, such as income and credit scores, and incorporates alternative data into their lending decisions.¹³⁵ This alternative data can range from personal health information to data gleaned from social media activity.¹³⁶ The prediction value that comes from this data allows fintech companies to extend credit to borrowers who might not qualify absent this additional social data.¹³⁷ Tapping this new stream of customers increases their revenues.¹³⁸

The examples discussed thus far are uses of prediction value to increase revenues that are mutually beneficial to companies, consumers, and at times, society more broadly. But other uses of prediction value to increase revenues can produce more ambiguous outcomes. A central example is the way that social data can enable predatory pricing behaviors.

These predatory pricing behaviors take a couple of flavors. The first is price discrimination. Economist Joseph Stiglitz explains, “Data can be used to extract consumer surplus by charging different customers different prices Companies that prosper are not those that are most efficient and that do the best job satisfying customers but those that are

132. Laney, *supra* note 99, at 40.

133. *Id.*

134. See Christopher K. Odinet, *Consumer Bitcredit and Fintech Lending*, 69 *Ala. L. Rev.* 781, 788–90 (2018) (providing an overview of fintech lending business models).

135. Marco Di Maggio, Dimuthu Ratnadiwakara & Don Carmichael, *Invisible Primes: Fintech Lending With Alternative Data 1* (Nat’l Bureau of Econ. Rsch., Working Paper No. 29840, 2022), https://www.nber.org/system/files/working_papers/w29840/w29840.pdf [<https://perma.cc/2XEF-6MHZ>] (identifying the use of alternative data as a “key feature” of fintech business models).

136. Odinet, *supra* note 134, at 785.

137. Di Maggio et al., *supra* note 135, at 26.

138. This is also an example of reducing costs because the better decisionmaking reduces rates of default. *Id.*

best at exploitation, at extracting this consumer surplus.”¹³⁹ Social data gives companies the capacity to predict consumer behavior in a way that allows them to charge close to the maximum amount that the consumer would be willing to pay, versus setting the same price for all consumers.¹⁴⁰ In *Texas v. Google*, for instance, the State alleged that Google induced advertisers to bid their true value to develop detailed predictions of each advertiser’s personalized willingness to pay.¹⁴¹ The company then overrode preset price floors to use advertisers’ true bids against them by secretly generating unique and custom per-buyer floors based on what buyers had bid in the past.¹⁴² While each individual transaction between customer and company under such conditions is still, in theory, locally efficient, it does mean that at least some consumers will face higher prices for identical products. Moreover, as Stiglitz notes, the general practice of such price discrimination can be harmful to overall market efficiency in the longer term, especially in the context of proprietary exchange mechanisms where such prices are not subject to public scrutiny.¹⁴³

Companies use various forms of social data to predict consumer behaviors in order to engage in price discrimination. Online consumers can be identified through a variety of means, such as cookies, which allow companies to track consumers’ digital footprints. These digital footprints allow companies to predict purchasing preferences.¹⁴⁴ Merely knowing the physical location of the online shopper can allow for price discrimination because the online seller can tailor their price to those in local brick-and-mortar stores.¹⁴⁵ Amazon has the capacity to collect data on behaviors such as a consumer hovering a mouse over a product or viewing a product multiple times, which can then allow Amazon to display higher prices for the consumers it believes are most likely to purchase the product.¹⁴⁶ And

139. Julia Angwin, *How AI Could Undermine an Efficient Market Economy*, The Markup: Hello World Newsl. (June 25, 2022), <https://themarkup.org/newsletter/hello-world/how-ai-could-undermine-an-efficient-market-economy/> [https://perma.cc/9DZY-S49V] (internal quotation marks omitted) (quoting Joseph E. Stiglitz).

140. Id.; see also Andy Fitch, *Different People, Different Prices: Talking to Joseph E. Stiglitz*, L.A. Rev. Books (Jan. 24, 2020), <https://blog.lareviewofbooks.org/interviews/different-people-different-prices-talking-joseph-e-stiglitz/> [https://perma.cc/6AQA-B4S6].

141. See Second Amended Complaint at 86, *Texas v. Google LLC*, No. 4:20-CV-957-SDJ (E.D. Tex. Filed Aug. 4, 2021), 2021 WL 4146613 (“Google deployed a bid optimization scheme based on predictive modeling . . . [and] [w]ith this new bid optimization, . . . Google re-engineered its ability to trade ahead of its rivals.”).

142. Id. at 65–66.

143. For further discussion of the impact of data and AI on market efficiency, see generally Joseph E. Stiglitz, *People, Power, and Profits: Progressive Capitalism for an Age of Discontent* (2019).

144. See Akiva A. Miller, *What Do We Worry About When We Worry About Price Discrimination? The Law and Ethics of Using Personal Information for Pricing*, 19 J. Tech. L. & Pol’y 41, 45 (2014).

145. See Fitch, *supra* note 140.

146. See Rory Van Loo & Nikita Aggarwal, *Amazon’s Pricing Paradox*, 37 Harv. J.L. & Tech. 1, 27 (2023).

while evidence on this behavior is extremely limited, some have even speculated that digital companies can listen to conversations via devices like Alexa to determine the price a consumer is willing to pay for a product.¹⁴⁷

Another application of targeted price discrimination is companies identifying competitors' customers and offering those customers lower prices than they offer their established customers to lure them away from their competitors.¹⁴⁸ This ability to use price discrimination to lure in new customers is particularly problematic given yet another element of predatory pricing—manipulating customer behavior in ways that reduce customers' ability or likelihood to exit. Amazon, for example, uses both data from customers' past purchases and data on general purchasing behaviors to dynamically manage product price and presentation to exploit customers' behavioral biases.¹⁴⁹ This dynamic pricing and presentation, in turn, hinders customers' ability to find the lowest-priced products available on Amazon's website and to comparison shop with other retailers.¹⁵⁰ This inability to comparison shop both within and outside the platform limits consumers' ability to use traditional exit mechanisms to discipline Amazon. Thus, Amazon's use of social data value to engage in personalized pricing and presentation strategies stabilizes the company's revenues but also provides more opportunities to increase revenues via price discrimination.

This section has explored just a few of the ways in which companies are able to use prediction value to increase and stabilize revenues, thus converting prediction value into exchange value. The following section explores another means of indirect conversion of prediction value into exchange value—expanding business operations.

c. Expanding Business Operations. — Companies can leverage prediction value to expand business operations in a few different ways.¹⁵¹

147. Christopher R. Leslie, *Predatory Pricing Algorithms*, 98 N.Y.U. L. Rev. 49, 70–71 (2023); see also Maurice E. Stucke & Ariel Ezrachi, *How Digital Assistants Can Harm Our Economy, Privacy, and Democracy*, 32 Berkeley Tech. L.J. 1239, 1265–67 (2017) (“Given its ubiquity in the home, a digital assistant will have even more personal data, more opportunities to observe how users respond . . . and more opportunities to learn the right price point for that user.”).

148. Leslie, *supra* note 147, at 72–73.

149. See Van Loo & Aggarwal, *supra* note 146, at 23–29.

150. *Id.*

151. See Laney, *supra* note 99, at 68 (identifying developing new products and services as a method of monetizing data); MIT Tech. Rev. Custom, *The Rise of Data Capital 2* (2016), http://files.technologyreview.com/whitepapers/MIT_Oracle+Report-The_Rise_of_Data_Capital.pdf [<https://perma.cc/5YMF-88C5>] (“Data is now a form of capital, on the same level as financial capital in terms of generating new digital products and services.”); Sadowski, *When Data Is Capital*, *supra* note 29, at 6 (identifying “build[ing] stuff” as one of the ways that value can be derived from data capital (emphasis omitted)). For discussion of the use of big data in new products through a case study of an electronics company, see

They could use insights from social data to better cultivate new products and services in order to enter new business lines within their own industries. Or they could use social data to expand their operations into entirely new industries.

With respect to developing new products and services within their own industries, prediction value allows companies to determine which offerings will be most desirable to customers by taking into account knowledge that the companies have on their preferences and needs from analyzing social data.¹⁵² The prediction value derived from vast amounts of social data replaces traditional product-development strategies that relied on expert intuition and smaller data-gathering activities such as focus groups. This leads to a lower-risk and higher-efficiency approach to product development.¹⁵³ The companies are then able to garner revenues through these new products and services, all while needing to price in less risk and uncertainty associated with entering a new venture—thus indirectly converting prediction value into exchange value.

There are a multitude of examples of companies using social data to effectively expand within their own industries. Netflix became a creator of entertainment content rather than simply a streaming service for third-party content. Netflix entered the world of content creation with the advantage of social data on the viewing habits of millions of its subscribers. Through this social data, Netflix was able to predict factors that would lead to the success of newly created shows, such as the appeal of different subject matters and actors.¹⁵⁴ Its first original series, *House of Cards*, debuted in 2013 to great success, and the company continues to leverage prediction value when creating original content.¹⁵⁵ Food delivery services, such as DoorDash, have moved outside of restaurant delivery into grocery delivery. When DoorDash launched DashMart in 2020, commentators highlighted that its customer data, as well as its internal data on optimizing delivery, would give DoorDash a competitive edge in providing this new type of food delivery service.¹⁵⁶ Fintech is another area in which using

generally Yuanzhu Zhan, Kim Hua Tan, Yina Li & Ying Kei Tse, *Unlocking the Power of Big Data in New Product Development*, 270 *Annals Operations Rsch.* 577 (2018).

152. Zhan et al., *supra* note 151, at 580 (“Companies that are able to recognise customers’ latent needs and to have this data inform new product features or entire products will be much more likely to develop successful novel products.” (citations omitted)).

153. See Rogers, *The Digital Transformation Playbook*, *supra* note 84, at 4–9 (describing ways in which digital business models diverge from traditional models particularly by enabling a dynamic customer feedback loop to foster rapid innovation).

154. Jon Markman, *Netflix Harnesses Big Data to Profit From Your Tastes*, *Forbes* (Feb. 25, 2019), <https://www.forbes.com/sites/jonmarkman/2019/02/25/netflix-harnesses-big-data-to-profit-from-your-tastes/?sh=1aa093d266f> (on file with the *Columbia Law Review*) (explaining that Netflix relied on algorithmic predictions of success based on subject matter, viewership trends, and appeal of actors when contracting for *House of Cards*).

155. *Id.* (“Netflix is quietly transforming the entertainment industry with data.”).

156. Jessica Dumont, *DoorDash Launches New Online Convenience Store DashMart*, *RetailDive* (Aug. 6, 2020), <https://www.retaildive.com/news/door-dash-launches-new->

social data gathered through one service has been used to develop new products. Square began as a merchant-services provider before expanding into loans for small businesses.¹⁵⁷ Square uses data gathered from its merchant services technology, such as a business's volume and frequency of sales, to make its lending decisions.¹⁵⁸

One noteworthy example of the use of prediction value to both streamline operations and better predict and supply new products is the use of social data by the fast-fashion company Shein. Like other online retail companies, Shein collects detailed information from its website and app on what items shoppers search for; which pages they spend time on; and how they respond to similar items shown by the site—which are typically personalized suggestions based on the shopper's prior searches on Shein or other sites.¹⁵⁹ This data, combined with data obtained through digital advertising, informs generalized insights linking changes in shopping patterns to predictions about new and growing trends. Shein is distinct in that it uses predictions about shifts in demand far more aggressively and systematically to streamline its operations.¹⁶⁰ It experiments with products identified as growing trends, and it then place larger orders with manufacturers based on analysis of which products show the greatest promise.¹⁶¹ Manufacturers in turn have access to systems that monitor real-time demand and signal what items are likely to be ordered next.¹⁶² Shein's combination of detailed customer behavioral tracking with data-intensive, on-demand manufacturing techniques allows the company to transform an uptick in shopper search keywords into a responsive style available for purchase in less than two weeks.¹⁶³

Prediction value from social data gathered in one industry can also be used by companies to gain a competitive advantage as they expand into completely different industries. For example, an airline has leveraged

online-convenience-store-dashmart/583047/ [https://perma.cc/ZD39-K7X4]; see also Chris Albrecht, With DashMart, DoorDash Is Creating Its Own Ghost Convenience Stores, *The Spoon* (Aug. 5, 2020), https://thespoon.tech/with-dashmart-doordash-is-creating-its-own-ghost-convenience-stores/ [https://perma.cc/K7ND-HS7B] (“In addition to keeping all the revenue from sales through DashMart, DoorDash also gets more data around customers and purchases and can better optimize its own inventories.”).

157. Molefe Choane, How Square Capital Uses Traditional and Non-Traditional Data Sources to Extend Loans to Merchants, HBS Digit. Initiative (Mar. 23, 2021), https://d3.harvard.edu/platform-digit/submission/how-square-capital-uses-traditional-and-non-traditional-data-sources-to-extend-loans-to-merchants/ [https://perma.cc/C8XD-D63W].

158. *Id.*

159. The ‘Secret Sauce’ that Helps Shein Predict the Next Hot Trend Has Nothing to Do With AI, Company Exec Says, *Bus. Insider* (Aug. 18, 2023), https://www.businessinsider.com/the-secret-sauce-behind-sheins-on-demand-fashion-2023-8 (on file with the *Columbia Law Review*).

160. *Id.*

161. *Id.*

162. *Id.*

163. *Id.*

prediction value stemming from social data in its loyalty program, combined with data from users' wearable devices, to expand into the insurance industry.¹⁶⁴ TikTok and other social media companies have attempted to integrate shopping directly into their platforms, using data on users to predict the products they are most likely to buy.¹⁶⁵ This business model is referred to as "social commerce."¹⁶⁶ Verily, a life sciences company owned by Alphabet Inc., Google's parent company, entered the data-driven life sciences space with the advantage of Google's data processing power¹⁶⁷ and has since used its data to expand into the health insurance market.¹⁶⁸ Section II.B below addresses in more detail how using prediction value to gain competitive advantage in other industries drives many aggressive merger and acquisition strategies.

Companies using social data from one industry as an opportunity to expand into entirely new industries exemplifies the stem-cell-like nature of social data discussed in Part I above. Social data's predictive capacities arm the companies that accumulate it with advantages across broad swaths of the economy. Companies use prediction value stemming from social data to lower costs, increase and stabilize revenues, and expand business operations. Each of these methods of leveraging prediction value indirectly converts prediction value into exchange value by allowing companies to earn greater profits.

3. *Script Three: Converting Prediction Value Into Economic and Political Power.* — The first two scripts both involve social data being used to generate monetary exchange value for companies. Entities following these scripts leverage prediction value to achieve business profits, albeit in ways that may raise new legal and normative concerns or amplify preexisting ones as compared to how companies have pursued value in the past. The third script is different: Power is at the center of the third script.

In a sense, script three coming last in this Article's taxonomy is a bit misleading. The prior scripts can be viewed as a suite of options for companies to convert the power of prediction value into profitmaking activity: to exploit, spend, or use the general power stored via social data

164. Alex Koster & Konrad von Szczepanski, Building a Business From Data Is Hard—Here's How the Winners Do It, Bos. Consulting Grp. (June 24, 2020), <https://www.bcg.com/publications/2020/how-winners-build-business-from-data> [<https://perma.cc/4RPC-GPBK>] (noting that the airline paired frequent-flier and wearable device data to develop its insurance offering).

165. Wodinsky, *supra* note 108.

166. *Id.*

167. Sean Captain, Google Life Sciences Rebrands as Verily, Uses Big Data to Figure Out Why We Get Sick, Fast Co. (Dec. 7, 2015), <https://www.fastcompany.com/3054352/google-life-sciences-rebrands-as-verily-uses-big-data-to-figure-out-why-we-get-si> [<https://perma.cc/M87J-YSHW>].

168. Heather Landi, Alphabet's Verily Breaks Into Stop-Loss Health Insurance Market Backed by Swiss Re, Fierce Healthcare (Aug. 25, 2020), <https://www.fiercehealthcare.com/payer/alphabet-s-verily-breaks-into-stop-loss-health-insurance-market-backed-by-swiss-re> [<https://perma.cc/L2SC-FW3U>].

toward a specific (profit-enhancing) end. But script three describes how, in its primary form, prediction value confers the power to apprehend, shape, and thus exert some measure of control over people's (or other entities') future behavior on its holder.¹⁶⁹ This power does not need to be converted into exchange value via one of the strategies discussed above to be valuable to a firm. In the third script, the utility of social data as a factor of production is to produce and store a form of power that prefigures any single plan for how to put that power toward wealth creation.¹⁷⁰ Script three thus describes both the residual form of unconverted data value and prediction value's initial form.

Script three catalogues ways that the social data value is put toward strategies of innovation to *obtain*, *hold onto*, and *make the most of* market power.¹⁷¹ Companies are of course ultimately interested in generating wealth and profits. But this drives companies not only to engage in the daily activity of profitmaking, as described in scripts one and two above, but also to engage in meta-strategies meant to help hold onto and maximize those profits,¹⁷² their market position, and the business strategies they used to obtain such profits—in other words, to cultivate and retain market power.¹⁷³ The cultivation of data power via script three gives

169. Again, these people or entities can be the data subjects from whom data is being collected, but they don't have to be. See *supra* Introduction, Part I (discussing social data).

170. Some readers and other scholars upon whose work this Article draws may reject the assertion that the third script is distinct from the second script. They could argue that all companies ultimately exist to earn profits, and any delay in converting prediction to exchange value is merely a strategy to achieve greater exchange value in the future. For the reasons articulated in this section, this Article contends that power alone is a driver of company behavior. Yet, the central argument—that various legal fields are failing to properly recognize social data as a value form and that these failures pose significant potential harms—holds even in the absence of the third script being a separate driver of company behavior.

