

DATA MARKETS: BETWEEN THEORY AND PRACTICE

Jacopo Perego
Columbia University

Preliminary, do not circulate. Latest version [here](#).

October 19, 2023

ABSTRACT

This article discusses some of the recent developments and open questions in the burgeoning literature on consumer data and privacy. Consumers supply a crucial input for the modern economy: their personal data. Yet, they have very limited control over how it is used and by which firms. When this happens, they are imperfectly compensated in return. We describe some of the inefficiencies that this status quo may create. Recently introduced legislation (e.g., the European GDPR) grants consumers control over their data, laying the legal basis upon which markets for data can flourish. In these markets, firms will compete to acquire data directly from consumers. Will these markets promote more efficient outcomes than the status quo? We discuss potential obstacles to efficiency, i.e., data externalities and consumer engagement, as well as possible remedies, such as “data unions” and “data taxes.”

This article has been prepared as a discussion paper for my talk at the conference *Future of Capitalism*, on November 10th, 2023 at Columbia Business School. The article is non-technical and is intended for a broad and non-academic audience. The references provided in this article are a great source to learn more about the topic.

1 Introduction

Consumer data has become an indispensable productive input of the modern economy. Most notably, firms use it to learn consumers' preferences and reach them with targeted advertisements or personalized offers. In such a way, this data fuels many multi-billion-dollar industries, including the online-advertising industry and digital marketplaces. While consumer data is, for the most part, supplied by consumers themselves when they interact, for example, with digital platforms, consumers typically have minimal control over it: That is, they hardly control how, when, and by whom their data is used. Additionally, when this happens, consumers are often compensated with suboptimal mechanisms, like barter: Their data is collected in exchange for a service. In the worst-case scenario, consumer data is collected without the awareness of the consumers from which this data originates (Federal Trade Commission (2014)).

This current arrangement with which firms procure such essential input appears far from ideal. It should not be surprising that it has attracted the attention of academics and policymakers. Economists, in particular, have argued why such an arrangement could harbor inefficiencies and market failures. For example, this status quo could entrench positions of market power, it could inefficiently distort the supply of data, and it could fail to protect consumers from unwanted consequences that the undisciplined use of their data may create. Economists have urged policymakers to promote alternative market designs and increase regulation and oversight (e.g., see Posner and Weyl (2018), Stigler Report (2019), and Seim, Bergemann, Cremer, Dinielli, Groh, Heidhues, Schaefer, Schnitzer, Scott Morton, and Sullivan (2022)).

The first step towards a better design of these markets is to give consumers more control over how their data is used. This is the goal of recent legislation such as, for example, the European Union's General Data Protection Regulation (GDPR). It is believed that this legislation could lay the basis for the emergence of competitive *data markets*, i.e., markets where firms would need to compete to acquire data directly from consumers—or intermediaries that act on their behalf—in exchange for economic compensation (a monetary transfer and a service).

Yet, what is a market for data, and how should we expect it to function? In the past few years, the economics literature has started investigating the properties of these markets and the

extent to which they can promote efficient outcomes. This article discusses ideas and open problems from this literature. We begin by explaining the inefficiencies that arise from the rudimentary arrangement of the status quo described above. We then explain how new legislation could facilitate the establishment of data markets. When property rights are well-defined, it is compelling to embrace a Coasian perspective and think that the aforementioned inefficiencies can be resolved by simply letting firms and consumers freely negotiate a price for the exchange of data. We explain why this is unlikely to be the case, due to data externalities. Finally, we discuss potential remedies and alternative market designs, such as data unions and data taxes.

This article aims to provide a *non-technical* introduction to the economics literature on data markets, briefly touching on open problems and the role of policy. The readers wanting to gain a deeper understanding of these concepts are encouraged to consult the references provided below as well as more-specialized reviews of this literature, such as [Acquisti, Taylor, and Wagman \(2016\)](#), [Bergemann and Bonatti \(2019\)](#), and [Goldfarb and Tucker \(2023\)](#).

2 The Data Economy

In this article, we use the term *consumer data* to refer to any information that relates to a specific (i.e., identified or identifiable) consumer. This notion includes data about her preferences—such as her demographic characteristics—as well as data serving an identification purpose—such as her IP address or telephone number.

