ANNUAL REVIEWS

*Annual Review of Economics*

# The Economics of Digital Privacy

Avi Goldfarb and Verina F. Que

Rotman School of Management, University of Toronto, Toronto, Canada;
email: agoldfarb@rotman.utoronto.ca, f.que@rotman.utoronto.ca

Annu. Rev. Econ. 2023. 15:267–86

First published as a Review in Advance on March 27, 2023

The *Annual Review of Economics* is online at economics.annualreviews.org

https://doi.org/10.1146/annurev-economics-082322-014346

JEL code: L86

## Keywords

digital privacy, data externalities, privacy regulation

## Abstract

There has been increasing attention to privacy in the media and in regulatory discussions. This is a consequence of the increased usefulness of digital data. The literature has emphasized the benefits and costs of digital data flows to consumers and firms. The benefits arise in the form of data-driven innovation, higher-quality products and services that match consumer needs, and increased profits. The costs relate to the intrinsic and instrumental values of privacy. Under standard economic assumptions, this framing of a cost-benefit trade-off might suggest little role for regulation beyond ensuring consumers are appropriately informed in a robust competitive environment. The empirical literature thus far has focused on this direct cost-benefit assessment, examining how privacy regulations have affected various market outcomes. However, an increasing body of theory work emphasizes externalities related to data flows. These externalities, both positive and negative, suggest benefits to the targeted regulation of digital privacy.

## 1. INTRODUCTION

Privacy is increasingly in the media and is the subject of regulatory discussions. This rise of attention to privacy stems from the increased usefulness of data. Digitization has reduced the cost of collection, storage, transmission, and analysis of data (Goldfarb & Tucker 2019). This, in turn, has led to the expanded use of digital data in decision making (Brynjolfsson & McElheran 2016). For consumers, data enable personalized services and products at a much lower cost, which significantly enhances consumer welfare. The use of these data can help firms improve profits and it can help individuals get higher-quality products and services that better match their needs.

The use of such data, however, has some negative consequences. Some of these negative consequences are direct. Individuals have an intrinsic distaste for the collection and use of information about them. Firms face the direct costs of obtaining and protecting consumer data. Both individuals and firms may also find that this information is used against them.

Some of these negative consequences arise from externalities. Information about one individual can be informative about another. For instance, Erlich et al. (2018) show that a genetic database needs to cover only 2% of the target population to identify nearly everyone. This negative externality is similar to the data spillovers described by Tucker (2018). When people take photos of their car with geocodes to take note of their parking spot, they can record other people and cars. This information could be used in ways that harm people who did not take photos. These negative externalities generate many of the concerns about digital privacy.

In short, privacy affects economic outcomes.

Privacy is a difficult term to define. In the nineteenth century, privacy denoted "the right to be let alone," which was recognized as fundamental to human existence and inherent in human nature (Warren & Brandeis 1890). After World War II, technological developments led to debates on the trade-offs between privacy and surveillance. Westin (1968) describes privacy as the control over and safeguard of personal information. Altman (1975) refers to the boundaries between self and others, between private and shared or public features of one's life. More recently, Solove (2008) emphasizes that there are many different facets of privacy, including information collection, information processing, information dissemination, and invasion. Nissenbaum (2009), in contrast, depicts privacy to be "the right to appropriate the flow of personal information." The appropriateness of the information flow depends on the context-relative information norms, which determine how information should flow within particular social contexts (Bleier et al. 2020).

From an economic perspective, the digital privacy literature has focused on the benefits and costs of restricting information flows (see Acquisti et al. 2016 for a comprehensive summary of the history of economic analysis on the trade-offs associated with privacy). Data are fundamentally information (Farboodi et al. 2019), which is a tool to reduce uncertainty about unknown outcomes. In the following sections, we use data and information interchangeably and use the term "digital privacy" to denote a restriction on digital data flows. Because digital information can be copied for near zero marginal cost without degrading quality, it is nonrival in the absence of effort to exclude (Goldfarb & Tucker 2019). This nonrivalry can generate both positive and negative externalities from data flows. Our economics-focused discussion of digital privacy will move away from concepts such as control, autonomy, secrecy, and right to be let alone that are used in the classical privacy literature. Instead, we concentrate on the trade-offs for consumers and firms, both directly and in terms of externalities.

In the next section, we discuss the direct benefits and costs of data flows to consumers. We follow this with a similar discussion for firms. We then turn to externalities, both negative and positive. Next, we turn our attention to a discussion of current regulatory and engineering approaches to privacy, discussing consequences in light of the benefits, costs, and externalities. We conclude with a brief summary and a discussion of open questions.