171. This Article uses the term “market power” as defined by innovation economists: the ability of a company to insulate itself from conditions of competitive pricing. On this account, market power is the capacity for a firm to extract quasi-rents, defined as prices above what would be a competitive level. See Daniel Francis & Christopher Jon Sprigman, *Antitrust: Principles, Cases, and Materials* 49 (2023) (“Market power is the ability of an individual supplier to charge a price that is above the competitive level . . .”).

172. On firms' rational interest in creating conditions for the capture of quasi-rents and in pursuing strategies to increase the magnitude and extend the duration of quasi-rents, see Yochai Benkler, *Power and Productivity: Institutions, Ideology, and Technology in Political Economy*, in *A Political Economy of Justice* 27, 38 (Danielle Allen, Yochai Benkler, Leah Downey, Rebecca Henderson & Josh Simons eds., 2022).

173. In theory, market power may be obtained via (1) productivity-enhancing innovation that allows companies to do things their competitors cannot (i.e., technical or other forms of innovation that grow the proverbial pie); (2) distribution-shifting innovation, also called “sabotage,” that enables companies to weaken or erode competitive conditions, insulating them from market discipline (i.e., innovation that reallocates the existing pie); and (3) strategies to influence political and regulatory conditions that would otherwise prevent or constrain quasi-rent extraction (or interfere with some other political aim). On firms pursuing strategies of both productivity and sabotage as a means to obtain market power, see Yochai Benkler, *Structure and Legitimization in Capitalism: Law, Power and*

companies flexibility to choose among these various options, or even to pursue multiple at once.

Systematic evidence linking firms' usage of data with specific innovation strategies or excess profits is still being developed, and debates about what conclusions can be drawn about these behaviors are ongoing.¹⁷⁴ Nevertheless, evidence exists to associate companies' aggressive accumulation of data value with strategies of innovation focused

Justice in Market Society 9–11 (Oct. 26, 2023), <https://ssrn.com/abstract=4614192> [<https://perma.cc/NJ8U-WU6S>] (unpublished manuscript).

On the sabotage account of using data power for innovation, see Kean Birch, Margaret Chiappetta & Anna Artyushina, *The Problem of Innovation in Technoscientific Capitalism: Data Rentiership and the Policy Implications of Turning Personal Digital Data Into a Private Asset*, 41 *Pol'y Stud.* 468, 468–79 (2020) (detailing the role of data assets in allowing firms to engage in innovation strategies of “data rentiership,” extracting value from data via relationships of ownership and control rather than using it to deliver new products, services, and markets); see also Shapiro, *Platform Sabotage*, *supra* note 21, at 203–15 (applying the concept of sabotage to study relationships between gig workers and gig platforms). The notion that firms are indifferent between innovation that enhances productivity and forms of “sabotage” comes from Thorstein Veblen. See Thorstein Veblen, *Absentee Ownership and Business Enterprise in Recent Times* 181 (1923) (“[T]he business men who control industry [must] guard against an over-supply of their industrial output,—a running balance of sabotage on production with a view to maintain prices. It is the immediate effect of a surplus output . . . to inflate prices . . . reliev[ing] the need of this businesslike sabotage”); Thorstein Veblen, *On the Nature and Uses of Sabotage* 8 (1919) (“And the ways and means of this necessary control of the output of industry are always and necessarily something in the nature of sabotage—something in the way of retardation, restriction, withdrawal, unemployment of plant and workmen—whereby production is kept short of productive capacity.”).

174. For example, some argue that innovations from data value use should be considered a form of true productivity-enhancing technological advance, while others argue these innovations are a form of rentiership that has not improved products or services but simply structured existing relationships to better facilitate rentiership. See Evgeny Morozov, *Critique of Techno-Feudal Reason*, 133/134 *New Left Rev.* 89 (2022) (arguing against the thesis that the dominance of large technology companies that lock users and workers into relations of data extraction and dispossession herald a form of “technofeudalism” and arguing that several business tactics of large technology platforms indicate investment in genuine technological improvements characteristic of classic capitalist investment); see also Cédric Durand, *Scouting Capital's Frontiers: Reply to Morozov's 'Critique of Techno-Feudal Reason'*, 136 *New Left Rev.* 29 (2022) (describing how the technofeudal hypothesis complements alternative theories such as globalization and financialization); Rikap, *Capitalism as Usual?*, *supra* note 58, at 145 (arguing that the digital sector has created novel “relations of production” such that the capitalism-as-usual model characterizes not the digital economy as usual but rather a new form: “intellectual monopoly capitalism”); Timothy Erik Ström, *Capital and Cybernetics*, 135 *New Left Rev.* 23, 24 (2022) (“[N]either Morozov's same-as-ever capitalism nor Durand's technofeudalism succeed in grasping the novel dynamics of a capitalist sector founded on networked computing-machines, tracing its conception to the US military-industrial complex.”); Jodi Dean, *Same as It Ever Was?*, *New Left Rev.*: Sidecar (May 6, 2022), <https://newleftreview.org/sidecar/posts/same-as-it-ever-was> [<https://perma.cc/9VMU-KSEW>] (“[T]he ongoing process of separation is not a ‘going back’ to historical feudalism, as Morozov would have it, but a reflexization, such that capitalist processes long directed outward—through colonialism and imperialism—turn in upon themselves.”).

on rentiership, the excess profits of market power, and the persistence of oligopolies in the digital economy.¹⁷⁵ Companies also use the power cultivated via prediction value to evade or influence regulation.¹⁷⁶

Beginning with rentiership, Birch and his coauthors argue that firms' strategies to accumulate data value via ownership and control of data assets present a source of rent-based market power from which firms may then extract excess profits.¹⁷⁷ Revisiting some of the earlier examples helps to break down this claim of data rentiership as a form of market power in finer detail. Consider for example Google's sale of access to data value in the form of targeted ads discussed in the first script. The increased price Google can charge for placing a targeted ad (as opposed to the price of an untargeted ad) is a script one account of data value—data value directly converted into exchange value. But Google is separately motivated to obtain large pools of data¹⁷⁸ and enclose that data, to produce more accurate (and thus more valuable) prediction value than potential competitors. This aggregate data pool itself is not converted into exchange value; the “prediction service” it allows Google to offer, discussed in the first script, is.¹⁷⁹ The technical and legal strategies Google uses to collect and protect its data (namely, to extend ownership and control over that data) are concerned with producing and holding onto a source of market power: an innovation asset pool that (in theory) gives Google an enduring edge over its competitors not just in ad placement but for any number of applications and future innovations.

In other words, the social data value itself gives Google a competitive edge.¹⁸⁰ Google's store of social data value is a source of power, separable from various strategies the company can pursue to take advantage of that power, including strategies to convert data value into exchange value in

175. See, e.g., Majority Staff of H. Subcomm. on Antitrust, Com. & Admin. L., H. Comm. on the Judiciary, 117th Cong., Investigation of Competition in Digital Markets 145 (Comm. Print 2020) (noting that Google reported profit margins more than three times that of the average U.S. firm in nine of the previous ten years); Pierre Collin & Nicolas Colin, Task Force on Taxation of the Digital Economy 44 (2013) (Fr.), <https://nebula.wsimg.com/f722d8a16e3e827b1030e7608c1ff84e?AccessKeyId=44C040F42B9648A5BD88&disposition=0&alloworigin=1> [<https://perma.cc/DK6N-KZHK>] (noting high profit margins for most Big Tech firms); Shivaram Rajgopal, Anup Srivastava & Rong Zhao, Do Digital Technology Firms Earn Excess Profits?, Cal. Mgmt. Rev. (Nov. 19, 2020), <https://cmr.berkeley.edu/2020/11/do-digital-technology-firms-earn-excess-profits/> [<https://perma.cc/E8PY-WJY2>] (finding that digital firms achieve large profit margins as compared to other industries).

176. See *infra* notes 204–211 and accompanying text.

177. Birch et al., *Data as Asset?*, *supra* note 55, at 12–13.

178. These strategies include aggressive acquisition of would-be competitors, discussed in section II.B.

179. Zuboff, *Surveillance Capitalism*, *supra* note 30, at 97–98.

180. As noted at the beginning of Part II, data value is not the only factor in how companies produce and maintain market power; companies also deploy strategies that have little or nothing to do with social data value. This point is merely meant to distinguish the account of data value being used in direct or indirect money-making for Google from data value's role in Google's strategies to maintain market power.

the form of (possibly excessive) profits. Indeed, some argue that digital platform companies like Google, which have amassed sufficiently large datasets and monopolized access to them, have a “self-perpetuating and expanding” capacity for prediction value to beget and maintain market power.¹⁸¹ Rikap explains, “Data-driven intellectual monopolies base their innovations on processing big data with this artificial intelligence (AI) approach. Data-harvesting, centralization and analysis thus foster a cumulative advantage in terms of the ability to innovate.”¹⁸²

The trend of many dominant companies achieving extremely high market values despite running losses also suggests the general value of social data even absent its conversion into exchange value. For example, when Microsoft acquired LinkedIn in 2016, the firm was a loss company.¹⁸³ But it had a “network of 433 million professionals” and a massive amount of data on those users.¹⁸⁴ Microsoft paid \$26 billion to acquire the firm.¹⁸⁵ And Amazon’s market capitalization rapidly rose even in periods when it was reporting regular losses.¹⁸⁶ In other work, Birch and D.T. Cochrane identified the delayed expectation of future high profits typical of digital firms that have achieved market dominance as a new form of rentiership, which they describe as “expected monopoly rents.”¹⁸⁷ Relatedly, one may reasonably take investors to be making a bet that companies will eventually

181. Rikap, *Intellectual Monopoly Capitalism*, supra note 58, at 151. For others arguing that data strategies feature in tech platforms’ capacity to engage in enduring forms of rentiership stemming from intellectual monopolization, see Cédric Durand, *Techno-Féodalisme: Critique de L’économie Numérique* (2020); Cédric Durand, *Predation in the Age of Algorithms: The Role of Intangible Assets*, in *Accumulating Capital Today: Contemporary Strategies of Profit and Dispossessive Policies* 149, 149–62 (Marlène Benquet & Théo Bourgeron, eds., 2021); see also Malcolm Harris, *Are We Living Under ‘Technofeudalism’?*, *NY Mag.: Intelligencer* (Oct. 28, 2022), <https://nymag.com/intelligencer/2022/10/what-is-technofeudalism.html> [<https://perma.cc/3CWH-7H9U>] (discussing Durand’s technofeudalism thesis in English).

182. Rikap, *Intellectual Monopoly Capitalism*, supra note 58, at 151; see also Cecilia Rikap, *Amazon: A Story of Accumulation Through Intellectual Rentiership and Predation*, 26 *Competition & Change* 436, 437–38 (2022).

183. Kerry Flynn, *LinkedIn Earnings Are Just Fine Ahead of Microsoft Merger*, *Mashable* (Aug. 4, 2016), <https://mashable.com/article/linkedin-earnings-ahead-of-microsoft-merger> [<https://perma.cc/4HT7-UZVJ>] (reporting that LinkedIn posted losses of eighty-nine cents per share in 2016 and fifty-three cents per share in 2015).

184. Sarah McBride, *Microsoft to Buy LinkedIn for \$26.2 Billion in Its Largest Deal*, *Yahoo! Fin.* (June 14, 2016), <https://finance.yahoo.com/news/microsoft-buy-linkedin-26-2-001104954.html> [<https://perma.cc/M6BT-EJCJ>].

185. *Id.*

186. See Lina M. Khan, *Note, Amazon’s Antitrust Paradox*, 126 *Yale L.J.* 710, 748–49 (2017) [hereinafter Khan, *Amazon’s Antitrust Paradox*] (noting the trend of increasing company stock price in the face of losses and quoting an analyst as saying “Amazon’s stock price doesn’t seem to be correlated to its actual experience in any way” (internal quotation marks omitted) (quoting David Streitfeld, *Amazon Reports Unexpected Profit, and Stock Soars*, *N.Y. Times* (July 23, 2015), <http://www.nytimes.com/2015/07/24/technology/amazon-earnings-q2.html> (on file with the *Columbia Law Review*))).

187. Birch & Cochrane, supra note 56, at 50–51.

convert such prediction value into exchange value by following one of the strategies surveyed by the first and second scripts above.

Nevertheless, the fact that exchange value is realized at the investor level but not at the firm level is an atypical result that presents important challenges for legal regimes trying to govern such companies and their market behavior.¹⁸⁸ From the perspective of companies, it may incentivize the cultivation and accumulation of prediction value even in the absence of clear strategies for how to convert that value into profitmaking activity. Companies can frame their users and their users' data to investors as measurable assets in a process that scholars have described as "techcraft."¹⁸⁹ By presenting user metrics in a way that is legible to investors, companies transform social data and the prediction value it stores into an asset that investors take into account in the market valuation of a company.¹⁹⁰

A growing number of scholars link monopolized access to data not to any single profitmaking strategy but to generalized forms of market power. For example, Stiglitz argues that corporations are deriving excess power and wealth based on their ability to exploit data, leading to a host of negative implications for market competition and society more broadly.¹⁹¹ Rikap similarly argues that the concentration of knowledge in a handful of companies through their possession of large quantities of data has created "intellectual monopolies"; she has warned of the power implications of these monopolies for the future of innovation, knowledge production, and global development.¹⁹² Economist Jean Tirole identifies the "soft control" that private companies, governments, and other organizations can achieve through control over social data.¹⁹³

188. Exchange value is realized at the investor level but not the company level when prediction value translates into increased market capitalization of companies but does not translate into company profits. See Amanda Parsons, *The Shifting Economic Allegiance of Capital Gains*, 26 *Fla. Tax Rev.* (forthcoming 2024) (manuscript at 28–43), <https://ssrn.com/abstract=4152114> [<https://perma.cc/BZ4F-WCRG>] [hereinafter Parsons, *Shifting Economic Allegiance*] (explaining this phenomenon and its implications for tax law specifically).

189. See Birch et al., *Data as Asset?*, *supra* note 55, at 2.

190. See *supra* notes 54–59 and accompanying text (discussing the process of assetization of data).

191. See Stiglitz, *supra* note 143, at 120–37.

192. See Rikap, *Capitalism as Usual?*, *supra* note 58, at 159 ("Intellectual monopoly capitalism is therefore defined by a growing appropriation of society's knowledge, which enables the monopoly to exercise power over other firms and organizations."); Rikap, *Intellectual Monopoly Capitalism*, *supra* note 58, at 149 ("[M]ore than ever knowledge (cum innovation) is power and contemporary capitalism is driven by those monopolizing it.").

193. See Tirole, *supra* note 60, at 2011–12.

Legal scholars are also paying greater attention to the power of prediction value.¹⁹⁴ Julie Cohen constructs a forceful analysis of how technology, ideology, and the law have together produced power for informational capitalism's winners in her book *Between Truth and Power*.¹⁹⁵ In another work, Cohen argues that the power wielded by some platform firms has potentially tipped into a form of sovereignty.¹⁹⁶ Frank Pasquale highlights the ways in which tech firms have used “obfuscation and secrecy to consolidate power and wealth.”¹⁹⁷ Katharina Pistor notes the unexpected outcome of oligopolic power in the digital economy.¹⁹⁸ Under a Coasean framework, Pistor argues, the declining transaction costs in the digital economy should lead to the decline of the firm and a turn to

194. For further discussion in the legal academic literature of the relationship between data and power, see Kapczynski, *Informational Capitalism*, supra note 31, at 1515 (“Our legal order, intertwined with the architecture of digital networks, has enabled the creation of vast new firms that wield new forms of surveillance and algorithmic power, but it also has delivered us a form of neoliberal capitalism that is inclined toward monopoly, concentrated power, and inequality.”); Marian, supra note 31, at 550–51 (discussing the political power firms garner from the ability to engage in political microtargeting); Maurice E. Stucke, *Should We Be Concerned About Data-Opolies?*, 2 *Geo. L. Tech. Rev.* 275, 312–23 (2018) (outlining political power of “data-opolies” and its potential harms). See generally Lina M. Khan, *Sources of Tech Platform Power*, 2 *Geo. L. Tech. Rev.* 325 (2018) [hereinafter Khan, *Tech Platform Power*] (discussing forms of power held by platform businesses, including “information exploitation power,” and challenges for the legal system in addressing those powers).

The centrality of power is not just a matter of academic theorizing. It is embedded in the culture of informational capitalism's major corporate players. For example, in the early days of PayPal, the firm developed an app that tracked how many people opened new accounts, dinging each time a new account was opened. The firm called the app the “World Domination Index.” Max Chaffkin, *The Contrarian: Peter Thiel and Silicon Valley's Pursuit of Power 70* (2021). The concept of a “World Domination Index” is representative of the broader project of PayPal founder Peter Thiel—a project to shift the balance of power to these companies and their owners. And Thiel's project is not an idiosyncratic one—it has expanded to influence many in Silicon Valley. *Id.* at 14–17. The third script described in this section is in line with this project and worldview. See Will Davies, *The Road From Mont Pelerin to Silicon Valley*, *Pol. Econ. Rsch. Ctr.* (Oct. 17, 2022), https://www.perc.org.uk/project_posts/the-road-from-mont-pelerin-to-silicon-valley/ [<https://perma.cc/T4JA-Y3GU>] (discussing the desire for domination among Silicon Valley founders and the resulting belief that “there is the higher-order freedom of ‘founders’, which is far greater than the ordinary freedom to make choices, but really the freedom to construct whole social worlds”).

195. Cohen, *Between Truth and Power*, supra note 4.

196. Julie E. Cohen, *Law for the Platform Economy*, 51 *U.C. Davis L. Rev.* 133, 199 (2017) [hereinafter Cohen, *Platform Economy*] (“Dominant platforms’ role in the international legal order increasingly resembles that of sovereign states. And even as they evade the obligations of domestic legal regimes, platform firms are actively participating in the ongoing construction of new transnational institutions and relationships that are more hospitable to their interests.”).

197. Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* 14 (2015).

198. See Pistor, *Rule by Data*, supra note 31, at 101–04.

markets.¹⁹⁹ Yet instead, a small set of “Big Tech” firms have come to dominate the digital economy. Pistor links the power that Big Tech wields to their control of data and to data’s prediction value.²⁰⁰ She explains:

[I]n the world of big data controlled by Big Tech, data are not primarily objects of exchange transactions; rather, they are both the source for and the means of *control* by Big Tech and their clients over others: consumers of goods and services, workers, voters, members in organizations, or whatever other targets they might choose.²⁰¹

Pistor further explains that “[t]he worth of data does not lie in their exchange value but in the power they confer on data controllers”²⁰² and links this value to data’s capacity to create “asymmetries of power.”²⁰³

Social data power can also empower companies to achieve favorable political or regulatory goals.²⁰⁴ For example, consider Uber’s use of a program called Greyball.²⁰⁵ Beginning in 2014, Uber used Greyball (which was approved by Uber’s general counsel at the time) to skirt regulatory authorities by geofencing government buildings and “greyballing” users identified as (or suspected of being) law enforcement or city officials.²⁰⁶ Greyball allowed Uber to evade detection in cities like Boston, Paris, Portland, and in countries like Australia and China—all places Uber was formally restricted or banned. The aim was to evade detection long enough to rapidly grow its user base in these municipalities and gain a competitive edge against incumbent transportation providers—in violation of local laws.²⁰⁷ Once enough users began using Uber, the company could point to the popularity of the service as a *fait accompli* to regulators, making enforcement of existing regulations banning Uber unpopular and potentially politically costly. Indeed, in many instances,

199. See *id.* (outlining the Coasean theory and the patterns of firm growth in the digital economy that are incongruous with that theory).

200. See *id.* at 105 (“In fact, power seems a better explanation for the rise of Big Tech than the standard transaction cost arguments.”).

201. *Id.* at 104 (emphasis added).

202. *Id.* at 105.

203. *Id.* at 103. Economist Jean Tirole reached similar conclusions to Pistor. He identifies the “soft control” that private companies, governments, and other organizations can achieve through control over social data. See Tirole, *supra* note 60, at 2011–12.

204. For a more systematic account of how existing law has facilitated the cultivation of power among digital platforms as well as the way these companies are now attempting to use law to protect against countermovements to their rise in power, see Cohen, *Between Truth and Power*, *supra* note 4, at 11–13, 139–41 (summarizing the two parts of Cohen’s account); Cohen, *Platform Economy*, *supra* note 196, at 204.

205. Mike Isaac, *How Uber Deceives the Authorities Worldwide*, *N.Y. Times* (Mar. 3, 2017), <https://www.nytimes.com/2017/03/03/technology/uber-greyball-program-evade-authorities.html> (on file with the *Columbia Law Review*) [hereinafter Isaac, *How Uber Deceived*]; see also Mike Isaac, *Super Pumped: The Battle for Uber* 21–22, 270–71 (2019).

206. Isaac, *How Uber Deceived*, *supra* note 205.

207. *Id.*

Uber was able to successfully reach agreements to operate in cities after using Project Greyball to flout those cities' laws.²⁰⁸

Companies can use social data power to defy regulatory aims in more subtle ways. For instance, the example discussed above, drawn from research conducted by Rory Van Loo and Nikita Aggarwal, details how Amazon satisfies the technical requirement of offering low prices to consumers while using prediction value to hinder consumers' ability to easily search for and find the lowest priced goods.²⁰⁹ Again, Amazon's use of prediction value to guide which goods are shown to which customers—and for what price—describes a strategy covered under the second script: using prediction value to increase profits by imposing transaction costs on consumers with behavioral techniques that make these tactics difficult for consumers to detect or to discipline.²¹⁰ Yet in adopting this strategy, Amazon is also simultaneously using prediction value to achieve a regulatory objective: gaining the benefits of extracting excess profits while (arguably at least) hewing to the letter of antitrust law. Christopher Peterson and Marshall Steinbaum chronicle how gig companies similarly use prediction value to engage in fine-grained management and control of workplace conditions that escape regulatory scrutiny under existing consumer protection and antitrust law.²¹¹

There is a long history in legal scholarship and beyond of caring about market power precisely because of its close relation to political power. Within tax law, for example, the concept that controlling value resources confers power has colored debates around normative justifications for the consumption and wealth taxes as well as the corporate tax. Numerous tax scholars have argued that a person's consumption is the ideal tax base from both an efficiency and equity perspective.²¹² But a frequent critique of using consumption as a tax base is that it ignores the power the mere

208. *Id.* (detailing how two weeks after Uber began illegally dispatching drivers in Portland, Oregon, it reached an agreement with local officials to operate legally).

209. Van Loo & Aggarwal, *supra* note 146, at 11–23 (demonstrating how Amazon uses informational techniques, commonly regulated under consumer protection law, to manage its pricing strategies—allowing Amazon to maintain the perception of offering low prices while ensuring such low prices are difficult for consumers to access and find).

210. *Id.*; see also Dana Mattioli, Amazon Used Secret 'Project Nessie' Algorithm to Raise Prices, *Wall St. J.* (Oct. 3, 2023), <https://www.wsj.com/business/retail/amazon-used-secret-project-nessie-algorithm-to-raise-prices-6c593706> (on file with the *Columbia Law Review*).

211. Peterson & Steinbaum, *supra* note 123, at 637–58 (detailing how coercive practices enabled by platform informational techniques in the rideshare industry raise both consumer protection and anticompetitive concerns).

212. See, e.g., William D. Andrews, A Consumption-Type or Cash Flow Personal Income Tax, 87 *Harv. L. Rev.* 1113, 1165–77 (1974) (arguing for adoption of a consumption-style tax based on “considerations of fairness and efficiency,” as well as “feasib[ility]”); Richard L. Doernberg, A Workable Flat Rate Consumption Tax, 70 *Iowa L. Rev.* 425 (1985) (analyzing proposals for a consumption-style tax); Edward J. McCaffery, The Uneasy Case for Wealth Transfer Taxation, 104 *Yale L.J.* 283, 326 (1994) (“[I]f efficiency were the primary concern, then the consumption tax would seem to prevail under a technical analysis involving elasticities, general equilibrium, and so on.”).

possession of wealth brings with it, regardless of whether that wealth is actually consumed.²¹³ More recently, the power stemming from the mere possession or control of something that is valuable has been cited as a rationale for imposing a wealth tax on individuals.²¹⁴ Tempering the level of resources under the control of corporate management and the accompanying economic and political power that resource control brings has also been put forward as a justification for the corporate tax.²¹⁵ Legal scholars outside of tax law, notably scholars of market regulation, have likewise cited the link between concentrations of economic resources and power, and the potentially negative ramifications for American democracy.²¹⁶

* * *

The three scripts that firms follow to leverage prediction value—directly converting prediction value to exchange value, indirectly converting prediction value to exchange value, and converting prediction

213. See, e.g., Anne L. Alstott, *The Uneasy Liberal Case Against Income and Wealth Transfer Taxation: A Response to Professor McCaffery*, 51 *Tax L. Rev.* 363, 371 (1996) (“The unavoidable difficulty is that private wealth remains a source of current social, economic and political power that goes beyond the potential use of wealth for consumption.”); Barbara H. Fried, *Who Gets Utility From Bequests? The Distributive and Welfare Implications for a Consumption Tax*, 51 *Stan. L. Rev.* 641, 653 (1999) (highlighting the theory that people accumulate wealth “for the power or status that merely being wealthy brings them”); Edward D. Kleinbard, *Capital Taxation in an Age of Inequality*, 90 *S. Cal. L. Rev.* 593, 640 (2017) (arguing that power and prestige from wealth exist even absent the ability to consume that wealth in the future).

214. See, e.g., Jeremy Bearer-Friend, *The Great Democracy Initiative, Restoring Democracy Through Tax Policy* 10–11 (2018), https://rooseveltinstitute.org/wp-content/uploads/2021/08/GDI_Restoring-Democracy-Through-Tax-Policy_201812.pdf [<https://perma.cc/3JVG-CB9Y>] (arguing for federal wealth and transfer taxes to counter “concentrated political power of economic elites”); Ari Glogower, *Taxing Inequality*, 93 *N.Y.U. L. Rev.* 1421, 1445–50, 1467–68 (2018) (describing the relative economic power theory, which states that uneven concentrations of wealth and market power lead to uneven distributions of political and social power, and presenting a combined tax on income and wealth as a means to counter inequality).

215. See, e.g., Reuven S. Avi-Yonah, *Corporations, Society, and the State: A Defense of the Corporate Tax*, 90 *Va. L. Rev.* 1193, 1233–41 (2004) (outlining the mechanisms of corporate power and the role of political and economic power stemming from control over financial resources); see also Bearer-Friend, *supra* note 214, at 5 (“[T]he corporate tax also protects democratic values by serving as a regulatory device to counteract the unbridled powers of large businesses.”).

216. See Ganesh Sitaraman, *The Crisis of the Middle-Class Constitution: Why Economic Inequality Threatens Our Republic* 224 (2017) (“As wealth is concentrated in the hands of elites and corporations, they use their wealth and influence to rewrite laws and regulations in ways that help them amass even greater wealth and power.”); Jedediah Britton-Purdy, David Singh Grewal, Amy Kapczynski & K. Sabeel Rahman, *Building a Law-and-Political-Economy Framework: Beyond the Twentieth-Century Synthesis*, 129 *Yale L.J.* 1784, 1788–89 (2020) (“Government enacts the policy preferences of the rich over those of the majority . . . Citizen frustration with this intertwined and increasing concentration of economic and political power is visible on the right . . . and on the left . . .”).

value into economic and political power—have precipitated certain business models and strategies that have become prevalent in the digital economy. This section has explored many of those business models and practices. The following section describes and analyzes some other prevalent business models and practices, focusing more specifically on practices targeted at the accumulation of social data.

B. *The Business Models and Practices of the Digital Economy*

Many of the business models and practices that typify the digital economy focus on growth and expansion.²¹⁷ While in many instances these growth-focused business practices further a company's pursuit of the different scripts discussed above, a central role of growth-focused business practices is to allow companies to accumulate social data. This accumulation could take the form of building up user and customer bases, thus securing streams of social data from them. Or it could take the form of directly acquiring social data from other companies. These growth and expansion strategies typically eschew profits (at least in the short or medium term) in favor of building up social data. The firms can then leverage the prediction value stemming from this data to achieve greater profits and power in the future. This section analyzes three business models and practices frequently pursued by firms in the digital economy that allow them to accumulate social data. The first is offering free or low-cost services to build up user and customer bases. The second is creating ecosystems of products and services that capture users and customers. The third is pursuing aggressive merger and acquisition strategies.

It is important to note at the outset that the business models discussed in this subpart most obviously connect with the pursuit of script three. This is because script three involves storing prediction value as a form of market power rather than converting prediction value into exchange value. The business models discussed in this subpart center around accumulating social data as well as establishing strong networks and market dominance to accumulate more social data in the future. But as will be highlighted throughout this section, some of these growth-and-expansion-focused strategies also further the other scripts.

1. *The Business of "Free"*. — Profits are not the central motivator for emerging firms in the digital economy. These firms instead prioritize building up user and customer bases with the aim of achieving market dominance.²¹⁸ While market dominance brings with it many advantages

217. The aim of this section is not to provide an exhaustive account of the business models and strategies seen within informational capitalism. Instead, the section highlights some key strategies that are both particularly prominent within informational capitalism and significant for the legal regimes tasked with regulating these businesses.

218. See Collin & Colin, *supra* note 175, at 28–29 (explaining how firms in the digital economy prioritize gaining new users and achieving “traction”); Vijay Govindarajan, Shivaram Rajgopal & Anup Srivastava, *Why Financial Statements Don't Work for Digital Companies*, *Harv. Bus. Rev.* (Feb. 26, 2018), <https://hbr.org/2018/02/why-financial->

for companies, the opportunity to accumulate social data is one reason for this advantage. As discussed above,²¹⁹ the predictive power of social data relies on the ability to collect and analyze broad swaths of social data. Because “big” data requires systematic monitoring of large numbers of people, digital firms need to accrue large user and customer bases before they can fully realize the predictive capacity of social data. Building up a collection of data subjects is a necessary first step for firms to compete in an informational capitalist economy.

Another reason that firms prioritize building up user and customer bases over profits has to do with the importance of platform business models within the digital economy. Platforms are important to the accumulation of social data because their technologies structurally facilitate tracking of users and data collection.²²⁰ Platform businesses use technology to connect users in a wide variety of value-creating interactions.²²¹ And platforms are also heavily reliant on network effects for the success of their businesses.²²² When a new platform is launched, there is little reason for a new user to join the platform because there is not an existing network of users with whom to interact. For example, one would not want to join a social network platform without a robust network of users, such as family, friends, and celebrities, with whom to interact. After a critical mass of users is reached, however, positive network effects begin to take over, which can lead to rapid growth of the platform and market dominance.²²³ At the point that this critical mass is achieved, digital firms can, in theory, exploit their dominant market positions and begin to reap monetary profits.²²⁴ As political economist Jathan Sadowski explains,

statements-dont-work-for-digital-companies [<https://perma.cc/QAR5-ADZK>] (describing “achieving market leadership” as “the most important aim” for digital firms).

219. See *supra* notes 43–58 and accompanying text.

220. See Srnicek, *supra* note 109, at 42–43 (explaining that platforms emerged out of a need to “monopolise, extract, analyse, and use” data); Cohen, *Platform Economy*, *supra* note 196, at 140–43 (describing the role of platforms in “the datafication of everyday life”); Khan, *Tech Platform Power*, *supra* note 194, at 329–30 (discussing platforms’ ability to gather extensive amounts of data from their users’ activities).

221. See Geoffrey G. Parker, Marshall W. Van Alstyne & Sangeet Paul Choudary, *Platform Revolution* 108–11 (2016) (describing four forms of excess value that the platform generates for consumers, producers, and third-party providers).

222. See Carl Shapiro & Hal R. Varian, *Information Rules: A Strategic Guide to the Network Economy* 13–14 (1999) (discussing the role of positive feedback and network effects in the information economy); Stucke, *supra* note 194, at 281–83 (describing the role of network effects for Big Tech companies); Michael L. Katz & Carl Shapiro, *Systems Competition and Network Effects*, *J. Econ. Persps.*, Spring 1994, at 93, 93–95 (analyzing the role of network effects in consumer decisionmaking in communications and technology systems).

223. See Jeffrey H. Rohlfs, *Bandwagon Effects in High-Technology Industries* 27–28 (2013) (describing the positive feedback cycle that occurs once a network has achieved a critical mass); see also Tim Wu, *The Master Switch: The Rise and Fall of Information Empires* 49 (2010) (“In today’s automated world, the larger the network, the better it is, because you can reach more people in more places.”).

224. See *supra* notes 171–181 and accompanying text.

“The practice of acquiring data first—indeed, of designing things for the primary purpose of data extraction—and then (hopefully) figuring out how to valorise it later is now normal for organisations following the platform model.”²²⁵

As a result, growing a network of users and customers is essential for digital firms, both to lock in access to flows of social data and to achieve the network effects necessary to reach market dominance. To achieve this growth, digital firms eschew profits, often for very long periods of time. One way they do this is by offering free services.²²⁶ There is no fee to run a Google search, to post a photo to Instagram, or to stream music on Spotify. Firms offering free services to one group of customers will often still earn revenues in other ways, such as by selling targeted advertising²²⁷ or charging a fee for premium versions of their services (known as the “freemium” business model).²²⁸ Or they might engage in predatory pricing strategies, like those discussed above, to charge below-market prices to some customers and higher prices to others.²²⁹ Despite the existence of these strategies, profit maximization is not the central goal for these digital firms—growth and its accompanying data collection are key.²³⁰ For example, it was not until after its 2012 initial public offering (IPO) that Facebook began to expand its advertising sales.²³¹ The firm already had 800 million users at the time of the IPO.²³² This strategy of favoring growth over income can also be seen in acquisitions of digital firms when companies have sold for high market values despite not

225. Jathan Sadowski, *The Internet of Landlords: Digital Platforms and New Mechanisms of Rentier Capitalism*, 52 *Antipode* 562, 572 (2020) [hereinafter Sadowski, *Internet of Landlords*] (citing Marion Fourcade & Kieran Healy, *Seeing Like a Market*, 15 *Socio-Econ. Rev.* 9 (2016)).

226. See Zuboff, *Surveillance Capitalism*, supra note 30, at 52–53 (describing the provision of free services by digital firms and the framing of data surveillance as a quid pro quo for those services); Stucke, supra note 194, at 279 (“Most of Google’s and Facebook’s services for consumers are ostensibly ‘free.’”); see also Chris Anderson, *Free: The Future of a Radical Price* 20–33 (2009) [hereinafter Anderson, *Free*] (“[A]ll forms of Free boil down to variations of the same thing: shifting money around from product to product, person to person, between now and later, or into nonmonetary markets and back out again.”).

227. See supra notes 104–111 and accompanying text.

228. Anderson, *Free*, supra note 226, at 26–27 (describing the freemium business model).

229. See Leslie, supra note 147, at 75.

230. See Srnicek, supra note 109, at 97 (“Unlike in manufacturing, in platforms competitiveness is not judged solely by the criterion of a maximal difference between costs and prices; data collection and analysis also contribute to how competitiveness is judged and ranked.”).