In the past two decades, consumer data has become a crucial input into the modern economy. The reliance on consumer data is perhaps most striking in the multi-billion dollar industry of online advertising. In this industry, data is used by digital platforms to intermediate consumers and marketers. Platforms can offer marketers indirect access to the consumer data they have collected. This allows marketers to target consumers with unprecedented flexibility and precision. The great success of this business model, together with the decreased costs in data processing, have created incentives for firms to collect an ever-increasing amount of consumer data.

Online advertisement is hardly the only industry that heavily relies on consumer data. Online marketplaces, for example, extensively use consumer data to hone their services: Consumer data is used in pricing decisions, to determine which products to recommend to different consumers, and to inform the development of new products.

While data can be used in many ways and by many entities, this article focuses on consumer data (as opposed to, for example, machine data) used by firms (as opposed to, for example, governments). Specifically, our typical problem of interest is a firm using consumer data to generate personalized interactions with consumers. This can be in the form of targeted advertisements, product recommendations, personalized pricing or discounts, etc. While our scope in this article is narrow as described, we acknowledge that personal data is used in a myriad of other ways. Personal data is used in academic research, it is used by governments, and it is used for training algorithms, e.g., in the nascent generative AI industry. The breadth of ways in which personal data is used is a testament to its importance in our economy, and it is meant to increase in the foreseeable future.

This article focuses on three types of firms that collect and use consumer data: merchants (e.g., Nike), digital platforms (e.g., Google), and data brokers (e.g., Acxiom). These types of firms differ in how they collect and use consumer data. Of course, a firm could play more than one such role but distinguishing among them is conceptually useful.

- Merchants' primary business is to sell products or services to consumers. It is not to collect and use consumer data. Yet, these firms can collect data as they organically interact with consumers, for example, by observing their activity on their website. When these merchants use consumer data, they do so to personalize their offers or to hone their products and services. For our purpose, the important aspect to remark is that merchants' main reason for data collection is its direct use, as opposed to selling it to third parties.
- Digital platforms—e.g., Amazon, Meta, Google, Alibaba, and Tencent, etc.—are intermediaries in multi-sided markets. For example, they mediate interactions between consumers and merchants. Their services generate substantial value for the parties involved. Digital platforms collect consumer data in their repeated interactions with consumers,

who are attracted by the value of the services these platforms offer.¹ Platforms monetize the collected data by *indirectly* selling it to third parties, i.e., merchants. That is, platforms rarely transfer data to the merchants but, rather, they only allow them to temporarily use it for a specific purpose. For example, online advertising platforms grant advertisers access to consumer data by allowing them to bid on specific consumers' characteristics.

- Finally, data brokers—e.g., Acxiom, Corelogic, Experian, Oracle, etc.—are firms whose principal business goal is to collect consumer data. They do not sell products or serve as intermediaries. The typical data broker collects consumer data from public sources, online and offline, and packages them into individual profiles. They monetize this activity by selling these profiles to third parties, e.g. merchants. In many cases, consumers are unaware that their data is being collected and packaged by data brokers. **Federal Trade Commission (2014)** provides a detailed analysis of the data-brokerage industry.

The discussion above indicates that firms collect consumer data in two principal ways. Firms can obtain consumer data via barter, data in exchange for a service, or they can directly appropriate it without consumer knowledge and without compensation. In either case, consumers have very limited control over which data is collected, how it is used (e.g., whether it is used for advertisements or for improving the website), how long it is stored, and whether it can be re-sold to third parties.

3 Inefficiency of the Status Quo

The status quo arrangement described in the previous section is rather peculiar. The consumers, who supply one of the key productive inputs of the modern economy, appear to play a negligible role in controlling the supply of this input: They have limited control of what data is collected, how it is used, and by whom; they are imperfectly compensated when this happens; and, as a

¹See Allcott, Braghieri, Eichmeyer, and Gentzkow (2020) for an analysis of the consumer welfare effects in the case of social media platforms.

consequence, they are unable to reap the full value that their data generates for the firms that use it.