# 2. CONSUMERS' DECISION MAKING UNDER PRIVACY

## 2.1. Benefits of Privacy

In this section, we talk about the benefits of privacy. We start with the valuation of privacy and then discuss the digital privacy paradox.

**2.1.1. Valuation of privacy.** Privacy preferences can be divided into two types: those where privacy itself is treated as an intrinsic right (Warren & Brandeis 1890) and those where privacy is an instrument for protecting agents from revealing their type in a way that could impact the payoff of their economic activities (Stigler 1980, Posner 1981). The formal microeconomic models of privacy that started appearing in the early 2000s focused on the latter type, in which consumers care about privacy in order to avoid price discrimination in a repeat-purchase scenario (Taylor 2004, Acquisti & Varian 2005, Hann et al. 2008).

Specifically, online retailers have rich data on purchase history, address, and browsing history. This information could be used for price discrimination. There is an extensive stream of literature that investigates how the extraction and storage of consumer information could be utilized to design personalized pricing and targeted advertising strategies (Villas 1999, 2004; Taylor 2004; Zhang & Krishnamurthi 2004). For example, Taylor (2004) provides an early and influential perspective on analyzing consumer privacy and the market for customer information. He defines the value of customer information to be the firm's ability to identify individuals for personalized prices. This can harm or benefit consumers, depending on whether the firm has market power.

Privacy also interacts with advertising technology. Consumer preferences for privacy depend on the sophistication of the firm's advertisement targeting technology. Johnson (2013) finds that, without any intrinsic preference for privacy, consumer preferences for increased targeting are not monotone. Instead, consumer utility has a U shape in the accuracy of targeting. This shape highlights the consumer's attitude change toward advertising as technology advances. When targeting is not accurate, incremental improvement in targeting accuracy only leads to further frustration and prompts consumers to block ads. When targeting has improved sufficiently, consumers may eventually welcome it. The stage of the targeting technology therefore greatly influences consumer attitudes and preferences about privacy.

Building on Becker's (1980) framework, Lin (2022) models the intrinsic and instrumental components of consumer privacy preferences and empirically estimates them through a lab experiment. The intrinsic component of taste includes consumers' characteristics or behaviors to be kept secret. The instrumental part comes from consumers' anticipated surplus or economic loss from disclosing private information to the firm. In Lin's model, consumer $i$ has a vector of personal variables $D_i = [d_{i1}, d_{i2}, \ldots, d_{ik}]$ with $k$ types of data and a sharing decision with equal length $S_i$: Each sharing decision $(s_{i1}, \ldots, s_{ik})$ indicates whether the individual shares the associated variable. The decision $S_i$ brings an intrinsic privacy cost $C_i = [c_1, c_2, \ldots, c_k]$, a type-induced payoff from sharing, baseline compensation, and a random utility shock $\epsilon_{ik}$ to the consumer's utility specification:

$$U(\mathcal{S}_i; C_i, D_i) = \sum_k - \underbrace{c_k(X) \times s_{ik}}_{\text{intrinsic preference}} \tag{1.}$$

$$+ 1_{\text{inst}} \times 1_{k\in\{1,2\}} \times \underbrace{\beta \times p_i \times w_k \times \hat{E}[d_{ik}|\mathcal{S}_i, D_i]}_{\text{type-induced payoff}} \tag{2.}$$

$$+ \underbrace{\beta \times p_i \times s_{ik}}_{\text{utility from compensation}} + \epsilon_{ik}. \tag{3.}$$

In the above utility specification, each intrinsic privacy cost $c_k$ could be expanded to a function of observables $X$ in line 1. In line 2, $1_{\text{inst}}$ is an indicator for instrumental privacy concerns, and $1_{k \in \{1,2\}}$ indicates the information-sharing decisions influenced by the instrumental incentives. While data type $k$ could include many types, Lin emphasizes and measures two: income and purchase intent. In the type-induced payoff, consumer $i$ has two types of beliefs: first-order belief and higher-order belief. The consumer's first-order belief is $w_k$—their expected increase in the percentage winning probability for an adjacent, higher type. The consumer's higher-order belief is $\hat{E}[\cdot]$—their expectation of the firm's expectation about their type. In addition, $\beta$ stands for the marginal utility of monetary rewards, and $p_i$ represents the compensation offered for the data. In line 3, the utility from baseline compensation is $\beta \times p_i$, which in Lin's specification is proportional to the number of shared variables $s_{ik}$.