231. Collin & Colin, supra note 175, at 28; Rebecca Greenfield, *2012: The Year Facebook Finally Tried to Make Some Money*, *The Atlantic* (Dec. 14, 2012), <https://www.theatlantic.com/technology/archive/2012/12/2012-year-facebook-finally-tried-make-some-money/320493/> (on file with the *Columbia Law Review*).

232. Facebook, Inc., *Registration Statement (Form S-1) 1* (Feb. 1, 2012), https://www.sec.gov/Archives/edgar/data/1326801/000119312512034517/d287954ds1.htm#toc287954_3a [https://perma.cc/X4ES-AWMG].

earning income.²³³ While providing free services can lower digital firms' bottom lines, it allows them to secure large networks and streams of social data from those network participants.

Other digital firms might not offer entirely free services to potential users and customers but will offer low-cost services designed to build up their network and social data access, often taking substantial losses in the process. Amazon's business strategy is a leading example of this. Amazon launched Amazon Prime in 2005.²³⁴ At an initial annual cost of seventy-nine dollars, the program offered free two-day shipping for customers, and other features have been added onto the program over the years, such as streaming video services.²³⁵ Lina Khan has written that "[t]he program has arguably been the retailer's single biggest driver of growth," but "[a]s with its other ventures, Amazon lost money on Prime to gain buy-in."²³⁶ Khan cites an analyst who estimates that Amazon loses between \$1 to \$2 billion per year through the Prime program.²³⁷ This is part of a broader historical trend of Amazon eschewing profits in favor of growth for much of its corporate life.²³⁸ It is only in recent years, after establishing market dominance, that the firm has begun to report substantial profits.²³⁹ Even then, the majority of those profits are derived from its cloud computing business line, with its other operations seeing lower margins or even losses.²⁴⁰

This common practice in the digital economy of businesses choosing to not maximize their profits by offering free or low-cost services for extended periods of time defies the expectations of firm behavior. Various legal regimes assume that companies will aim to maximize profits for their shareholders after a reasonable start-up period. But with the growing importance of social data as a factor of production, this assumption that is at the center of legal scaffolding no longer holds true.

2. *Building Ecosystems.* — The building of ecosystems of products and services also typifies firm behavior within the digital economy.²⁴¹ Google is

233. See *supra* notes 183–187 and accompanying text.

234. Khan, Amazon's Antitrust Paradox, *supra* note 186, at 750.

235. *Id.*

236. *Id.* at 750–51.

237. *Id.* at 751.

238. See *id.* at 747–49 (describing Amazon's history of either losses or very low profit margins).

239. See Amazon.com, Inc., Annual Report (Form 10-K) 38 (Feb. 3, 2022), <https://www.sec.gov/Archives/edgar/data/1018724/000101872422000005/amzn-20211231.htm> [<https://perma.cc/PTQ2-9DCM>] (reporting a net income of \$12 billion in 2019, \$21 billion in 2020, and \$33 billion in 2021).

240. See *id.* at 24 (showing that most of Amazon's income in 2020 and 2021 was derived from Amazon Web Services).

241. See Srnicek, *supra* note 109, at 95–96 (describing ecosystem development by platform businesses); Birch & Cochrane, *supra* note 56, at 49–50 (describing ecosystem-building strategies by Big Tech companies and the ways in which those strategies create "enclave rents").

not just a search engine. The firm has created an ecosystem that includes an email service (Gmail), an online word processor (Google Docs), a web browser (Chrome), a phone and accompanying mobile operating system (Android), among many, many other products and services.²⁴² Apple's ecosystem ranges from hardware products, like the iPhone and Apple Watches, to services, like iCloud storage and iMessage, to apps available via the App Store.²⁴³

Ecosystem building is a mechanism for growth, allowing digital firms to build up user and customer bases and collect and amass social data, including its accompanying prediction value. Firms can create ecosystems in ways that lock in users to that ecosystem, such as apps that are only compatible with the firm's operating system.²⁴⁴ This lock-in guarantees flows of social data.²⁴⁵ Additionally, if a firm has access to more areas of a person's life, it can accumulate a greater variety of social data.²⁴⁶ For example, if Amazon has a map of your home, it can better anticipate what products you might be inclined to purchase for your home.²⁴⁷ In fact, creating an ecosystem spanning broad arrays of business lines is reportedly part of founder Jeff Bezos's vision to build the firm into a "utility" that would become essential to commerce."²⁴⁸ Locking users into an ecosystem, closing out competitors, self-preferencing their own products, and controlling access to the data collected from the ecosystem all contribute

242. Browse All of Google's Products and Services, About Google, https://about.google/intl/en_us/products/ [<https://perma.cc/V2KY-FYEA>] (last visited Feb. 20, 2024).

243. See, e.g., Thomas Ricker, First Click: Apple's Greatest Innovation Is Its Ecosystem, *The Verge* (Sept. 7, 2016), <https://www.theverge.com/2016/9/7/12828846/apple-s-greatest-product-is-its-ecosystem> [<https://perma.cc/M5BT-XJT2>] (describing various products and services in the Apple ecosystem); Ian Sherr, Apple's 'Walled Garden' Walls Will Get Even Higher With iOS 14, iPadOS 14 and MacOS Big Sur, *CNET* (July 2, 2020), <https://www.cnet.com/tech/mobile/apple-walled-garden-walls-will-get-even-higher-with-ios-14-ipados-14-macos-big-sur/> [<https://perma.cc/7UME-BZVX>] (same).

244. See Birch & Cochrane, *supra* note 56, at 49 (describing the approach of Big Tech companies "locking in users to their ecosystems, both legally (e.g. contractual agreements) and technically (e.g. interoperability restrictions)").

245. See Complaint at 27, *United States v. Apple Inc.*, No. 02:24-cv-04055 (D.N.J. filed Mar. 21, 2024), 2024 WL 1219405 (alleging that Apple sells keyword search data for one app to parties other than the app's owners as another way to increase revenue); Srnicek, *supra* note 109, at 96 (describing how ecosystem building creates monopolies of data access for digital firms).

246. See Srnicek, *supra* note 109, at 95 ("[A]ccess to a multitude of data from different areas of our life makes prediction more useful, and this stimulates centralisation of data within one platform.").

247. See Ron Knox, Amazon's Dangerous New Acquisition, *The Atlantic* (Aug. 21, 2022), <https://www.theatlantic.com/ideas/archive/2022/08/amazon-roomba-irobot-acquisition-monopoly/671145/> (on file with the *Columbia Law Review*) (discussing Amazon's purchase of the vacuum manufacturer iRobot and the implications of the data that Amazon could gather from smart vacuums for its retail business).

248. Khan, *Amazon's Antitrust Paradox*, *supra* note 186, at 754–55 (quoting Amazon employees).

to the firm market power consolidation that is the aim of the third script.²⁴⁹ By providing access to both a greater quantity and greater variety of social data, ecosystem building allows firms to cultivate greater prediction value and market power, fueling a positive feedback cycle of future growth.

In addition to generally facilitating accumulation of social data, ecosystem building can be useful for firms pursuing the second script—indirectly converting prediction value into exchange value. Ecosystem building increases and stabilizes revenues by locking customers into the system.²⁵⁰ And ecosystem building is particularly helpful for companies attempting to develop new products and expand into new business lines and industries. An expansive ecosystem of products and services provides firms with the opportunities to use prediction value accrued from one part of their ecosystem and monetize it through a product or service in another part of their ecosystem. As one tech entrepreneur explained: “At large companies, sometimes we launch products not for the revenue, but for the data. We actually do that quite often . . . and we monetize the data through a different product.”²⁵¹ This ability to monetize social data accumulated in one part of the company’s ecosystem in another part of the ecosystem provides ample opportunity for legal arbitrage. Companies might face strict regulations around the use of social data for a particular product or service. But if they have an expansive ecosystem of products, they could monetize this social data through another product or service.

3. *Aggressive Acquisitions.* — The digital economy has brought with it an uptick in acquisitions. This can be seen particularly in the context of digital firms. Big Tech cash expenditures on acquisitions averaged \$23 billion in the period between 2010 and 2019—approximately three times the average for the top 200 global firms.²⁵² As of April 2021, since their respective foundings, Apple had acquired 123 companies, Amazon had acquired 111, Facebook had acquired 105, and Google had acquired 268.²⁵³

249. See Zuboff, *Surveillance Capitalism*, supra note 30, at 179 (discussing the “unprecedented concentrations of knowledge and power” that companies like Google have achieved through ecosystem building); Birch et al., *Data as Asset?*, supra note 55, at 2 (describing the “societal dominance” achieved by Big Tech firms through, among other factors, ecosystem governance and control over access to social data); U.N. Conf. on Trade & Dev., *Trade and Development Report 2018: Power, Platforms, and the Free Trade Delusion*, at VI–VII, U.N. Doc. UNCTAD/TDF/2018, Sales No. E.18.II.D.7 (2018).

250. See Complaint, supra note 245, at 44 (alleging that Apple has limited the cross-platform development of digital wallets to increase switching costs for leaving Apple’s ecosystem).

251. Sadowski, *Internet of Landlords*, supra note 225, at 572 (internal quotation marks omitted) (quoting Stanford Graduate School of Business, Andrew Ng: Artificial Intelligence Is the New Electricity, YouTube, at 33:28 (Feb. 2, 2017), <https://www.youtube.com/watch?v=21EiKfQYZXc> (on file with the *Columbia Law Review*)).

252. Birch et al., *Data as Asset?*, supra note 55, at 11.

253. Chris Alcantara, Kevin Schaul, Gerrit De Vynck & Reed Albergotti, *How Big Tech Got So Big: Hundreds of Acquisitions*, Wash. Post (Apr. 21, 2021),

This business practice of aggressive acquisitions is part of the overall focus on growth and expansion within the digital economy. Commentators highlight that many of these acquisitions are largely driven by the desire to acquire data from target companies through so-called data-driven mergers.²⁵⁴ Facebook's acquisition of WhatsApp has been cited as one example.²⁵⁵ Google's acquisition of Waze is another.²⁵⁶

Of course, companies have other motivations for these acquisitions separate and apart from acquiring social data. These motivations might include foreclosing future competition and consolidating market dominance. But accumulating social data remains an important factor. And these other motivations may interrelate with accumulating social data. Stamping out competition is a means for companies to achieve and maintain their dominant market positions.²⁵⁷ Establishing dominant market positions provides companies with access to user bases and future streams of data from those users. Acquiring other firms to gain access to their users is part of the growth strategy of building dominant market positions.²⁵⁸ Microsoft's 2016 acquisition of LinkedIn is another example of the acquisition of a digital firm to gain access to a user base and their data. As discussed above,²⁵⁹ LinkedIn was posting losses when Microsoft paid \$26 billion to acquire it. But, as one analyst described, "[the acquisition was] a massive growth play for Microsoft."²⁶⁰ Microsoft highlighted the large customer base that the deal brought, as well as the potential for user data to improve its analytics and AI capacity.²⁶¹ As part of the digital economy's overall focus on growth, aggressive acquisitions are an important business strategy for firms following all three of the digital economy's scripts.

Acquisitions are particularly useful to firms pursuing the second script, specifically those companies that aim to use social data to create

<https://www.washingtonpost.com/technology/interactive/2021/amazon-apple-facebook-google-acquisitions/> (on file with the *Columbia Law Review*) (last updated Sept. 26, 2023).

254. See Maurice E. Stucke & Allen P. Grunes, *Big Data and Competition Policy* 1–3 (2016) (exploring the phenomenon of data-driven mergers and the failure of global competition policy to adequately respond to the trend).

255. *Id.* at 124.

256. *Id.* at 135.

257. See Majority Staff of H. Subcomm. on Antitrust, Com. & Admin. L., H. Comm. on the Judiciary, 117th Cong., *Investigation of Competition in Digital Markets* 6 (Comm. Print 2020) (detailing the trend of digital firms acquiring companies "to neutralize a competitive threat or to maintain and expand the firm's dominance"); C. Scott Hemphill & Tim Wu, *Nascent Competitors*, 168 U. Pa. L. Rev. 1879, 1880 (2020) (introducing the concept of nascent competition and its antitrust implications).

258. See Birch et al., *Data as Asset?*, *supra* note 55, at 11 ("[T]here are variations among Big Tech firms when it comes to acquisitions, although the core commonality of their business model is that they seek to strengthen their monopoly of users, user engagement, and access to users.").

259. See *supra* note 183 and accompanying text.

260. McBride, *supra* note 184 (quoting analyst Ted Schadler).

261. *Id.*

new business lines or expand into other industries. The majority of Big Tech's acquisitions have been acquisitions that expanded the firms outside of their original business lines and into new sectors. Seventy-eight percent of Apple's acquisitions, sixty-four percent of Amazon's acquisitions, seventy percent of Google's acquisitions, and seventy-three percent of Facebook's acquisitions have been of companies outside their original business lines.²⁶² Through these acquisitions, firms are able to take the prediction value that they have built up through collecting social data in one context and apply it in another context.

Google's acquisition of Fitbit is an example of an acquisition that allowed the company to expand into a new industry (as well as gain access to streams of social data and expand its ecosystem of products). In 2019, Google announced its intent to acquire Fitbit, a company that produces wearable fitness technology and had approximately 30 million active users and data on users' fitness and health spanning back a decade.²⁶³ At the time, Google was attempting to pivot into the healthcare industry.²⁶⁴ The merger sparked concerns from antitrust authorities across the globe about the implications of Google possessing that level of social data.²⁶⁵

Intuit's recent acquisition of Mailchimp is another useful example of this strategy. Intuit, a financial software firm, acquired Mailchimp, an email marketing platform, in 2021 for \$12 billion.²⁶⁶ Intuit's existing products included Credit Karma, Mint, and TurboTax, which provided the

262. Alcantara et al., *supra* note 253 (breaking down all known acquisitions by Amazon, Apple, Facebook, and Google through April 2021 by original or new business line); see also Khan, *Amazon's Antitrust Paradox*, *supra* note 186, at 754 ("Another key element of Amazon's strategy—and one partly enabled by its capacity to thrive despite posting losses—has been to expand aggressively into multiple business lines. . . . For the most part, Amazon has expanded into these areas by acquiring existing firms.")

263. Lucas Griebeler Da Motta, *Why We Should Be Careful About Google's Promises in the Fitbit Deal*, ProMarket (Aug. 21, 2020), <https://www.promarket.org/2020/08/21/why-we-should-be-careful-about-googles-promises-in-the-fitbit-deal/> [<https://perma.cc/UH4H-AX59>].

264. See *supra* notes 167–168 and accompanying text (discussing Verily).

265. See, e.g., Austl. Competition & Consumer Comm'n, *Statement of Issues: Google LLC—Proposed Acquisition of Fitbit Inc 2* (2020), <https://www.accc.gov.au/system/files/public-registers/documents/Google%20Fitbit%20-%20Statement%20of%20Issues%20-%2018%20June%202020.pdf> [<https://perma.cc/V2C4-FQ6W>] ("The accumulation of additional, individual user data via this transaction in an entity which already benefits from substantial market power in multiple markets may contribute to reduced competitive outcomes in the future."); European Commission Press Release IP/20/1446, *Mergers: Commission Opens In-Depth Investigation Into the Proposed Acquisition of Fitbit by Google* (Aug. 4, 2020), https://ec.europa.eu/commission/presscorner/detail/nl/ip_20_1446 [<https://perma.cc/VH5C-JMZY>] ("The data collected via wrist-worn wearable devices appears, at this stage of the Commission's review of the transaction, to be an important advantage in the online advertising markets.")

266. Press Release, Intuit Mailchimp, *Intuit Completes Acquisition of Mailchimp* (Nov. 1, 2021), <https://mailchimp.com/newsroom/intuit-completes-mailchimp-acquisition/> [<https://perma.cc/6WL5-Q82H>].

firm with data about individuals' personal finances and spending habits.²⁶⁷ Quickbooks was another existing Intuit product, which provided the firm with customer sales data from small and mid-sized businesses.²⁶⁸ Social data from Intuit's existing products on personal finance, spending habits, and customer sales could be used to predict and influence consumer behavior, and the firm can now use these insights to design more effective targeting of messages in an entirely new industry—email marketing.²⁶⁹ As Intuit explained in an investor presentation on the acquisition, "Customer data and purchase data brought together creates actionable insights and opportunities for small business and mid-market growth."²⁷⁰ Acquisitions are a means through which digital firms can convert the prediction value that they accrue through one business activity into exchange value in an entirely new industry.

Informational capitalism has brought about seismic changes to the economy. As social data has emerged as a new mode of production, prediction value has emerged as a new value form. Firms have responded to these changes by following three basic scripts: (1) directly converting prediction value to exchange value through methods such as targeted advertising, (2) indirectly converting prediction value to exchange value by using prediction value to improve upon or develop products and services, and (3) using prediction value as a means to establish power. As companies pursue these scripts, several business practices and methods have become commonplace in the modern economy. As the next Part will discuss, these scripts, and the business practices that have emerged alongside them, run counter to many of the assumptions at the heart of a variety of legal regimes. As a result, various areas of the law are struggling to effectively govern the digital economy.

267. Intuit: If Successfully Integrated, Credit Karma and Mailchimp Are Game Changers, Seeking Alpha (June 26, 2022), <https://seekingalpha.com/article/4520384-intuit-successfully-integrated-credit-karma-mailchimp-game-changers> (on file with the *Columbia Law Review*) [hereinafter Intuit: If Successfully Integrated] (identifying the types of data to which Intuit's products provide the firm access).

268. Gene Marks, On CRM: How Intuit's Purchase of Mailchimp Will Kill Your Monthly Newsletter, *Forbes* (Sept. 22, 2021), <https://www.forbes.com/sites/quickerbetteertech/2021/09/22/on-crm-how-intuits-purchase-of-mailchimp-will-kill-your-monthly-newsletter/> [<https://perma.cc/88A8-PNRT>].