This status quo has raised the concern of scholars and policymakers on the ground that it could harbor inefficiencies, i.e., it may prevent data from being efficiently supplied and used. We discuss four main issues that this status quo generates.

- *Supply Distortions.* If consumers are imperfectly compensated for the data they supply, they may end up supplying an inefficient quantity. For example, when the platform’s service cannot be tailored sufficiently well to the type of consumer, the absence of a monetary transfer could lead some of them to supply too much data, and others to supply too little data. In other situations, i.e., when the platform tracks consumers’ activity to infer their willingness to pay, consumers may have incentives to distort their online behavior in order to obtain a different treatment. This friction decreases the quality of the database the platform collects potentially reducing social welfare. This distortion can lead to inefficiencies (see, also, [Taylor \(2004\)](#), [Villas-Boas \(2004\)](#), [Acquisti and Varian \(2005\)](#), and [Calzolari and Pavan \(2006\)](#)).
- *Market Power.* Efficiency requires that data should flow to firms that make the best use of it. If consumers had full control over their data, they would temporarily “rent” it to whichever firm promises them the best treatment. When a better alternative materializes, they would withdraw their data and employ it in such an alternative. Competitive forces should lead data toward firms that make the best use of it. In contrast, when consumers do not have control over who uses their data, a firm may not have the incentive to sell it to a more efficient competitor. In this situation, a first-mover who collected a large database may enjoy a competitive advantage relative to potential competitors. Such a situation can lead to positions of market power, which hinder trade and create inefficiencies (see, also, [Stigler Report \(2019\)](#)).
- *Non-Rival Data.* When data is nonrival, it can be used by many firms simultaneously. As a consequence, there could be social gains to data being used broadly across firms (especially in the presence of increasing return to scale). When firms hoard data and limit

its widespread use, the market mechanism can generate inefficient data allocations, i.e. too little data is shared (see [Varian \(2009\)](#), [Jones and Tonetti \(2020\)](#), [Farboodi, Mihet, Philippon, and Veldkamp \(2019\)](#), and [Veldkamp and Chung \(Forthcoming\)](#)).

- *Unwanted Usage*. If consumers have limited control over how their data is used and by whom, they may be unable to escape the unwanted consequences that such use entails. For example, suppose a data broker sells a consumer's data to a third-party firm. This firm uses this consumer's data to send unsolicited personalized offers or for price discrimination. As a consequence, this consumer may experience disutility.

4 Property Rights, Data Markets, and Data Externalities

A natural step towards addressing the issues discussed in the previous section is to give consumers control over their data. If consumers owned their data, for example, they could ultimately decide who uses it and how. When property rights are well-defined, it is compelling to embrace a Coasian perspective and think that the aforementioned inefficiencies can be resolved by simply letting firms and consumers freely negotiate a price for the exchange of data. In practice, these property rights are implemented through legislation that protects consumers' *data privacy*. [Warren and Brandeis \(1890\)](#) describe privacy as "the right to be let alone." If a consumer does not like how, when, and by whom her data is used, she can object to it and take it back.

Along these lines, the European Union's General Data Protection Regulation (GDPR) recently introduced legislation that protects consumers' right to privacy. This represents the most consequential step in the direction of empowering consumers and giving them control over how, when, and by whom their data is used. With GDPR, EU's consumers obtained several data rights. The list includes (a) the right to delete the data a firm has collected about them, (b) to object to this data being used, and (c) to transfer it to another firm. The legislation also regulates how firms should treat and store consumer data, which data can be collected, and under what conditions. The enforcement of such groundbreaking legislation comes with its own challenges. As of today, almost 5 billion US dollars in fines have been imposed on firms that

did not comply with the legislation. The vast majority of these fines have been issued to large digital platforms, including Amazon, Google, Facebook, Instagram, and TikTok.

In the United States, a growing list of States, including large ones such as California, Connecticut, Indiana, Texas, and Virginia, have passed similar laws. Today, about one out of three Americans resides in a State that has passed a dedicated data privacy law. However, a US federal law is likely to be years away.

A nascent empirical literature has started studying the effects of GDPR on firms' practices and outcomes. [Johnson \(2023\)](#) reviews this literature. This literature finds that—as can be expected—GDPR hurt firm performance by imposing costs, decreasing revenue, and thereby hurting profitability, with stronger effects on smaller firms. Yet, relatively little is known today about GDPR's consequences on consumers, due to inherent measurement challenges.