Lin (2022) finds heterogeneous and right-skewed intrinsic preferences of consumers with a mean valuation of \$10 for sharing a demographic profile and a 97.5% quantile of \$30. Whether and how consumers opt to share data depend on the heterogeneity and correlation of the two main components in their preferences. This framework, which explicitly recognizes intrinsic and instrumental aspects of consumer privacy preferences, underlies our interpretation of much of the rest of the literature.

Tang (2019) estimates the value of privacy for online borrowers using large-scale field experiments. Her structural model, by linking individuals' disclosure, borrowing, and repayment decisions, is able to quantify the monetary value of personal data. She shows that individuals value privacy and measures the intrinsic part of the privacy: The social network ID and employer contact information are valued at \$32, accounting for 8% of the value of a foregone loan.

#### 2.1.2. The digital privacy paradox.

Tang (2019), Lin (2022), and others provide evidence that consumers do care about digital privacy. Nevertheless, privacy preferences are context dependent and have changed over time. For example, Goldfarb & Tucker (2012b) use survey data with 3 million responses from 2001 to 2008 to document that older consumers are more privacy sensitive than younger consumers and that overall privacy concerns are rising over time. Privacy sensitivity is measured by refusal to provide personal information about income, age, or zip code in a survey. This change in privacy concerns over time appears to be driven by an expansion of the types of data that consumers consider to be private. Specifically, consistent with Nissenbaum's (2004) concept of contextual integrity, privacy concerns in nonpersonal contexts (e.g., entertainment, consumer-packaged goods) grew more rapidly as cross-context data exchange became more common. Acquisti et al. (2015) also measure privacy using refusal to provide information in a survey, and they similarly demonstrate an increasing concern about privacy over time. Both studies document that privacy concerns grow as consumers are immersed in more sophisticated data-sharing practices when using digital products and services.

Despite these measured benefits of privacy to individuals, and despite evidence of increasing concern for privacy, consumers continue to give out large quantities of personal information. There is often a gap between consumer's stated and revealed privacy preferences. This phenomenon is labeled the privacy paradox (Norberg et al. 2007). Individuals' valuation of privacy is affected by contextual and nonnormative factors. Acquisti et al. (2013) establish a notable gap between individuals' willingness to accept (WTA) and willingness to pay (WTP) in a field experiment. They show that the cash-for-privacy exchange is larger when individuals consider getting money from trading out their data, and it is smaller when people pay for privacy. Athey et al. (2017) convincingly establish a digital privacy paradox through three main empirical findings—dubbed small money, small costs, and small talk. "Small money" recognizes that people are willing to relinquish private data quite easily when they face small incentives, though they claim they care about privacy. In the study, people gave up the email addresses of their friends in exchange for a slice of pizza. "Small costs" speaks to the fact that small frictions in navigation costs can efficiently reduce technology adoption, even when the consequences are transparently presented. People typically did not select a privacy-protecting option from a list of four when the privacy-protecting option was at the bottom of the list. "Small talk" shows that an irrelevant aspect of privacy in the particular context studied, encryption, could provide an illusion of protection and reduced privacy-enhancing behavior.

The privacy literature has suggested several explanations to understand the privacy paradox. Burtch et al. (2015) emphasize consumer ignorance or lack of attention about how data might be used, demonstrating that delayed presentation of privacy policies increases revenue in an online fundraising context. Adjerid et al. (2016) also find that reminders about privacy policies tend to lead individuals to opt out. They point out that privacy reminders reduce the usage of health information exchanges, unless combined with subsidies for adoption.

Related to the role of consumer ignorance in explaining the privacy paradox is the impact of giving consumers the perception of control. Tucker (2014) examines the relationship between users' perception of control over their personal data and the likelihood of them clicking on Facebook's advertisements. As shown in **Figure 1**, reproduced from Tucker (2014), personalized advertising was relatively ineffective before the introduction of policy that increased consumers' perceived control over personal data flows. After the policy was introduced, personalized advertising was nearly twice as effective at attracting users, even though these controls were not directly related to the way the data were used.