269. See Intuit: If Successfully Integrated, *supra* note 267 (noting the potential for Intuit to use its existing data to improve the email marketing service provided by Mailchimp); Marks, *supra* note 268 (same).

270. Intuit, Inc., Investor Presentation: Intuit's Acquisition of Mailchimp 15 (2021), https://s23.q4cdn.com/935127502/files/doc_presentations/2021/Intuit%27s-Acquisition-of-Mailchimp-Presentation.pdf [<https://perma.cc/2R8T-UHXL>].

III. LEGAL COLLISIONS

A. *Two Camps of Legal Collisions*

The challenges of grappling with social data and prediction value creates issues across several legal regimes. This section focuses on two: tax law and privacy and data protection law. These fields represent two “camps” of legal failings in the face of informational capitalism.

This first camp consists of fields of law that have historically been tasked with governing and regulating value creation. These fields are struggling to integrate value creation from social data into their existing regulatory regimes. This Article argues that these struggles stem from the failure to recognize prediction value as a distinct and separate value form that does not readily translate into exchange value. In addition to tax law, other legal fields included in this camp include antitrust law and financial regulation.

The second camp consists of fields of law that have not historically viewed themselves as having a role in governing and regulating value creation. This Article argues that the advent of social data as a factor of production and prediction value as a key and distinct mode of value creation has made regulating value creation an imperative for these fields. But these fields are still grappling with their new role as primary governors of value creation under informational capitalism. As a result, while recent shifts in scholarly trends promise otherwise, these fields have not yet developed a positive agenda for regulating value creation. Recognizing prediction value as a form of value creation separate from exchange value can help inform this positive regulatory agenda. In addition to privacy and data governance law, other legal fields included in this camp include First Amendment law.

B. *Taxing Prediction Value*

Tax law is in the business of governing value creation. This business of governing value creation is in pursuit of three basic goals: to raise government revenues, to redistribute income and wealth, and to regulate private sector behavior.²⁷¹ In pursuit of these goals, the tax system strives to allocate burdens across taxpayers in a way that is equitable, efficient, and administrable.²⁷² International tax law is further tasked with ensuring

271. Reuven S. Avi-Yonah, *The Three Goals of Taxation*, 60 *Tax L. Rev.* 1, 3 (2006) (“What are taxes for? . . . [T]axes are needed to raise revenue for necessary governmental functions Taxation can have a redistributive function Taxation also has a regulatory component: It can be used to steer private sector activity in the directions desired by governments.”).

272. See *id.* at 26 (identifying equity, efficiency, and administrability as “the three traditional policy grounds” of tax law); Allison Christians, *Introduction to Tax Policy Theory* 10–11 (May 29, 2018), <https://ssrn.com/abstract=3186791> [<https://perma.cc/C7P9-XP KU>] (unpublished manuscript) (“[M]ost tax scholarship argues that to achieve the desired

that the choice of which country is allowed to tax cross-border income is made in an equitable, efficient, and administrable manner.²⁷³

The digital economy and the accompanying rise of prediction value as a key value form is colliding with tax law in two distinct ways. The first is a conceptual collision. Tax law scholars and policymakers are not recognizing and understanding prediction value as a new value form distinct from exchange value and, as a result, are inappropriately attempting to address tax law's failures in the face of the digital economy through an exchange value lens. The second collision relates not to prediction value itself but to the digital economy's scripts and resulting business practices. The existing tax system produces results incongruent with the underlying goals of tax law when applied to these new and unfamiliar scripts and business practices.

1. *A Conceptual Collision.* — Modern tax law is grounded in exchange value. At the end of the day, tens of thousands of pages of code and regulations, countless judicial and administrative decisions, and thousands of bilateral treaties boil down to numbers on a tax form, and these numbers represent the monetary value of income or, in the case of estate and gift taxation, wealth.²⁷⁴ The Internal Revenue Code does not contain a standard definition of “value.”²⁷⁵ But the hundreds of references to “value” in the code predominantly refer to fair market value—the price, or exchange value, that an asset would demand in an open market transaction.²⁷⁶

Tax law's conceptual equation of “value” with exchange value is stymying efforts to adapt tax law to the digital economy. There is widespread agreement that tax law, particularly international tax law, is

distribution of costs and benefits through taxation, societies should be guided by three principles: equity (or fairness), economic efficiency, and administrative capacity.” (emphasis omitted)).

273. See Michael J. Graetz, *Taxing International Income: Inadequate Principles, Outdated Concepts, and Unsatisfactory Policies*, 26 *Brook. J. Int'l L.* 1357, 1392, 1406, 1410 (2001) (identifying equity, efficiency, and simplicity as considerations in designing international tax policy).

274. See Allison Christians & Laurens van Apeldoorn, *Taxing Income Where Value Is Created*, 22 *Fla. Tax Rev.* 1, 10–11 (2018) (“It is on the basis of the idea of market value or fair market value . . . that income tax systems assign income to parties in all kinds of transactions involving exchanges of tangible and intangible goods and services.”); Marian, *supra* note 31, at 514 (“[I]ncome taxation is imposed on the aggregation of taxpayers’ ‘savings’ and ‘consumption.’ Both concepts rely on our ability to identify . . . the monetary value of income.” (footnote omitted)); *Understanding Federal Estate and Gift Taxes*, CBO (June 2021), <https://www.cbo.gov/publication/57272> [<https://perma.cc/9P6H-3Q7X>] (“The value of the estate’s assets is usually determined as the fair-market value on the owner’s date of death . . .”).

275. Shu-Yi Oei, *United States*, in *Taxation and Value Creation* 669, 669 (Werner Haslehner & Marie Lamensch eds., 2021) (“There is no singular coherent definition of value in the Code.”).

276. *Id.*

failing in the modern economy.²⁷⁷ Assertions by politicians and governments that multinational corporations, particularly Big Tech companies, are not paying their “fair share” of taxes are common.²⁷⁸ This concern over companies paying their fair share is a matter not only of the total amount of tax paid but also to which countries those taxes are paid. The need to align the place of taxation with the place of “value creation” has been a frequent refrain among politicians and policymakers.²⁷⁹

This political push to use the concept of value creation to determine which country gets to tax companies has been broadly criticized by academics.²⁸⁰ Academics have described the concept of value creation as

277. See, e.g., Collin & Colin, *supra* note 175, at 2 (“The failure of tax law to keep pace with economic transformation is especially obvious in the case of the digital economy.”); Lilian V. Faulhaber, *Taxing Tech: The Future of Digital Taxation*, 39 *Va. Tax Rev.* 145, 149 (2019) (arguing that recent reform proposals indicate that “countries around the world believe that the international tax system is out of step with the current economy”); Mitchell Kane, *A Defense of Source Rules in International Taxation*, 32 *Yale J. on Regul.* 311, 312 (2015) (“The body of law generally labeled ‘international taxation’ is widely perceived to be in shambles.”).

278. See, e.g., The Associated Press, *The G-7 Nations Agree to Make Big Tech Companies Pay Their Fair Share of Taxes*, NPR (June 5, 2021), <https://www.npr.org/2021/06/05/1003563505/the-g-7-nations-have-agreed-to-make-big-tech-companies-pay-their-fair-share-of-t> [<https://perma.cc/J44H-9K5X>] (quoting Rishi Sunak, then the Chancellor of the Exchequer, as characterizing the G-7 agreement on international tax reform as “requiring the largest multinational tech giants to pay their fair share of tax in the UK”); Richard Lough, *Explainer: Macron’s Quest for an International Tax on Digital Services*, Reuters (Aug. 22, 2019), <https://www.reuters.com/article/us-g7-summit-digital-tax-explainer/explainer-macrons-quest-for-an-international-tax-on-digital-services-idUSKCN1VCOVH> (on file with the *Columbia Law Review*) (describing frustration among political leaders regarding their inability to tax tech companies on profits they believe to be derived from business activities in their countries).

279. See, e.g., Communication from the Commission to the European Parliament and the Council: Time to Establish a Modern, Fair and Efficient Taxation Standard for the Digital Economy, at 4, COM (2018) 146 final (Mar. 21, 2018), https://eur-lex.europa.eu/resource.html?uri=cellar:2bafa0d9-2dde-11e8-b5fe-01aa75ed71a1.0017.02/DOC_1&format=PDF [<https://perma.cc/3LZW-QWKT>] [hereinafter European Commission, *Modern Taxation*] (explaining that a disconnect has emerged in the digital economy “between where the value is created, and where taxes are paid” and proposing reforms to correct this disconnect); see also Werner Haslehner & Marie Lamensch, *General Report on Value Creation and Taxation: Outlining the Debate*, in *Taxation and Value Creation*, *supra* note 275, at 3, 35 (“There can be no doubt that ‘taxing income where value is created’ has proved to be a powerful rallying cry to instigate a global tax reform.”); Andrew Hayashi & Young Ran (Christine) Kim, *Taxing Digital Platforms*, *Va. J.L. & Tech.*, Spring 2023, no. 3, at 1, 10 (explaining that some recent reform proposals have been justified based on user value creation and noting that that principle is one that “countries’ finance ministries, the OECD, and the EC recite and seem to endorse”).

280. See Werner Haslehner, *Value Creation and Income Taxation: A Coherent Framework for Reform?*, in *Taxation and Value Creation*, *supra* note 275, at 39, 40 (describing the academic response to the concept of value creation as “generally very critical of the concept’s meaning and usefulness to drive a coherent reform of the international tax system”).

“a phrase that has [no] meaning in modern economics,”²⁸¹ an “unhelpful” and “fuzzy notion,”²⁸² and “not even conceptually coherent as a theory.”²⁸³ The conversation surrounding aligning taxation with value creation exemplifies how exchange value continues to be at the center of tax policy discussions, despite the vital role of prediction value in the modern economy. It also exemplifies the harms caused by that continued focus. The concept of value creation is arguably unhelpful, fuzzy, and incoherent when value creation is viewed exclusively through the lens of exchange value. But, if academics and policymakers expand their notion of value creation to include prediction value, the move to align the place of taxation with the place of value creation would become more conceptually coherent.

The problem is not that academics and policymakers are not recognizing the growing importance of social data in the economy. Many have noted that it is unfair for digital companies to collect and exploit data from a country’s residents without being subjected to tax in those jurisdictions.²⁸⁴ This perceived unfairness has led to user participation proposals: reforms that would allocate taxing authority over digital companies’ income to users’ jurisdictions based on their contributions of data as well as content.²⁸⁵ But, while the importance of social data is being recognized, the emergence of prediction value as a new value form is not. Conversations around reform are still trying to fit prediction value into the familiar exchange value mold. Critics have cited difficulties in measuring

281. David Quentin, Corporate Tax Reform and “Value Creation”: Towards Unfettered Diagonal Re-Allocation Across the Global Inequality Chain, 7 *Acct. Econ. & L.*, no. 20160020, 2017, at 1, 3.

282. Wolfgang Schön, Ten Questions About Why and How to Tax the Digitalized Economy, 72 *Bull. for Int’l Tax’n* 277, 288 (2018).

283. Allison Christians, Taxing According to Value Creation, 90 *Tax Notes Int’l* 1379, 1379 (2018) [hereinafter Christians, Value Creation].

284. See, e.g., European Commission, Modern Taxation, *supra* note 279, at 4 (citing the failure of current tax laws to acknowledge value stemming from user data as an impetus for reform); see also Collin & Colin, *supra* note 175, at 53–54 (explaining the centrality of data to digital business models and highlighting that although digital companies’ data arises from the free labor of French users, France nonetheless is not able to tax these digital companies).

285. See, e.g., OECD/G20 Base Erosion & Profit Shifting Proj., Addressing the Tax Challenges of the Digitalisation of the Economy 9 (2019), <https://web.archive.org/2019-02-19/507498-public-consultation-document-addressing-the-tax-challenges-of-the-digitalisation-of-the-economy.pdf> [<https://perma.cc/GZ4Y-K8D8>] (presenting a reform proposal that would allocate taxing rights to users’ jurisdictions based on their critical role in value creation for digital businesses, including through the generation of data); see also HM Treasury, Corporate Tax and the Digital Economy: Position Paper Update 10–11 (2018) (UK), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/689240/corporate_tax_and_the_digital_economy_update_web.pdf [<https://perma.cc/2T6S-64UQ>] (citing creation of data as one of the ways in which user participation creates values for companies and advocating reforms to the taxation of digital businesses for whom the collection of user data is central to the businesses’ value creation strategies).

and attributing income to users' data creation as a barrier to user participation proposals and taxation based on value creation more generally.²⁸⁶ That difficulty of measurement stems from the conceptual incoherence of trying to translate prediction value into monetary exchange value.

Outside of the "taxing where value is created" debate, much of the discussion in policy and academic circles surrounding the appropriate taxation of the data economy is also seen through the exchange value lens. Assigning an accurate market value to data to tax it is an oft-cited challenge,²⁸⁷ as is the "cashless" nature of transactions between data subjects and data collectors.²⁸⁸ These discussions show a continued focus on fitting the square peg of prediction value into the round hole of the exchange-value-based tax system. A notable exception to this focus comes from Omri Marian.²⁸⁹ In a 2021 article, Marian proposes moving away from trying to fit the data economy to the existing income tax system by assigning monetary value to data because, he argues, doing so is "an insurmountable, if not a logically incoherent, task."²⁹⁰ On the policy level, the New York State Senate has proposed a personal consumer data excise tax that would tax data collectors based on the number of residents from

286. See, e.g., Johannes Becker & Joachim Englisch, *Taxing Where Value Is Created: What's 'User Involvement' Got to Do With It?*, 47 *Intertax* 161, 168 (2019) (discussing difficulties surrounding valuation of user data contributions); Christians, *Value Creation*, *supra* note 283, at 1381 ("[T]he idea that a given item of income produced through international trade and commerce can be fragmented geographically plainly is not true, has never been true, and no amount of normative rhetoric surrounding valuation can make it true."); Itai Grinberg, *User Participation in Value Creation*, 2018 *Brit. Tax Rev.* 407, 420–21 (highlighting administrability issues with the U.K. user participation reform proposal).

287. See, e.g., Adam B. Thimmesch, *Transacting in Data: Tax, Privacy, and the New Economy*, 94 *Denv. L. Rev.* 145, 174 (2016) ("Perhaps the biggest barrier to applying our existing tax instruments to personal-data transactions is the problem of how to value the personal data and the digital products being traded."); Aqib Aslam & Alpa Shah, *Tec(h)tonic Shifts: Taxing the "Digital Economy"* 51–54 (IMF, Working Paper No. 20/76, 2020) (discussing the impact of data valuation issues on countries' approaches to taxing the digital economy).

288. Commentators have explored the possibility of these exchanges being treated as taxable barter exchanges. See, e.g., Louise Fjord Kjærsgaard & Peter Koerver Schmidt, *Allocation of the Right to Tax Income From Digital Intermediary Platforms—Challenges and Possibilities for Taxation in the Jurisdiction of the User*, 2018 *Nordic J. Com. L.* 146, 159–60; Hillel Nadler, *Taxing Zero*, 26 *Fla. Tax Rev.* (forthcoming 2024) (manuscript at 4), <https://ssrn.com/abstract=4449094> [<https://perma.cc/ZWF9-C3DG>] ("Zero-price transactions are a form of barter exchange: consumers receive something valuable from businesses and businesses receive something valuable in return."); Thimmesch, *supra* note 287, at 162–63, 169 ("[A]pplication of [the market-exchange] model suggests that data aggregators should be viewed as engaging in taxable barter exchanges through which they sell access to their digital products in exchange for consumer data.").

289. Marian, *supra* note 31; see also Reuven Avi-Yonah, Young Ran (Christine) Kim & Karen Sam, *A New Framework for Digital Taxation*, 63 *Harv. Int'l L.J.* 279, 335–40 (2022) (commending Marian's argument and proposal and presenting an alternative reform that would also use data volume, rather than income, as a tax base).

290. Marian, *supra* note 31, at 561.

whom they collect data.²⁹¹ These efforts to push the tax system out of the exchange-value-based mold are commendable and exciting but unfortunately remain in the minority.

2. *Colliding With the Digital Economy's Scripts.* — Beyond the conceptual challenge of integrating prediction value into a tax system centered around exchange value, the digital economy's scripts and associated business practices are also colliding with tax law. These scripts and practices were beyond the historical imagination of the original architects of the tax system, and many of the assumptions about business practices that these lawmakers accepted no longer hold true. As a result, the existing tax system, when applied to the digital economy, precipitates tax outcomes that are inconsistent with the underlying norms and goals of taxation. This section briefly explores three examples of tax law failing in the informational capitalist environment. The first is the continued use of income as a tax base when companies focus on growth over profits. The second is the opportunity for advantageous tax deferral offered to companies who are not immediately converting prediction value into exchange value. The third is the international tax implications of prediction value manifesting as exchange value via an increase in company market capitalization compared to company profits.

a. *Income as a Tax Base.* — Firms are taxed on their income, not on the size of their user bases or the amount of social data and resulting prediction value they have amassed. Until this growth and expansion translates into exchange value, it exists outside the current tax system.²⁹² As explained in Part II, firms often do not earn income, instead focusing on growth through business practices such as freemium business models. Governments are unable to collect tax revenue from these digital firms despite the fact that they provide benefits and resources without which the firms would not be able to operate—benefits and resources that are funded by tax revenues. Digital firms' focus on growth over income also frustrates tax law's redistributive goals. The rise of Big Tech oligopolies has sparked concerns among various scholars, particularly regarding the concentration of prediction value and the accompanying economic and political power it brings to those firms.²⁹³ But prediction value is not part of the tax base; therefore, the tax system cannot redistribute prediction value and temper this concentration of economic and political power. Finally, the focus on growth over income frustrates the regulatory purpose of taxation. The deductions and credits offered by the tax code as a means

291. See Robert D. Plattner, *The Virtues of a Simple Excise Tax on Personal Consumer Data*, Tax Notes (Dec. 12, 2022), <https://www.taxnotes.com/special-reports/digital-economy/virtues-simple-excise-tax-personal-consumer-data/2022/12/09/7ffirb> [<https://perma.cc/9SQE-3K5W>] (describing the structure of the New York data excise tax).