With legislation like GDPR, consumers gain rights that give them some control over their data. To what extent they will proactively use these rights is still to be determined (see also [Section 6](#)). However, this legislation lays the legal foundation upon which markets for consumer data can be established. In a market for data, firms will compete to acquire data from consumers. In some cases, consumers will receive monetary compensation when their data is used, on top of the service that the firm promises to offer. In theory, this compensation can be positive or negative (i.e., the consumer would have to pay the firm to access the service without having her data collected).

As of today, these markets hardly exist. Only a handful of firms have experimented with the idea of directly paying consumers for their data. None of them has reached a reasonable scale. One of the largest is the “Microsoft Rewards” program. This is a data collection program that provides monetary incentives to use Microsoft Bing, as well as other Microsoft products. When a consumer searches a query on Bing, she receives points that can be later converted into prizes.

The economics literature has recently started investigating the properties that such data markets will have. For example, should we expect these markets to solve the issues described in [Section 4](#)? Will these markets promote efficient outcomes? In a market for data, each consumer owns her data and decides which firm to sell it to, if any. In exchange, firms will offer

this consumer a service (e.g., intermediation with a merchant) and a transfer (a price for the data). This literature has identified several (non-pecuniary) *data externalities* that these markets may suffer from. These externalities pertain to the particular nature of the good that is traded in these markets—namely, information—and how it is used. The literature shows that when a consumer chooses to sell her data, her decision can directly affect the welfare of other consumers. It has been shown that these externalities can lead to market inefficiency. In some cases, this inefficiency can be quite severe, causing the market to unravel. Below, we briefly describe two data externalities that have been identified by the literature.

- *Externality originating from the nature of data.* The first externality originates from the fact that, due to the “social” nature of consumer data, the data of different consumers may be correlated. For example, consider two consumers, A and B, who are friends with each other. Suppose also that the incomes of people who are friends tend to be positively correlated. In such a case, if Consumer A sells her data to a firm, this firm can infer something about Consumer B’s income. Consumer A’s decision to sell can (indirectly) hurt Consumer B’s in several ways. First, if Consumer B has an intrinsic preference for privacy, she is hurt by the fact that the firm knows more than she would like. Second, the firm could use this additional knowledge about her income for price discrimination. Third, upon learning about Consumer B’s income via Consumer A, the firm will have a lower willingness to pay for Consumer B’s data. Were Consumer B to decide to sell her data, she would receive a lower compensation. Recent work by [Choi, Jeon, and Kim \(2019\)](#), [Acemoglu, Makhdoumi, Malekian, and Ozdaglar \(2022\)](#), [Ichihashi \(2021\)](#), [Bergemann, Bonatti, and Gan \(2022a\)](#) explore this externality and shows it can lead to inefficiencies in different settings, with and without competition between firms.
- *Externality originating from how data is used.* Regardless of whether consumers’ data is correlated, another kind of externality can arise ([Galperti, Levkun, and Perego \(2023a\)](#)). This occurs when the firm that acquires it is an intermediary, like Google and Amazon. In their role as information intermediaries, these platforms provide merchants access to consumer data. By their nature, however, these platforms rarely find it optimal to reveal all that they know to the merchants. Instead, it is often optimal to withhold some infor-

mation from them.² The platform withholds information by pooling together consumers with different characteristics. It then only reveals to the merchant that a consumer belongs to one such pool. The exact composition of the pool determines what a merchant knows about the consumer and, thus, the price it sets for such a consumer or the bid it submits to the auction. When a consumer sells her data to the platform, her decision affects the composition of the platform’s pools. In turn, this affects the behavior of the merchant, which in turn, affects the payoff of all other consumers in the pool. In other words, when a consumer sells her data to the platform, she creates an externality on other consumers. Recent work by Galperti, Levkun, and Perego (2023a) and Galperti, Liu, and Perego (2023b) explores this externality and shows it can lead an otherwise perfectly competitive market for data to fail.