Chen et al. (2021) emphasize correlated preferences between benefiting from data flows and the desire to protect privacy. Combining survey and behavioral data on the Alipay platform—one of the largest Chinese online payment and lifestyle platforms, with more than 900 million active users as of 2022—they find that users with stronger privacy concerns in the survey tend to give out authorization and use the app services more frequently and extensively. They argue this is driven

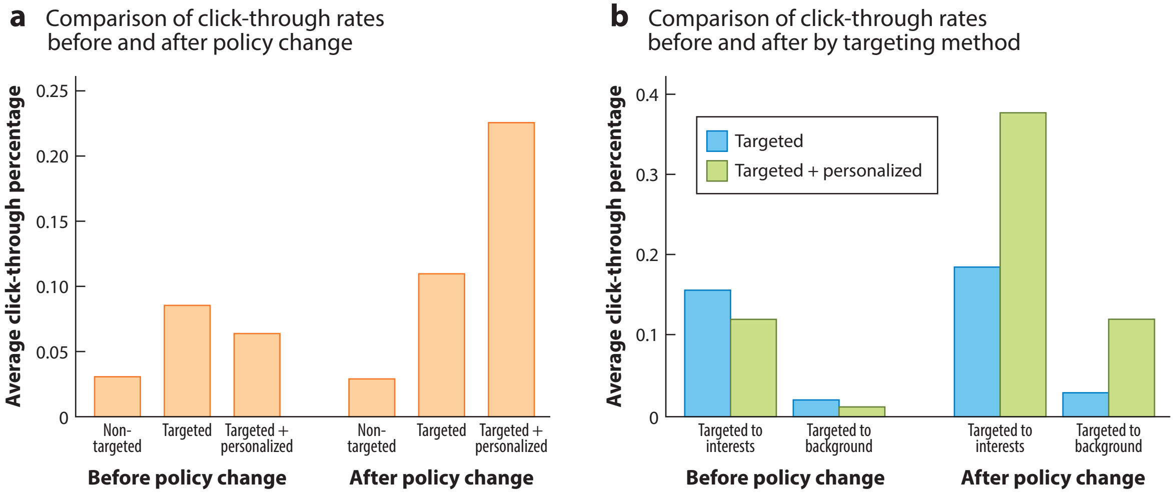

**Figure 1**

Consumer click-through rates before and after the introduction of policy.

by an instrumental value of privacy. It is the intense use of digital services by the most active users that creates a stronger preference for privacy. It is an open question whether this correlation is exogenous or whether digital preferences cause increased privacy concerns.

Solove (2021) highlights a methodological explanation. The observed behavior is measured in very specific contexts, while self-reported privacy concerns tend to come from general surveys. Thus, the latter may not correlate closely with the former.

The literature overall suggests a privacy paradox, in the sense that individuals claim to care a great deal about how their data are used but appear to act as if they did not care. A number of explanations have been put forward related to consumer ignorance, correlated preferences between privacy and the benefits of data, and methodological issues.

## 2.2. Benefits of Data Flows to Consumers

Data flows in the digital world bring substantial economic benefits directly to consumers. Consumers enjoy new services such as search engines and recommendation systems, personalized advertising and offers, and targeted products and services. Moreover, customized communication can reduce information overload and assist customers in making informed decisions (Ansari & Mela 2003). When firms have access to data, prices can fall. For example, Kummer & Schulte (2019) use data from 300,000 apps on the Google Play Store and document that paid apps ask for less consumer data than free apps. Apart from the economic benefits, consumers may experience direct psychological benefits from sharing data. Tamir & Mitchell (2012) discover that human self-disclosure activities—i.e., sharing information with others—engage neural and cognitive mechanisms associated with rewards.

**2.2.1. Better service and personalization.** Open data flows grant firms more information to tailor their products, services, and communications to an individual customer, which we refer to as personalization. Personalization can reduce information overload, which aids consumers in making efficient decisions. By allowing firms to learn their preferences, consumers benefit from reduced customer search costs (Goldfarb & Tucker 2019), so that the right products and messages could be delivered to the right person, at the right time.

To understand the value of personalization to consumers, Sun et al. (2021) conduct a large-scale field experiment on the Alibaba e-commerce platform, involving a random sample of 555,800 customers. By banning the use of personal data in the homepage recommendation algorithm, they observe a sharp decrease in both customer engagement (click-through rate and product browsing) and market transactions (sales volume and amount). Specifically, the customers' click-through rate on the recommended products drops immediately by 75%, and the customers' browsing behavior on the homepage is subsequently reduced by 33%. As a result of the two combined effects, purchases fall 81%. The analysis indicates that the value of personalization in e-commerce is large for the whole consumer group. Moreover, it disproportionately benefits newer customers, those with less purchase power, females, and those from developing regions.