292. Marian, *supra* note 31, at 561 (arguing that data, rather than income, should serve as the primary tax base in light of the rise of the data economy); Thimmesch, *supra* note 287, at 174 (chronicling the ways in which the data economy escapes taxation).

293. See *supra* section II.A.3.

to shape firm behavior are less effective when firms do not have significant income or tax liabilities to offset.

b. *Tax Deferral Opportunities.* — One possible argument against the concerns about the continued reliance on income as a tax base is the claim that all companies will *eventually* convert prediction value into exchange value. While social data as a factor of production tends to lead companies to defer short- or medium-term profits in favor of building up greater prediction value, a company's purpose is to earn profits for their shareholders, and it will eventually achieve this purpose. These profits might be earned through script one's direct conversion of prediction value to exchange value through means such as targeted advertising revenues. Or these profits might be earned by indirectly converting prediction value into exchange value through profits earned from the new or improved products and services companies are able to offer as a result of the prediction value they have accrued. Why does it matter that the tax system is not capturing prediction value when it will eventually be converted to exchange value, which the tax system will capture?

This argument is flawed in a couple of ways. First, it ignores the existence of the third script in which companies never fully convert prediction value into exchange value but instead use prediction value as a means to gain power—power that may or may not be used to create exchange value. As discussed in Part II above, the power that stems from merely possessing something of value usually justifies the taxation of wealth as well as the relationship between income versus consumption. This same rationale carries over to justify taxing companies pursuing the third script.²⁹⁴

Even if one rejects the idea that any company would pursue the third script and never fully monetize prediction value, this argument ignores a foundational consideration for evaluating the effectiveness of a tax system: the value to the taxpayer of deferring tax liabilities. The benefit of tax deferral is a fundamental concept taught to students in basic tax law classes.²⁹⁵ If a taxpayer is able to push off their tax liability into some point in the future (either by deferring income inclusion, accelerating deductions, or both), they are able to put the amount that they would have paid in taxes to productive use in the intervening period. This concept is known as the “time value of money.”²⁹⁶ For example, if a taxpayer can expect a rate of return on investment of seven percent annually, \$1 saved in taxes this year has a future value to the taxpayer of \$1.97 in ten years.²⁹⁷

294. See *supra* notes 213–214 and accompanying text.

295. See, e.g., Joseph Bankman, Daniel N. Shaviro, Kirk J. Stark & Edward D. Kleinbard, *Federal Income Taxation* 191–92 (18th ed. 2019) (describing the importance of tax deferral and the time value of money); Michael J. Graetz & Anne L. Alstott, *Federal Income Taxation: Principles and Policies* 297–303, 627 (9th ed. 2022) (same).

296. Bankman et al., *supra* note 295, at 191–92.

297. *Future value* = $PV(1+r)^n$. *PV* equals present value, *r* equals the interest rate, and *n* equals the number of periods the interest is held.

This benefit of tax deferral is an essential driver of tax planning. Tax expenditure policies, such as defined contribution retirement plans, use the benefits of tax deferral as a carrot to encourage individuals to save for retirement.²⁹⁸ Deferral is also a key feature of many tax shelters, which are designed to artificially accelerate the timing of deductions and defer the timing of income inclusion.²⁹⁹ Because tax law conceptualizes “value” in terms of exchange value,³⁰⁰ the benefit of tax deferral has historically been framed in monetary terms. But the same principle applies when a taxpayer can defer tax on prediction value. When a company is allowed to build up prediction value for extended periods without forcing any type of distributive mechanism, the company benefits by being able to accrue even greater levels of prediction value and its resulting economic and political power.³⁰¹

It is not unique to the digital economy for companies to forgo income while they build and invest in their businesses and, as a consequence, defer tax liabilities while simultaneously building economic value. But what is unique to the digital economy is the extent of this deferral. Longer periods of tax deferrals produce greater advantages to the taxpayers and greater harms to the tax system. These lengthy tax deferrals are another way in which tax law is colliding with the digital economy.

c. International Tax Implications. — Finally, the focus on growth over income within the digital economy often leads to prediction value manifesting as an increase in the market valuation of a company. To the extent that prediction value is reflected in the market value of a company, it is then converted into exchange value when an investor sells their shares and realizes capital gains income. The tax system is then able to tax the investor’s capital gains income. This is beneficial because it allows the tax system to raise government revenues and accomplish redistributive goals. There are, however, troubling normative implications for tax law when prediction value is only taxed when it converts to exchange value in the form of capital gains income at the investor level.

One of these problems emerges in the context of international tax law, specifically the determination of *which* country will have taxing rights over the digital economy’s value creation. To prevent double taxation,

298. See Graetz & Alstott, *supra* note 295, at 296–303 (describing the interaction between tax deferral, the time value of money, and retirement accounts).

299. See Bankman et al., *supra* note 295, at 504 (identifying deferral as one of the typical features of a tax shelter transaction); Stanley S. Surrey, *The Tax Reform Act of 1969—Tax Deferral and Tax Shelters*, 12 B.C. Indus. & Com. L. Rev. 307, 310 (1971) (explaining the central role of deferral to the tax shelter industry).

300. See *supra* notes 274–276 and accompanying text.

301. The exact design of a distributive mechanism for prediction value is a rich topic for future research but is beyond the scope of this Article. For further discussion, see generally Amanda Parsons, *Defining the Goal of a Data Tax*, Eur. L. Open Symposium Issue: *Taxing Data as an Instrument of Economic Digital Constitutionalism* (forthcoming 2024) (on file with the authors).

international tax law divides taxing rights over cross-border income among countries based on a system of classification and assignment.³⁰² This classification and assignment system was first developed by members of the League of Nations in the 1920s and has remained largely unchanged since.³⁰³ The system generally grants taxing rights over active business income to the source country (the country in which the business operates) and taxing rights over passive investment income, including capital gains income, to the investor's residence country.³⁰⁴

This choice in the 1920s to assign taxing rights over active business income to source countries and passive investment income to residence countries was influenced by tax law's underlying normative principles, which continue to be influential today.³⁰⁵ It was also influenced by assumptions about the nature of business activities that no longer apply in the digital economy.³⁰⁶ One of these normative principles was the benefits principle, which justifies taxation based on the benefits and resources that a country provides to taxpayers.³⁰⁷ And one of these assumptions was that any firm that increased in market value would also earn business income.³⁰⁸ Under this assumption, even though only the residence country

302. See Steven A. Dean, *A Constitutional Moment in Cross-Border Taxation*, 1 *J. on Fin. for Dev.*, no. 3, 2021, at 1, 1–3 (describing international tax law's classification and assignment system).

303. See Michael J. Graetz & Michael M. O'Hear, *The "Original Intent" of U.S. International Taxation*, 46 *Duke L.J.* 1021, 1023 (1997) ("Despite massive changes in the world economy in the last seventy years, the international tax regime formulated in the 1920s has survived remarkably intact."). For a thorough history of the development of these model treaties, see generally Sunita Jogarajan, *Double Taxation and the League of Nations* (2018).

304. See Reuven S. Avi-Yonah, *All of a Piece Throughout: The Four Ages of U.S. International Taxation*, 25 *Va. Tax Rev.* 313, 322 (2005) [hereinafter *Avi-Yonah, Ages of U.S. International Taxation*] (characterizing the international tax system as generally allocating active business income to the source country and passive investment income to the investor's residence country). For two model conventions, see *Org. for Econ. Co-op. & Dev. [OECD], Model Tax Convention on Income and on Capital* (Nov. 21, 2017), https://read.oecd-ilibrary.org/taxation/model-tax-convention-on-income-and-on-capital-2017-full-version_g2g972ee-en#page1 [<https://perma.cc/8MYD-VS5E>] [hereinafter *OECD, Model Tax Convention*]; *U.N. Dep't of Int'l Econ. & Soc. Affairs, U.N. Model Double Taxation Convention Between Developed and Developing Countries*, U.N. Doc. ST/ESA/PAD/SER.E/213 (2017), https://www.un.org/esa/ffd/wp-content/uploads/2018/05/MDT_2017.pdf [<https://perma.cc/JX96-A2BL>].

305. See Parsons, *Shifting Economic Allegiance*, *supra* note 188 (manuscript at 12–21, 25–28) (describing the norms driving the original design of the international tax system and how current debates in international tax law reveal their continued importance).

306. See *id.* (manuscript at 20) (explaining the assumption of the original designers of the international tax system that "a company whose value was increasing would also be earning income in the country in which they were operating").

307. See *id.* (manuscript at 12–21) (detailing the influence of benefits theory on the original design of the international tax system).

308. In a seminal report commissioned by the League of Nations during the 1920s negotiations (the recommendations of which were largely followed by the original designers of the international tax system), the authors stated in their analysis of which country should

would be able to tax investors on capital gains income from the sale of shares of a successful company, the country in which the company was operating, the source country, would still be able to collect tax revenues from the company itself because that company would be earning active business income, which is generally taxed in the source country.³⁰⁹ Therefore, the source country would be compensated for the resources and benefits that it provided to the company, and the benefits principle would be satisfied.

As explained in Part II above, this assumption that an increase in the value of a company will always be accompanied by income no longer holds in the digital economy under the growth- and expansion-focused business strategies of digital firms. As a result, the benefits principle goes unfilled. Countries can provide digital companies with benefits and resources that the firms rely on to achieve growth, such as infrastructure and the education of users. But, without company-level income, the source countries are unable to collect tax revenues, even when that growth is translated into exchange value when investors sell their appreciated shares. For example, Company A could have millions and millions of users in Argentina, building out its network, providing the firm with a steady stream of social data, and, in turn, contributing to a rise in the firm's market value. But when a U.S. investor in Company A goes to sell their appreciated shares, only the United States (the residence country) is able to tax that income.³¹⁰ Argentina does not get a bite at the tax apple unless Company A earns income, which it often does not under the prominent business models of the digital economy that eschew income in the short or medium term in favor of growth. This growth-without-income phenomenon and business model was beyond the historical imaginations of the original designers of the international tax system in the 1920s.³¹¹ The business model is clashing with existing international tax law, leading to outcomes that violate the normative goals of international tax law and

be granted taxing rights over capital gains income, "Corporate shares would, indeed, be worth nothing if the company had no earnings." *Econ. & Fin. Comm., Report on Double Taxation*, at 36, League of Nations Doc. E.F.S.73.F.19 (1923); see also Reuven S. Avi-Yonah, *The Structure of International Taxation: A Proposal for Simplification*, 74 *Tex. L. Rev.* 1301, 1305–07 (1996) (describing the history of the League of Nations report).

309. See OECD, *Model Tax Convention*, *supra* note 304, at M-27 (generally granting taxing rights over business profits to the source country); see also Avi-Yonah, *Ages of U.S. International Taxation*, *supra* note 304, at 322 (explaining that taxing rights over active business income are typically granted to the source country).

310. See Avi-Yonah, *Ages of U.S. International Taxation*, *supra* note 304, at 322 (explaining that taxing rights over passive investment income are typically granted to the residence country).

311. See Parsons, *Shifting Economic Allegiance*, *supra* note 188 (manuscript at 20–21) ("The imaginations of the four economists and other participants in the 1920s Compromise could not predict the rapid technological advances of past decades and the ways in which they have transformed the global economy.").

contributing to a broad sentiment that the current international tax system is unfair.³¹²

This section explores the conceptual disconnect of tax law scholars and policymakers failing to recognize prediction value as a new and distinct form of value creation and then explains a few of the ways in which the unfamiliar and unexpected business practices associated with the digital economy have collided with existing tax law. These collisions have resulted in tax law's failing to achieve its underlying goals of revenue-raising, redistribution, and regulation and have raised questions about the effectiveness of tax law in the modern economy. Both a conceptual understanding of prediction value as a value form that is distinct from, and does not always translate neatly into, exchange value and an understanding of the types of business activities that prediction value has precipitated are essential first steps for tax law to adequately respond to the challenges presented by the digital economy.

C. *Governing Social Data Value*

Data privacy (and the related field of data protection law³¹³) is the legal field historically focused on the project of governing social data. Indeed, privacy and data governance law is tasked with protecting individuals from the very same surveillance of an economic system that rewards—indeed depends on—that surveillance.³¹⁴ Given that data privacy law is the primary regime that regulates how data about people is collected, processed, and used, it not only guards against privacy violations but also serves as one of the primary legal regimes that regulates social data value.³¹⁵

312. *Id.* (manuscript at 21) (describing how the nature of value creation in the digital economy is causing outcomes that are in conflict with the underlying norms and goals of international tax law).

313. Note that much of what is called “data privacy” or “information privacy” in the United States also includes elements of data protection law. In the European Union, these are separate, though related, legal regimes. In this Article, “data privacy law” is generally used to refer to the broader category of both related regimes. For a discussion of the differences between privacy and data protection (and the tendency of U.S. law to favor the former), see Anupam Chander, Margot E. Kaminski & William McGeeveran, *Catalyzing Privacy Law*, 105 *Minn. L. Rev.* 1733, 1747–49 (2021).

314. Cohen, *Between Truth and Power*, *supra* note 4, at 40 (noting that digital platforms are designed fundamentally for “data-based surplus extraction”).

315. This depends on one's account of privacy interests and privacy law. Privacy is a “big tent” concept experiencing a high point of concept pluralism, as the past decade has seen the expansion of informational wrongs characterized within the language of privacy law wrongs. See María P. Angel & Ryan Calo, *Distinguishing Privacy Law: A Critique of Privacy as Social Taxonomy*, 124 *Colum. L. Rev.* 507, 522–29 (2024) (“[I]nformation-based discrimination and algorithmic manipulation have come to be recognized . . . as privacy problems.”); María P. Angel, *Privacy's Algorithmic Turn* 30 *B.U. J. Sci. & Tech. L.* (forthcoming 2024) (manuscript at 2–4), <https://ssrn.com/abstract=4602315> [<https://perma.cc/PZG5-WWB9>] (“In the context of the transition to artificial intelligence (‘AI’) and algorithmic decision-making systems, a big portion of scholars . . . have begun to

Data governance law faces the “mirror image” of tax law’s challenges discussed in section III.B. The problem for data privacy law is that existing laws are not designed with the task of regulating value creation, of any kind, in mind. Data privacy law’s traditional role in the commercial sphere was to grant a set of procedural guarantees: to protect individuals against personal data being collected against their will or used for purposes that exceed the boundaries of their consent.

As social data value emerges as a primary goal of production, data privacy finds itself thrust into—and grappling with—the role of regulating this value creation. This introduces a host of challenges that arise from this mismatch between data privacy’s traditional role and its role structuring social data value. The result is a legal regime poorly equipped to respond programmatically to the systematic pressures placed on privacy in a surveillance-fueled economy or to develop a positive agenda for how to manage the social stakes of prediction value.

Yet data privacy law is also comparatively well positioned among legal regimes to meet this challenge. Over the past several years, scholarly work in privacy law has begun to systematically respond to these conceptual and programmatic challenges. This section argues that distinguishing between prediction value and exchange value can provide a helpful way to translate recent pioneering work in privacy law into legal action. Thinking of the relevant tasks of data privacy law in the language of exchange value and prediction value can both identify and regulate harmful practices of prediction value production as well as foster and facilitate socially beneficial uses of social data value.

1. *Privacy Law Background.* — Privacy and data governance law governs the commercial cultivation of social data value in two ways. First, private data collection is primarily governed via interpersonal, quasi-contractual relations of individual control and consent rights. Privacy law has traditionally only contemplated social concerns regarding such data’s prediction value, and its capacity to coerce action and remake social relations, if or when it falls into the hands of public actors. This implicates the second way, which is the public regime governing privately collected social data: Fourth Amendment protection of a reasonable expectation of privacy against state intrusion.³¹⁶

consider new privacy harms.”); M. Ryan Calo, *The Boundaries of Privacy Harm*, 86 *Ind. L.J.* 1131, 1139–42 (2011) (acknowledging the difficulty of conceiving a singular definition of privacy, due to the many subconcepts seemingly covered by this term, but arguing that there is still a “need for principles that delimit privacy harm”).

316. This Article uses the Fourth Amendment as shorthand for both federal and several state constitutional privacy protections. While the Fourth Amendment grounds a substantial majority of public privacy law, courts also derive privacy rules from other provisions of the Constitution, and several state constitutions incorporate additional privacy rights. See William McGeveran, *Privacy and Data Protection Law 3* (2016) (“The word ‘privacy’ does not appear in the United States Constitution. Yet concepts of private information and decisionmaking are woven through the entire document, and courts have

While the concept of privacy itself is considerably older, U.S. digital privacy law began in the 1970s as Congress passed a rash series of bills in response to the early wave of computerization.³¹⁷ The highly influential Fair Information Practice Principles (FIPPs), first laid out in a 1973 report, canonized best principles regarding information processing and deeply informed privacy statutes in the United States and abroad.³¹⁸

These principles, as enacted in agency policies and laws, focus on proper data hygiene, data subject consent, and preventing privacy harms to individuals from which data is collected.³¹⁹ While the specifics of how the FIPPs were operationalized vary from law to law, the standard package of privacy protections they provide includes two aspects. First, negative individual rights *against* overreaches in data collection, accompanied at times by narrowly tailored inalienable data subject rights against downstream misuses of their data.³²⁰ These elements grant data subjects their privacy rights, ensuring data is collected with their consent and that certain decisions regarding how their data is used are not undertaken without additional consent.³²¹ Second, privacy laws may also include

developed a substantial jurisprudence of constitutional privacy.”). For an example of a state constitutional right to privacy, see Cal. Const. art. I, § 1. Civil public data collection is of course also widespread. It is also governed by its own set of statutory requirements that impose both procedural conditions of notice and transparency regarding what data is collected as well as substantive constraints on how publicly collected data may be processed, used, and shared between and beyond public entities. See, e.g., Margaret Hu, *The Ironic Privacy Act*, 96 Wash. U. L. Rev. 1267, 1277–79 (2019) (describing the provisions of the Privacy Act of 1974, which sets basic limits on how the federal government collects citizens’ information). While the federal provisions governing civil data processing done in the course of agency business are too many to enumerate here, the foundation and legal backbone of federal agency data collection is the Privacy Act of 1974. *Id.* at 1276–77; see also 5 U.S.C. § 552a (2018).