To clarify the nature of this externality, it is useful to consider the following analogy. A school (i.e., the platform) needs to persuade employers (i.e., merchants) to hire its students (i.e., consumers). Suppose that the school policy is to provide employers with no information about the students’ GPAs upon graduation (i.e., information withholding). The employers know that, on average, the school’s students have a GPA of 3.5 and, thus, make a congruent salary offer to each one of them. Now suppose a new student whose GPA is 3 (resp. 4) joins the school (i.e., an additional consumer sells her data to the intermediary). While the employers won’t be able to tell this student apart from the rest, their belief about the average GPA will be lower (resp. higher) and so will their salary offers. Thus, the decision of the additional student to join the school creates an externality (positive or negative) on all other students. Crucially, this externality is enabled by the fact that the school withholds some information from the employers. Moreover, note that this externality is present even if ex-ante students’ GPAs are uncorrelated.

²For example, if a platform runs a second-price ad auction, it faces a tradeoff: more information improves the efficiency of the allocation but creates higher information rents for the merchants. As a consequence, it is optimal to withhold some information from the merchants (see Bergemann, Heumann, Morris, Sorokin, and Winter (2022b)). Relatedly, Google’s “quality score” effectively pools the queries of different consumers to increase competition among advertisers (see, e.g., Sayedi, Jerath, and Srinivasan (2014)); Uber can sometimes withhold riders’ destinations from drivers, so as to increase riders’ welfare.

5 Possible Remedies and Open Questions

The previous section illustrated that data markets can suffer from data externalities, which can lead to market failures. Does it imply that regulators should not promote the emergence of markets for data? The answer is negative, for at least two reasons.

First, it is important to note that the inefficiency of a data market does not imply that the outcomes induced by such market allocation (e.g., the level of consumer welfare) are worse than the status quo. It is indeed possible that an inefficient data market where consumers have control over their data leads to consumer welfare that is higher than that of a market where consumers are expropriated. Most of the literature so far has focused on identifying inefficiencies, but the comparison between these “imperfect” market designs needs to be further explored.

Second, markets with externalities are a fact of life. The inefficiencies they generate can be addressed, for example, with regulation. Below, I discuss three possible kinds of interventions that have just started to be analyzed by economics literature and will require more work in the future.

- *Data Unions*. At the heart of the externalities discussed in Section 4 is the fact that consumers can fail to coordinate. If consumers could coordinate their decisions to sell or not their data, and perhaps redistribute the compensations they receive when they do, the externalities would be reduced if not eliminated. A *data union* addresses this coordination failure (see Posner and Weyl (2018), Seim et al. (2022)).³ In a data union, consumers voluntarily relinquish their data to the union, which manages their data on their behalf. The union has a fiduciary duty to the consumers, it monetizes their data (following pre-existing agreements with such consumers), and then redistributes back to them the net proceeds from the sale. Galperti, Liu, and Perego (2023b) show that the union solves the coordination problem and guarantees that the data market is efficient.

The literature on this institution is, however, still thin, and more research is needed in

³A data union would also address the issue that individual consumers lack bargaining power vis a vis a large monopolistic platform.

this direction.

- *Increased Competition.* Regulatory pressures in the coming years may lead to increased competition among digital platforms (e.g., search-sponsored advertisements and online marketplaces). It is natural to wonder whether competition is likely to resolve the problems discussed in Section 4. The current evidence points to the direction that even perfectly competitive markets for data will suffer from the externalities discussed above (see [Acemoglu et al. \(2022\)](#) and [Galperti et al. \(2023b\)](#)).
- *Personalized Data Prices and Taxes.* An economy with externalities can be efficient if markets are sufficiently complete ([Arrow \(1969\)](#)). This requires the price system to be sufficiently rich. [Galperti et al. \(2023b\)](#) show that if data prices depend not just on the type of data that is being traded but also on its intended use, a competitive data market is efficient. Interestingly, the GDPR recognizes the importance of determining how data is intended to be used *before* it is traded: “The specific purposes for which personal data are processed should be [...] determined at the time of the collection” (Regulation 2016/679 (39)). It should be noted, however, that the richness in the price system required for efficiency is substantially more burdensome than what the law requires.⁴ Finally, taxes can be used to partially decentralize efficient allocations without the need for unrealistically rich price systems. The tax would help consumers internalize the external effect that their decision creates.