In addition, Chan et al. (2022) show the great benefit of expanded credit access from digitally verified data. They document that the better-verified data increase average loan origination rate by 35.5%, without substantially raising the interest rates charged on these loans. The effect is especially significant for deep subprime and subprime consumers, with a 146% and a 44% raise in the loan rate, respectively. On the lender side, lenders also enjoy an estimated 19.6% increase in profit from the expanded credit access.

Similar results have been found in health care. When data flows are easier, electronic medical record (EMR) adoption is higher, and patients benefit. Miller & Tucker (2011) show that the improved monitoring capacity resulting from the adoption of EMR can reduce neonatal mortality rates. Furthermore, Derksen et al. (2021) find that the introduction of an EMR system to track down HIV patients in Malawi immediately enhances the number of patients actively in care and reduces patient mortality. The use of this system is limited by privacy permissions. Therefore, patients' privacy preferences can inhibit the effectiveness of the EMR system significantly. In this sense, data flows improve patient health.

**2.2.2. Price discrimination and data flows.** Price discrimination is central to the instrumental value that consumers receive from privacy. However, in equilibrium, under certain circumstances data flows can increase consumer surplus.

Conitzer et al. (2012) recognize that one way for data flows to increase consumer surplus is related to the Coase conjecture (Coase 1972). Consider a monopolist who can track individual past purchasing patterns in order to price discriminate, and consumers can in turn conceal their personal data at a cost. When maintaining anonymity is costly, the seller has better capability to identify old customers and to price discriminate. Knowing that, consumers will hesitate to make the initial purchase. Anticipating this, the seller is constrained to provide a lower initial price, which dominates the profit increases arising from future price discrimination. As a result, the seller would prefer to commit to a no-price-discrimination case. When the cost of maintaining anonymity is low—say, consumers can freely anonymize themselves—all individuals will choose to do so, resulting in the highest profit for firms. Illustrating this hide-and-seek game, Conitzer et al. (2012) provide a distinct perspective to the debate. When privacy is costly for consumers, they can be better off. As such, providing privacy protection can reduce consumer surplus and social surplus when the cost of maintaining anonymity is low.

Absent these dynamic considerations, competition can serve to prevent price discrimination via personalized pricing. While a monopoly firm with access to consumer data can make consumers worse off through the improved match values and more aggressive pricing, Loertscher & Marx (2020) demonstrate that in this setting, if the price is regulated, a reduction in privacy will always benefit the consumer because of the improved matches. In this model, maintaining competition is more important than privacy protection to advance consumer surplus. Furthermore, Miklós-Thal & Tucker (2019) demonstrate that better consumer information can decrease collusion and foster competition in the market. With better demand forecasting, colluding firms face a higher temptation to deviate to a lower price. The overall effect suggests that better forecasting from data flows leads to lower prices and higher consumer surplus.

Data flows may also incentivize prosocial behavior in a way that reduces prices and maximizes consumer surplus. Usage-based car insurance (UBI) can provide individualized price discounts based on driving behavior. Safe drivers self-select into the UBI program to pay a lower premium. The program with its economic incentives can motivate UBI participants to adopt better driving habits (Jin & Vasserman 2021). UBI has a measurable effect on reducing fatal auto accidents (Reimers & Shiller 2019) and promoting good driving habits in the long run. The average daily hard-brake frequency dropped by 21% for UBI customers after six months. Young drivers and female drivers show more improvements and benefit more from the program (Soleymanian et al. 2019). When consumers choose to drive safer to get a lower price, they also increase the total social welfare to the extent that safe driving generates positive externalities.

Thus, while price discrimination can give rise to instrumental privacy concerns, there are situations in which open data flows may increase consumer surplus even when those data flows facilitate price discrimination. Ultimately, the impact of data flows on consumer surplus (in an instrumental privacy sense) depends on the particular context.

# 3. FIRMS

## 3.1. Benefits of Data Flows to Firms

Data flows from consumers have created new opportunities for firms. Firms can set personalized pricing, send targeted advertisements, and improve customer relation management. Data have created new markets and, in some circumstances, have increased market power.

**3.1.1. Personalized pricing.** Consumer data allow personalization. This suggests a potential for first-degree price discrimination (Shapiro & Varian 1998, Smith et al. 2001). A broad theoretical literature has arisen suggesting opportunities for firms (and welfare losses for consumers) from digital price discrimination (Chen & Iyer 2002, Taylor 2004, Acquisti & Varian 2005, Hermalin & Katz 2006). However, as noted by Goldfarb & Tucker (2019), the theoretical literature on the use of data for enabling digital price discrimination appears to be more developed than both the empirical studies and the industry practices. Well-documented examples of first-degree price discrimination are limited.