317. Several experts consider the year 1974—when Congress passed the Privacy Act—to be a turning point in U.S. privacy law. Declining to follow recommendations of the HEW Secretary’s Advisory Committee to pass an all-encompassing privacy law based on the Committee’s Fair Information Practice Principles (FIPPs), Congress, while maintaining FIPPs as the Privacy Act’s underlying framework, limited the Act’s scope to systems of records held by federal government agencies. See Dep’t of Health, Educ. & Welfare, *supra* note 24. Congress’s failure to legislate nongovernmental activity left other sectors on a path to piecemeal sectoral privacy laws operating against a backdrop of general consumer protection standards, a path already begun with sectoral statutes such as the Fair Credit Reporting Act (1970), Bank Secrecy Act (1970), and Family Educational Rights and Privacy Act (1974). For an overview of this history, see generally Daniel J. Solove & Paul M. Schwartz, *Privacy Law Fundamentals* 42–52 (2015).

318. Dep’t of Health, Educ. & Welfare, *supra* note 24, at xxiii.

319. For an overview of the FIPPs, see Fair Information Practice Principles (FIPPs), Fed. Priv. Council, <https://www.fpc.gov/resources/fipps/> [<https://perma.cc/WS7N-KP63>] (last visited Mar. 6, 2024).

320. For example, the Fair Credit Reporting Act (FCRA) prohibits credit score decisions on the basis of incorrect or out-of-date information. 15 U.S.C. § 1681 (2018).

321. See Margot E. Kaminski, *The Case for Data Privacy Rights (or, Please, a Little Optimism)*, 97 Notre Dame L. Rev. Reflection 385, 387–88 (2022) (“[T]he core of the [Fair Information Practices (FIPs)] are its individual data privacy process rights: notice and access,

provisions that can be understood as data protection requirements.³²² These elements impose proper processing obligations onto businesses that collect and handle data to ensure that data requests are tailored to the purposes for which data is being collected, honor the intentions of the data subject in any further sharing of their data, and impose protocols to enhance the security of data resources.

In practice, much of actual privacy management (and regulation) occurs not via courts or regulators but in the private actions of entities that develop internal compliance systems in the shadow of these rarely enforced laws.³²³ Users in turn are tasked with legitimizing these privacy practices via click-through consent, a legal approach Daniel Solove refers to as “privacy self-management.”³²⁴

From the perspective of privacy law, whether social data resides with public (as opposed to private) actors is normatively and legally significant. Constitutional privacy is drawn from several portions of the document, but the “oldest and largest body of [public] privacy law” concerns use of evidence obtained in violation of the Fourth Amendment, found to violate a defendant’s reasonable expectation of privacy.³²⁵ Notable recent case law extended privacy rights against government intrusion to some privately held data, which had long been excluded from Fourth Amendment protection under the third party doctrine—a rule that defendants no longer enjoy a reasonable expectation of privacy in information they had already disclosed to third parties (including, until *Carpenter v. United States*, just about any company that collects social data).³²⁶

In short, data privacy law, focused as it is on protecting individuals against either interpersonal harms that may arise from improper information collection and management or state abuse of prediction value, has not traditionally understood its primary aim as that of governing and managing commercial social data value creation.³²⁷ While the near-exclusive focus on prediction value via state surveillance in the field is

coupled with a (limited) opportunity to be heard. The FIPs are . . . not about protecting against the gathering and circulation of substantively sensitive data.” (footnote omitted)).

322. See, e.g., Children’s Online Privacy Protection Act, 15 U.S.C. §§ 6501–6506; 45 C.F.R. pts. 160, 162, 164 (2023) (HHS rule implementing HIPAA); see also Chander et al., *supra* note 313, at 1747–49.

323. Waldman, *supra* note 90, at 4–5.

324. Daniel J. Solove, Introduction: Privacy Self-Management and the Consent Dilemma, 126 *Harv. L. Rev* 1880, 1880 (2013).

325. McGeeveran, *supra* note 316, at 3; see also *Katz v. United States*, 389 U.S. 347, 360 (1967) (Harlan, J., concurring).

326. *Carpenter v. United States*, 138 S. Ct. 2206, 2217 (2018) (holding that “an individual maintains a legitimate expectation of privacy[.]” for Fourth Amendment purposes, “in the record of his physical movements as captured through [cell-site location information]”). On the third-party doctrine, see, e.g., *California v. Greenwood*, 486 U.S. 35, 40–41 (1988).

327. Viljoen, *Relational Theory*, *supra* note 6, at 578–79.

shifting, both popular and doctrinal conceptions of socially coercive privacy harm remain focused on public, rather than private, actors.

This results in two shortcomings. First, it ignores many salient concerns regarding informational power that arise as social data is imbricated into the strategies of commercial actors canvassed in Part II and the supporting role privacy and data governance law plays in facilitating this form of value creation. Second, it marginalizes the legal task of fostering the potential social benefits of prediction value, when cultivated fairly and responsibly. Each is addressed in turn below.

2. *Data Privacy as a Value Regulation Regime.* — The basic insight that information about people confers power is not new. Social data's capacity to endow its holders with power over others has long been the central preoccupation of privacy and surveillance scholars and many other observers of the digital economy. As discussed in Part I above, several lines of scholarship across surveillance studies, communication, and privacy law proceed from the notion that social data confers a form of power onto its cultivators.³²⁸ A growing subset of scholars in these fields also study the market imperative of social data production and the significance of this production for capitalism's informational turn.³²⁹ Indeed, as discussed above, the capacity for digital information about people to confer power on its holder animated twentieth-century concerns over information, particularly its use for scaled and systemic social control.³³⁰ Such concerns motivated the first rash of federal privacy laws.³³¹

Yet to date, the predominance of privacy accounts (and broader accounts of surveillance and digital control) regarding the control power latent in data accumulation still focus attention on the risks of social data's *political* power, when, by coercion or contract, it falls into the hands of state actors. In this classic account, social data accumulation is a source of potential concern because it can be used to impose forms of social control and remake social relations to better suit the aims of state actors.³³²

328. For a discussion of scholars in law as well as information and communication that have studied the power and control of information systems and their predictive capacity, see *supra* Part I.

329. For extended examples of scholars working on this, see *supra* section I.B.

330. See Beniger, *supra* note 63, at 16–21 (detailing how computer technologies facilitate innovations in bureaucratic organization, communication and transport infrastructure, and systemic communication by mass media, all of which produce a revolution in societal control). For a paradigmatic example of twentieth-century concerns over social control, and the perils and promise of technologies of behavioral influence, see generally B.F. Skinner, *Beyond Freedom and Dignity* 3–9 (1971).

331. McGeveran, *supra* note 316, at 619 (noting that concerns over the powerful capacity of new databases led to the passage of the 1974 Privacy Act).

332. For example, Michel Foucault's famous account of the panopticon in *Discipline and Punish* establishes the centrality of surveillance (information gathering and control) for discipline, which he defines as techniques for reordering human complexity into prosocial behavior like docility. Michel Foucault, *Discipline and Punish: The Birth of the Prison* 215 (1977). *Discipline and Punish* is considered a foundational text of surveillance studies.

Examples of this preoccupation with privacy harm as (risk of) public misuse in action abound. Consider *Sanchez v. Los Angeles Department of Transportation*, a recent case brought by the Northern and Southern California ACLU against the City of Los Angeles (the City) for its regulation of e-scooter and e-bike vendors like Uber and Lime.³³³ The regulation in question imposes a licensing scheme on these companies that requires them to share location data with the City to ensure compliance with the Americans with Disabilities Act and with other license conditions such as maintaining an equitable distribution of scooters across neighborhoods.³³⁴ The ACLU argued that in requiring companies to hand over scooter location data, the City's licensing scheme violates the Fourth Amendment (applying *Carpenter*).³³⁵ This kind of litigation choice by the ACLU reflects the view that privacy harm is a matter of state interference. The privacy risk contemplated here is not that companies like Uber and Lime are collecting the very same potentially sensitive and revealing location data (and indeed, unlike the City, are readily able to link scooter location data to people via riders' registered accounts and ride histories). Instead, the moment when legally significant privacy risks are introduced is when such data crosses the private–public divide.

Popular accounts making the case against commercial surveillance also emphasize that what makes commercial surveillance harmful is the risk of private data falling into the hands of public actors. For example, the collection of location data on gaming applications poses a risk because that information could fall into the hands of U.S. Immigration and Customs Enforcement (ICE).³³⁶ Similarly, sharing location data with prayer apps is bad because that data might fall into the hands of the military.³³⁷ People should also be wary using fertility apps in the wake of the *Dobbs* decision because that data may be used to prosecute them for an illegal abortion.³³⁸

To be clear, this Part's argument is not that one ought to dismiss or ignore these examples of privacy harm, nor to abandon the argument that state actors can pose a serious threat to privacy. Instead, it is to suggest that

333. Complaint, *Sanchez v. L.A. Dep't of Transp.*, No. 2:20-cv-05044, 2021 WL 1220690 (C.D. Cal. Feb. 23, 2021).

334. *Id.* at 3–4.

335. *Id.* at 4–5, 16.

336. See Blest, *supra* note 17.

337. Joseph Cox, How the U.S. Military Buys Location Data From Ordinary Apps, *Vice News* (Nov. 16, 2020), <https://www.vice.com/en/article/jgqm5x/us-military-location-data-xmode-locate-x> [<http://perma.cc/7RNC-8ZV6>] (detailing how the U.S. military buys location data from many sources, including a Muslim prayer app that has been downloaded over ninety-eight million times).

338. Sara Morrison, Should I Delete My Period App? And Other Post-*Roe* Privacy Questions, *Vox* (July 6, 2022), <https://www.vox.com/recode/2022/7/6/23196809/period-apps-roe-dobbs-data-privacy-abortion> [<http://perma.cc/U2J8-R5SN>] (arguing that if one wants to keep reproductive health data private and is worried about a criminal investigation, one should not put it in an app, especially since several apps have been subject to privacy scandals).

the prevailing account of how social data value cultivated by *private actors* may be abused is one of its potential misuse by *state actors*. Perhaps as a result, existing privacy law remains rather neutral about the same capacity for social control wielded instrumentally by commercial actors in service of private (exchange value-enhancing) ends. This overlooks the role privacy law plays in legitimating and structuring private exercises of prediction value, some of which arguably raise similar kinds of normative concerns over power that animate privacy talk about data (mis)use in public settings.

a. *Market Power and Political Power: Knitting the Two Accounts Together.* — Using social data value to exert political power can, and often does, coexist with strategies to cultivate market power (or pricing power). As the examples in Part II show, economic power and political power are not easily separated or understood in isolation from one another. What from one perspective looks purely like commercial activity can reallocate political power between groups. For example, the use of prediction value to manage workers canvassed in script two above are deployed to grow profits but also re-make workplace relations in ways that erode politically won workplace protections and reallocate workplace power from workers to employers.³³⁹ Strategies can exist in the interplay between the two, since political agents and public entities are also market participants. For example, Meta pursuing its standard business model (placing targeted ads) but servicing political clients (politicians) can affect voter perceptions and thus the political process.³⁴⁰ And yet other exercises of social data power by companies occur outside the bounds of direct commercial relationships and do not use social data to produce business profits directly but instead to secure favorable regulatory outcomes.³⁴¹

The idea that control over something valuable brings with it power is not an unfamiliar concept in legal fields focused on wealth and market regulation, nor is the idea that people or firms might accrue valuable resources for the sake of garnering power. But to date such arguments have predominantly focused on the power conferred by monetary wealth—namely, wealth in the easily-recognized form of exchange value. The medium such power takes either is money or is easily defined in terms

339. See *supra* section II.A.3; *supra* notes 118, 120; see also Brishen Rogers, *Data and Democracy at Work 2–3* (2023) [hereinafter Rogers, *Data and Democracy*] (arguing that “technological development and . . . the degradation of work[] were completely intertwined, in the sense that companies increasingly used new technologies to limit workers’ power”).

340. As was famously at issue in the Cambridge Analytica scandal of 2016. See, e.g., Mark Zuckerberg Testimony: Senators Question Facebook’s Commitment to Privacy, *N.Y. Times* (Apr. 10, 2018), <https://www.nytimes.com/2018/04/10/us/politics/mark-zuckerberg-testimony.html> (on file with the *Columbia Law Review*); Zeynep Tufekci, *Opinion, Facebook’s Surveillance Machine*, *N.Y. Times* (Mar. 19, 2018), <https://www.nytimes.com/2018/03/19/opinion/facebook-cambridge-analytica.html> (on file with the *Columbia Law Review*).

341. For an example, see *supra* notes 205–208 and accompanying text (discussing Project Greyball); see also Isaac, *How Uber Deceived*, *supra* note 205.

of money, such as the cash in someone's bank account or a company's financial statements. So, compared to these more traditional accounts of how wealth confers power (and why that may present social and legal problems), the story of how private power is cultivated via social data is more complex from the perspective of fields like antitrust and tax law. So, the aim here is to knit these two accounts together: layering the power latent in social data accumulation with the concerns associated with how the accumulation of value confers onto private actors economic and political power.

b. *Demand-Side Versus Supply-Side Data Regulation.* — As written, none of the privacy rules canvassed above bear on the economic motivations for collecting data that place growing pressure on privacy law's system of individual rights and corporate compliance. In short, they do not address the fact that entities violate privacy and misuse data *because* they are pursuing prediction value—and indeed, face significant market pressure to do so.

Several scholars have criticized U.S. privacy law as overly focused on individual privacy rights—what can be considered the “supply” side of the social data market.³⁴² To be clear, imposing greater data processing requirements, or expanding liability for data misuse, should (in theory) increase the cost of cultivating data value and thus exert some effect on companies' pursuit of social data value. Existing approaches, however, are likely inadequate to constrain the market imperatives to cultivate social data value.

Scholarly criticism is warranted. Sectoral privacy laws overwhelmingly confer weak rights, operationalized by systems of private compliance, and are grossly underenforced.³⁴³ In response, lawmakers have (understandably) enacted solutions that strengthen existing approaches: higher standards of consent, more expansive lists of data subject rights, and more robust enforcement mechanisms.³⁴⁴ This approach, while

342. Sebastian Benthall & Salomé Viljoen, *Data Market Discipline: From Financial Regulation to Data Governance*, 8 J. Int'l & Compar. L. 459, 466 (2021) (introducing the concept of “supply side” and “demand side” data market regulation). Privacy scholars have comprehensively covered the shortcomings of current individual consent-based laws. See, e.g., Elettra Bietti, *Consent as a Free Pass: Platform Power and the Limits of the Informational Turn*, 40 Pace L. Rev. 310, 313 (2019); Neil M. Richards & Woodrow Hartzog, *The Pathologies of Digital Consent*, 96 Wash. U. L. Rev. 1461, 1463 (2019); Katherine J. Strandburg, *Free Fall: The Online Market's Consumer Preference Disconnect*, 2013 U. Chi. Legal F. 95, 96–97.

343. See Katherine Strandburg, Helen Nissenbaum & Salomé Viljoen, *The Great Regulatory Dodge*, Harv. J.L. & Tech. (forthcoming) (manuscript at 22) (on file with author) (“The [California Consumer Protection Act] imposes relatively weak limitations on information flow and use, giving consumers only a limited right to opt out of information sales and sharing for cross-contextual behavioral advertising and of certain uses of ‘sensitive’ data.”).

344. For example, the California Consumer Protection Act (CCPA) is considered a more robust consumer-protection-style U.S. privacy law. It retains the basic package of rights

admirable, still falls short of the steps required to transform data privacy law into an effective regulator of prediction value.

In *Between Truth and Power*, Cohen uses the analogy of corn production, which will be borrowed and expanded on here to illustrate the basic point: What data privacy law currently offers is roughly akin to a set of rules ensuring that corn is properly and ethically planted, grown, and harvested.³⁴⁵ While this is, in and of itself, a perfectly legitimate set of goals, such rules would be wholly inadequate to govern and regulate the commodity derivatives and futures markets that assetize corn at scale to produce billions of dollars in downstream value.³⁴⁶ Moreover, such rules would be inadequate to manage the effects that such processes of accumulation have on the general landscape of corn production: the transformational market pressures of industrial scale production and engineered modification to make corn-as-commodity more predictable and stable to grow, harvest, store, transport, and refine. Such modifications make corn well suited to its role as a key input in maximizing derivatives exchange value but leave corn decidedly less suited to certain (previously central) use values: namely, as a food.³⁴⁷

One way of describing this problem is that current and proposed privacy laws still lack an explicit focus on creating “demand-side” checks on social data production—in other words, regulation that directly manages the economic incentives and motives that drive companies to want social data. Companies’ excessive demand for social data is due to the (arguably artificially) low costs and cheap risks associated with surveillant practices. Entities can cultivate maximum prediction value all while externalizing the current costs and future risks of doing so.

but expands the scope of actions and data covered by these rights, strengthens the usual individual rights beyond consent and access to information, and imposes higher and more frequent consent requirements. See Cal. Civ. Code § 1798 (West 2023). For a discussion and overview of the CCPA provisions, see Chander et al., *supra* note 313, at 1747–49. For an analysis of how these provisions compare to existing FIPPs-style law, see Strandburg et al., *supra* note 343.