6 The Role of Consumers and Firms

In practice, giving consumers control will be useful only insofar as they care about their privacy. It is therefore natural to ask the question: Do consumers even care about their privacy? And if they do not, why should the policymakers?

⁴For example, an interpretation of the law is that it requires that consumers know whether their data will be used for, say, advertisement as opposed to product development. This would be sufficiently vague to be inadequate at correcting the inefficiencies discussed in the previous section. A narrower purpose would improve matters but it also seems impractical.

A nascent empirical literature, both in economics and marketing, has started to tackle this question. An influential paper by [Athey, Catalini, and Tucker \(2017\)](#) found that consumers' *intrinsic* value for privacy is low. In the context of a lab experiment administered to undergraduate students, the paper finds that, while students say they care about their privacy, they are willing to relinquish their data for a negligible monetary amount (a slice of pizza). Similarly, [Lin \(2022\)](#) finds that, on average, consumers' intrinsic value for privacy is small. She also finds a substantial heterogeneity in demographics and type of data and that the valuations of consumers at the right tail of the distribution often exceed the firm's valuation of individual consumer data. [Collis, Moehring, Sen, and Acquisti \(2023\)](#) finds that low valuations belong to demographic groups that are traditionally associated with low levels of digital literacy.

While this research area is very active and with many open questions, what lessons can we draw from these early findings as they pertain to the discussion of the previous sections?

First, it is important to understand that the inefficiencies discussed in [Section 3](#) do not hinge on the fact that consumers actually care about their privacy (e.g., think of the discussion about market power). In fact, even if consumers did not intrinsically care about their privacy, they could still respond to monetary incentives (e.g., being paid by a competing firm to transfer their data). Moreover, despite not valuing their privacy in an intrinsic way, consumers could still care about the *instrumental* effects that giving up their data may entail. Second, the value consumers assign to their privacy and their involvement in controlling their data is likely to increase over time, as consumers become more aware of the importance that data plays in our economy. The idea is that, as the quantity of collected data increases over time, the risk of potential misuse grows larger and so will consumer awareness. Relatedly, [Collis et al. \(2023\)](#) study that consumers' privacy values significantly increase when they become informed about how their data is used and monetized.

Finally, how do firms react in a world where consumers' privacy concerns are heightened? After all, if privacy is truly in demand, a competitive market should supply it, regardless of whether or not consumers have control over their data. Firms could gain market share by differentiating away from competitors and embracing business models that are more privacy-conscious. These business models may be less remunerative in the short term but could be a

viable niche strategy attracting a subset of consumers. Most notably, Apple has made consumer data privacy a core feature of many of its products and introduced features such as privacy labels (information), app tracking transparency (third-party cookies), and private relay (private browsing). Other Google’s competitors such as Brave (Internet browser) and DuckDuckGo (search engine) have built software that minimizes data collection and how it is monetized.

References

- ACEMOGLU, D., A. MAKHDOUNI, A. MALEKIAN, AND A. OZDAGLAR (2022): “Too much data: Prices and inefficiencies in data markets,” *American Economic Journal: Microeconomics*, 14, 218–256.
- ACQUISTI, A., C. TAYLOR, AND L. WAGMAN (2016): “The Economics of Privacy,” *Journal of Economic Literature*, 54, 442–92.
- ACQUISTI, A. AND H. R. VARIAN (2005): “Conditioning prices on purchase history,” *Marketing Science*.
- ALLCOTT, H., L. BRAGHERI, S. EICHMEYER, AND M. GENTZKOW (2020): “The Welfare Effects of Social Media,” *American Economic Review*, 110, 629–76.
- ARROW, K. J. (1969): “The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Non-Market Allocation,” *The Analysis and Evaluation of Public Expenditures: the PPB System*, Joint Economic Committee, Congress of the United States, Washington, D.C., 47–64.
- ATHEY, S., C. CATALINI, AND C. TUCKER (2017): “The Digital Privacy Paradox: Small Money, Small Costs, Small Talk,” *Discussion Paper 23488, National Bureau of Economic Research*.
- BERGEMANN, D. AND A. BONATTI (2019): “Markets for Information: An Introduction,” *Annual Review of Economics*, 11, 85–107.
- BERGEMANN, D., A. BONATTI, AND T. GAN (2022a): “The economics of social data,” *The RAND Journal of Economics*, 53, 263–296.
- BERGEMANN, D., T. HEUMANN, S. MORRIS, C. SOROKIN, AND E. WINTER (2022b): “Optimal Information Disclosure in Classic Auctions,” *American Economic Review: Insights*.
- CALZOLARI, G. AND A. PAVAN (2006): “On the Optimality of Privacy in Sequential Contracting,”

- Journal of Economic Theory*, 130, 168–204.
- CHOI, J. P., D.-S. JEON, AND B.-C. KIM (2019): “Privacy and personal data collection with information externalities,” *Journal of Public Economics*, 173, 113–124.
- COLLIS, A., A. MOEHRING, A. SEN, AND A. ACQUISTI (2023): “Information Frictions and Heterogeneity in Valuations of Personal Data,” *Working Paper*.
- FARBOODI, M., R. MIHET, T. PHILIPPON, AND L. VELDKAMP (2019): “Big Data and Firm Dynamics,” *AEA Papers and Proceedings*, 109: 38–42.
- FEDERAL TRADE COMMISSION (2014): *Data Brokers: A Call for Transparency and Accountability*, A Report by the Federal Trade Commission, May.
- GALPERTI, S., A. LEVKUN, AND J. PEREGO (2023a): “The Value of Data Records,” *Review of Economic Studies*, rdad044.
- GALPERTI, S., T. LIU, AND J. PEREGO (2023b): “Competitive Markets for Personal Data,” *Working Paper*.
- GOLDFARB, A. AND C. TUCKER (2023): *The Economics of Privacy*, NBER Conference Volume.
- ICHIHASHI, S. (2021): “The economics of data externalities,” *Journal of Economic Theory*, 196, 105316.
- JOHNSON, G. A. (2023): “Economic research on privacy regulation: Lessons from the GDPR and beyond,” in *The Economics of Privacy*, ed. by A. Goldfarb and C. Tucker, NBER Conference Volume.
- JONES, C. I. AND C. TONETTI (2020): “Nonrivalry and the Economics of Data,” *American Economic Review*, 110, 2819–58.
- LIN, T. (2022): “Valuing Intrinsic and Instrumental Preferences for Privacy,” *Marketing Science*, 41(4), pp. 235–253.
- POSNER, E. AND E. G. WEYL (2018): *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*, Princeton University Press.
- SAYEDI, A., K. JERATH, AND K. SRINIVASAN (2014): “Competitive Poaching in Sponsored Search Advertising and Its Strategic Impact on Traditional Advertising,” *Marketing Science*, September, 33 (4), 586–608.
- SEIM, K., D. BERGEMANN, J. CREMER, D. DINIELLI, C. C. GROH, P. HEIDHUES, D. SCHAEFER, M. SCHNITZER, F. M. SCOTT MORTON, AND M. SULLIVAN (2022): “Market Design for Personal

- Data,” *Policy Discussion Paper*, No. 6, Tobin Center for Economic Policy, Yale University, April.
- STIGLER REPORT (2019): “Stigler Committee on Digital Platforms,” *Final Report*, available at <https://research.chicagobooth.edu/stigler/media/news/committee-on-digitalplatforms-final-report>, September.
- TAYLOR, C. R. (2004): “Consumer privacy and the market for customer information,” *Rand Journal of Economics*.
- VARIAN, H. R. (2009): “Economic Aspects of Personal Privacy,” in *Internet Policy and Economics*, ed. by W. H. Lehr and L. M. Pupillo, New York: Springer.
- VELDKAMP, L. AND C. CHUNG (Forthcoming): “Data and the Aggregate Economy,” *Journal of Economic Literature*.
- VILLAS-BOAS, J. (2004): “Consumer Learning, Brand Loyalty, and Competition,” *Marketing Science*.
- WARREN, S. D. AND L. BRANDEIS (1890): “The Right to Privacy,” *Harvard Law Review*, Vol. IV December 15.