345. Cohen, *Between Truth and Power*, *supra* note 4, at 67–68.

346. The corn derivatives market was valued at over \$75 billion as of 2022. See Kiran Pulidindi, *Industry Analysis: Corn and Corn Starch Derivatives, Market Size by Product (Corn Starch, Corn Starch Derivatives (Corn Syrup, Maltodextrin, Others), Corn Flour, Corn Meal, Corn Protein), by Application (Pharmaceutical, F&B, Biodiesel, Textiles, Others), 2023–2032*, Glob. Mkt. Insights, <https://www.gminsights.com/industry-analysis/corn-and-corn-starch-derivatives-market> [<https://perma.cc/L7GR-655Y>] (last visited Feb. 20, 2024).

347. Only one percent of corn planted in the United States is sweet corn, the kind grown to be eaten as a vegetable by people. The rest of U.S. corn is a grain, primarily used for livestock feed, ethanol production, and manufactured goods (which does not include the small portion of grain corn used for human consumption as cereal, corn starch, corn oil, and corn syrup). In Iowa, the state that produces the most corn, fifty-seven percent of corn goes to ethanol production. See Corn Facts, Iowa Corn Growers Ass’n, <https://www.iowacorn.org/media-page/corn-facts> [<https://perma.cc/AV9N-FZ24>] (last visited Feb. 20, 2024) (finding that “99 percent of corn grown in Iowa is ‘Field Corn’”).

In other spheres of commercial activity, regulation to address issues of excessive demand typically aims to discipline speculative behavior. Applied to social data value, such regulatory models might be adapted to address companies' (arguably excessive) accumulation of social data for "techcraft"—data held by an entity to attract investment but where the entity lacks a clear technical or business case for how and why such social data fits into a strategy to produce better products or services of general social value.³⁴⁸ Other possibilities include licensing schemes that require companies to state up front their purposes for data collection or that impose obligations to share data with other stakeholders who can, in turn, exert discipline on excessive, spurious, or pernicious prediction value use.³⁴⁹

Again, this isn't to say that privacy and data protection law does not, in fact, serve as a demand-side regulation—data privacy law is necessarily a regime engaged in the regulation of social data value creation. But until privacy and data governance law addresses the economic causes behind its challenges, the pressure being placed on existing data privacy law will only grow.

Finally, the lack of systematic regulation of prediction value can also prevent frank assessment of where and how data governance law can facilitate positive cultivation and use of prediction value. As work law scholars have argued, expanding access to work-generated data, under the right circumstances, can empower workers—if workers gain not only access to existing data but rights to have a say over what aspects of the workplace are datafied, and for which purposes such workplace data are used, in addition to substantive protections against exploitative uses of social data.³⁵⁰ Better regulatory means for disciplining speculative prediction value can help to ensure that social data resources are channeled towards productive uses of prediction value—tamping down on social data collection that has little social utility while fostering productive social data production. Importantly, only regulating social data prediction value when its use constitutes a form of public harm can potentially limit or constrain the capacity of more systematic regulation to foster social data value's benefits.³⁵¹

348. On the concept of "techcraft" as a means of attracting investment, see Birch et al., *Data as Asset?*, supra note 55, at 2–3.

349. See, e.g., Frank Pasquale, *Licensure as Data Governance*, Knight First Amend. Inst.: Data and Democracy (Sept. 28, 2021), <https://knightcolumbia.org/content/licensure-as-data-governance> [<https://perma.cc/V99W-QSVR>] (proposing a licensure regime for data and AI that "flips the presumption" of firms' data practices being presumed legal until there is evidence of wrongdoing).

350. Rogers, *Data and Democracy*, supra note 339, at 143–50; Dubal, supra note 120, at 44–45, 48–50 (detailing worker efforts to use existing data protection laws and experiments with data cooperatives to pool and use workplace data while also noting the importance of placing substantive bans on certain employer use of workplace data).

351. Rory Van Loo, *Privacy Pretexts*, 108 *Cornell L. Rev.* 1, 60–62 (2022) (arguing for the benefits of data management over pretextual enforcement of privacy).

3. *Promising Horizons for Regulating Social Data Value in Data Privacy (and Beyond?)*. — In his excellent synthesis of recent trends in data privacy, Daniel Susser notes that while policy responses have been uneven, trends in data privacy scholarship signal growing attention to privacy law's role as a value-regulating legal regime.³⁵² In particular, he notes two relevant shifts. The first shift is from privacy as an individual interest, and laws to strengthen rights in that interest, to the social and relational nature of privacy and the need for structural approaches to secure privacy for everyone. The second shift is from a primary focus on public actors and a rights-based model against public overreach to growing concerns over the surveillance practices of private firms that incorporate a political economy perspective.³⁵³

Both of these trends signal an openness among data privacy law scholars to view the proper role of their field as that of directly regulating prediction value across the public–private divide. Taking the second trend first, data privacy scholars that focus on the business models and scripts canvassed in Part II take as their object of inquiry the economic causes of privacy erosion—namely, the market imperative to cultivate prediction value.³⁵⁴ Transforming market imperatives to cultivate, accumulate, and exploit prediction value necessarily relates to the other trend canvassed by Susser: the move from individual privacy rights to structural and systemic solutions.

Distinguishing between social data's prediction value and exchange value can lend clarity to these programs. Distinguishing prediction value from exchange value can explain broad trends in what aspects of life are datafied. Social data whose prediction value is more convertible into exchange value under scripts one and two, such as data subjects' clicks on relevant advertisements or expressions of purchasing preferences, may be more extensively produced than social data whose prediction value is not as readily converted into exchange value. While this may seem rather obvious, it suggests that shifting trends in datafication may in turn hint at shifting technological capacities and business strategies to transform prediction value into exchange value.

The distinction can be particularly helpful in charting a path toward a positive agenda for prediction value regulation. At least some undercultivated social data (not readily convertible into exchange value

352. Daniel Susser, *From Procedural Rights to Political Economy: New Horizons for Regulating Online Privacy*, in *The Routledge Handbook on Privacy and Social Media* 281, 282 (Sabine Trepte & Philipp Masur eds., 2023).

353. *Id.*

354. See, e.g., Cohen, *Between Truth and Power*, *supra* note 4, at 25; Srnicek, *supra* note 109, at 30–31; Zuboff, *Surveillance Capitalism*, *supra* note 30, at 10 (“Surveillance capitalism’s products and services are not the objects of a value exchange. They do not establish constructive producer-consumer reciprocities. Instead, they are ‘hooks’ that lure users into their extractive operations in which our personal experiences are scraped and packaged as the means to others’ ends.”).

under any script) may nevertheless be of great predictive value for certain applications. Data privacy scholars are necessarily attuned to the harms of surveillance overreach and how excessive datafication creates personal and social disruption. It is thus not surprising that data privacy scholars and activists commonly diagnose (albeit often implicitly) the problem of the digital economy as one of too much datafication.³⁵⁵

But this is perhaps not exactly right. It is true that there is almost certainly too much datafication of consumptive choices. But there is also almost certainly too little datafication of, for example, local climate data.³⁵⁶ Distinguishing exchange and prediction value can home in on the maldistribution of social data resources here: One's shoe purchasing preferences are readily transformed into script-one-style or script-two-style exchange value, while citizen-collected rainwater data might not be. Nevertheless, detailed real-time rainwater data is of profound predictive value for understanding climate effects.³⁵⁷

Data privacy scholars are increasingly interested in distinguishing the good from the bad when it comes to scholarly accounts of datafication.³⁵⁸ Distinguishing exchange from prediction can aid the conceptual and programmatic agenda of this shift. It can help to more precisely describe current practices and identify gaps in the task of regulating prediction value. Namely, this distinction can help not only to prevent problems of speculative and excessive social data production—which result in excessively risky or ill-gotten prediction value production—but also to identify areas that actually suffer from an under-production of prediction value.

355. Daniel Susser, *Data and the Good?*, 20 *Surveillance & Soc'y* 297, 297 (2022) [hereinafter Susser, *Data and the Good?*] (finding that data privacy scholars have “an aversion . . . to articulate a positive vision” for a “data-driven society”).

356. Christopher Flavelle & Rick Rojas, *Vermont Floods Show Limits of America's Efforts to Adapt to Climate Change*, *N.Y. Times* (July 11, 2023), <https://www.nytimes.com/2023/07/11/climate/climate-change-floods-preparedness.html> (on file with the *Columbia Law Review*) (detailing how the United States lacks a comprehensive current precipitation database that could help assess flood risks). Local rainfall data is also an excellent example of how nonhuman data—this is data about rain, after all—can still be social data, insofar as such data is used to inform how and where people may safely live.

357. In comparison to local rain levels, the EPA does collect real-time local air quality data. It makes this data available to people via AirNow, an app run by the agency that people can use to assess current air quality in their area. This information was of great predictive value to people during recent periods when smoke from Québecois wildfires drifted across large swathes of the United States. See *AirNow.gov*, <https://www.airnow.gov> [<https://perma.cc/FA7V-VHDJ>] (last visited Feb. 21, 2024). It helped one of the authors who lived in the affected area during the 2023 summer wildfire season.

358. See, e.g., Solow-Niederman, *supra* note 5, at 423 (arguing for reframing privacy governance as a network of organizational relationships to manage—not merely dataflows to constrain); Susser, *Data and the Good?*, *supra* note 355, at 298 (calling for surveillance scholars to move past critique and put forward alternative conceptions of a good digital society); Birch, *supra* note 56 (considering the role of data in replacing markets and neoliberalism).

This project, while promising for governing social data value, also raises questions for future work. Data privacy law's role, as traditionally conceived, is to protect individuals' interest in data collected about them and related forms of informational overreach that may arise.³⁵⁹ This is rather distinct from the political, economic, and systemic focus of recent scholarship and the task of regulating social data value creation.³⁶⁰ The primary question concerns the proper conceptual understanding of the relationship between these two programs.³⁶¹ While both strains index important elements of informational life, what is less clear is where the categorical fault lines lie between the "traditional" conceptual terrain and legal program of data privacy and the agenda of social data value regulation canvassed here.

Though this scholarly inquiry is far from settled, the increased focus on the causes of privacy erosion—what this Article diagnoses as the cultivation of prediction value—signals a promising conceptual shift from which to develop laws better attuned to governing the production of social data value.

CONCLUSION

This Article shows how separately analyzing data's prediction and exchange values may prove helpful to understanding the challenges law faces in governing social data production and the political economy organized around it.

Part I lays out the theoretical account of social data as a materialized store of prediction value and describes how this value form diverges from traditional conceptions of value in law and beyond. Part II develops the case for the legal (and normative) relevance of this conceptual gap. As this Article shows in Part II, distinguishing prediction and exchange value is helpful in capturing with greater precision how and why entities go about cultivating, storing, and exploiting social data for gain. Part III considers how social data value fares under current legal regimes. As it shows, both areas of law that are not typically considered regimes of value regulation (like data privacy and data protection law) and those that squarely focus on regulating value (like tax law) struggle with social data value, albeit in different ways. Part III considers how the cultivation and accumulation of social data value meets, challenges, and transforms legal forms.

359. See *supra* section III.C.1.

360. To be clear, it is not this Article's contention that such rights ought not exist, or that such individual interests are not valid. It is undoubtedly the case that individuals have legal privacy interests and that privacy rights secured against public overreach remain squarely within the realm of privacy, properly understood. See *supra* notes 24–31 and accompanying text.

361. On the topic of tensions within data privacy's conceptual capaciousness, see *supra* note 315 and accompanying text.

While this Article focuses on data privacy law and tax law, its analytic approach should prove fruitful in other areas of law as well, or notably for free expression and First Amendment law, antitrust, and financial regulation. These areas are similarly grappling with the changes to economic activity that derive from the cultivation and accumulation of prediction value.³⁶² The Article's analytic separation of prediction value and exchange value is helpful in other ways, too.

First, distinguishing the value of social data cultivation and accumulation from priced exchange value helps lay bare how much of alleged prediction value creation is mere speculation with little behind the curtain. As Shapiro notes, when it comes to understanding the way prediction value is capitalized by platforms into market valuation, there is a considerable "gap between what platforms do and what they say they do."³⁶³ Clarifying the two modes of value production (and how they relate to each other) thus helps regulators or other observers assess when such claims are plausible, and when they are not.

Bringing this gap into legal view is particularly important for areas of law that manage and regulate value creation. Such regimes have an interest in distinguishing between speculative and productive activity to channel social resources away from the former and toward the latter. Distinguishing between these forms of value also matters for areas of law meant to mitigate harms arising from such gaps and from the social disruptions caused by entities pursuing growth based on dubious claims of value in either form. Speculative and harmful practices escape scrutiny and continue to flourish in the digital economy when these two forms of value are confused and obscured.

Second, while it is not this Article's aim to develop a normative account of how social data production should be regulated, this Article's work to distinguish prediction and exchange value is helpful for such efforts. The Article does not engage in a normative evaluation of when (under what conditions) and why (for what reasons) the use of prediction value may be wrongful. But this is not to say that prediction value is not cultivated, hoarded, or used in wrongful ways, nor, indeed, that certain wrongful actions are not widespread among corners of the digital economy. Separating the cultivation of prediction value from its transformation into exchange value further clarifies normative critiques lodged against social data production.

Some accounts appear to critique data production insofar as it is directed by exchange value; they take issue with the commodification of social data. Reducing complex social and ethical considerations to exchange values may degrade or violate fundamental principles. For example, Rahel Jaeggi points out that child labor is considered wrong not because of the risk that children's labor is likely to be systematically

362. See *supra* notes 79–93 and accompanying text.

363. Shapiro, *Platform Sabotage*, *supra* note 21, at 204.

undervalued by the market (a “quantitative” harm) but because of the social conviction that making a labor market for children is itself violative (a “qualitative” harm).³⁶⁴ Similarly, critics argue that a priced market in adoptions, or organs, would be wrongful even if such markets might increase allocative efficiency.³⁶⁵ Some feminists make a similar point about sex and use this normative diagnosis to argue against sex work.³⁶⁶ Is assigning a priced value to social data, or certain subclasses or uses of social data, wrong in the same way? The recent FTC proposal to ban Meta from monetizing children’s data suggests such a theory.³⁶⁷

Other accounts appear to take issue with data production in virtue of it serving as a material store of prediction value. In other words, some argue there is something particular to the cultivation of prediction value that is, or can be, wrongful. For example, Zuboff, both in her early work on “informating” and in her later work on surveillance capitalism, suggests such a diagnosis. Philip Agre diagnosed informational harm as a process of “capture,” whereby greater portions of human activity are forced into market competition with other humans through the collective project of institutions measuring them against one another.³⁶⁸ Gandy’s panoptic sort is a “disciplinary” system of power that, if left unchecked, can result in amplifying loops of growing mistrust and amplified surveillance in which “each cycle pushes us further from the democratic ideal.”³⁶⁹ Is the cultivation of material stores of predictive value independently wrongful? Legal reforms to ban outright certain forms of surveillance, such as facial recognition and other forms of biometric surveillance, suggest such a theory.³⁷⁰

Different observers may come to different conclusions. But disambiguating the two aspects of data value makes distinguishing such critiques, and the relation they bear to one another, clearer.

364. Jaeggi, *supra* note 72, at 54–55.

365. For a famous example defending the allocative efficiency of nonpriced markets in altruistic goods, see generally Richard Titmuss, *The Gift Relationship: From Human Blood to Social Policy* (1970).

366. See, e.g., Andrea Dworkin, *Prostitution and Male Supremacy*, 1 *Mich. J. Gender & L.* 1, 3 (1993) (“Prostitution in and of itself is abuse of a woman’s body.”).

367. Press Release, FTC, *FTC Proposes Blanket Prohibition Preventing Facebook From Monetizing Youth Data* (May 3, 2023), <https://www.ftc.gov/news-events/news/press-releases/2023/05/ftc-proposes-blanket-prohibition-preventing-facebook-monetizing-youth-data> [https://perma.cc/5P3R-F9JW].

368. Philip E. Agre, *Surveillance and Capture: Two Models of Privacy*, 10 *Info. Soc’y* 101, 105, 117–18 (1994). The authors thank Dan Greene, Nathan Beard, Tamara Clegg, and Erienne Weight for pointing to this example, which is cited in their recent article. See Greene et al., *supra* note 64, at 4.

369. Gandy, *Panoptic Sort*, *supra* note 65, at 260.

370. In 2020, the city of Boston enacted a local ban on the use of facial recognition technology, becoming the tenth U.S. city to do so. See *Bos., Mass.*, Code § 16-62 (2020); see also S. 2963, 191st Gen. Ct. § 26 (Mass. 2020). In 2021, Massachusetts restricted the use of facial recognition by police but did not ban it. *Mass. Gen. Laws Ann. ch. 6, § 220* (West 2021).

The digital economy has reshaped and remade both people's social lives and the laws that structure them. The business models and tactics of accumulation pursued for social data value have produced significant wealth and power, as well as significant social disruption. This Article's primary ambition is to provide conceptual language better tailored to the specificities of how information produces value and to in turn better equip the various legal regimes tasked with regulating that value and enacting social goals related to the direction and shape of the information economy.