**3.1.2. Targeted advertising.** Data flows enable targeted advertising, which benefits firms, particularly small firms (Goldfarb 2014). Unlike personalized pricing, there is a great deal of evidence showing that firms use data flows to target online advertising to consumers. Targeting allows the firm to endogenously increase differentiation in the market and avoid wasted advertising. In other words, targeting improves advertising effectiveness (Iyer et al. 2005). For example, Rafieian & Yoganarasimhan (2021) explore the targeted advertising in the mobile in-app advertising context. Their proposed machine learning framework of an efficient targeting policy is estimated to improve the average click-through rate by 66.80% over the current system.

Targeted advertising can affect market power in ways that benefit advertisers (and perhaps consumers) at the expense of firms. Athey & Gans (2010) use a model with local and general outlets to analyze the impact of targeting on the supply and the price of advertising. The authors specify that the improved efficiency of information allocation from targeting leads to a demand increase. However, it may reduce the market power of each individual publisher if advertising space or advertiser capacity is not constrained. Bergemann & Bonatti (2011) model competition between online and offline media, whose main difference lies in the targeting ability based on consumer data. Better targeting improves consumer-product matches and thus the social value of advertising. At the same time, greater targeting amplifies the concentration of firms advertising in each market, which eventually leads to lower advertising prices received by the advertising market due to lack of competition between the advertisers.

**3.1.3. Customer relationship management.** Data enable firms to understand customer needs. Consumers' granular activity data can aid firms in implementing proactive retention strategies. For example, in subscription services, consumer data can reveal with customers are at risk of stopping their subscriptions. These data can also reveal the marginal impact of different interventions aimed at retaining the customer (Ascarza 2018).

**3.1.4. New types of firms.** Digital data flows have enabled a market for data in which a new type of firm, the data intermediary, plays an important role. Aside from the direct data flows from consumers, firms also benefit from the third-party data collected, aggregated, and organized by data intermediaries. Bergemann & Bonatti (2019) survey the growing literature on data markets and emphasize the role of data intermediaries that sell user information, which ranges from direct sales of lists of consumers with certain characteristics to indirect sales of data through sponsored search and retargeting. A growing theory literature models these data intermediaries and their optimal mechanisms in interacting with consumers and advertisers (Bergemann & Bonatti 2015, Bergemann et al. 2018, Yang 2022).

In addition to being third-party data brokers, the data sellers can be a platform with advertisers and consumers on both sides. De Corniere & De Nijs (2016) consider a setting where a platform makes a decision on disclosure or privacy—that is, whether to sell the consumer information gathered from one side of the platform to the advertisers on the other side of it. The model shows that disclosure improves the match between advertisers and consumers but raises prices, even without price discrimination. Disclosing information, in certain conditions, could increase the total profits of the platform and the advertisers while leaving an information rent to the winning bidder. The results are in line with Bergemann & Bonatti's (2015) finding that it is not optimal for data intermediaries to disclose the finest consumer information to firms, since the informational rent is passed on to firms. There are certain conditions in which the intermediary optimizes its profits with an intermediate level of privacy.

## 3.2. Benefits of Privacy to Firms

To the extent that consumers value privacy and purchase from firms that have strong privacy policies, firms benefit from privacy directly. In addition, firms can benefit through the reduced costs associated with data and through market power.

**3.2.1. Direct cost of data flows.** Collecting and securely storing data is costly. Companies face challenging legal obligations and compliance requirements. They incur costly investments in protecting stored consumer data from malicious access from third parties, such as cyberattacks. The data protection comes in the forms of application programming interface (API) updating, improved firewalls, and vulnerability checking from company-hired hackers. The benefits of data may be small relative to this cost (Shy & Stenbacka 2016). For example, Chiou & Tucker (2017) investigate whether larger quantities of historical data affect the accuracy of subsequent searches and, thus, the firm's ability to maintain market share. Historical data are costly to store, and they create security risks. The authors find no empirical evidence that reducing the length of data retention would harm the accuracy of search results. Similarly, Yoganarasimhan (2020) demonstrates that the returns to search personalization are concavely increasing with the length of user history data. In a field study, Neumann et al. (2019) find third-party consumer profiling often to be economically unattractive due to the high additional costs of targeting solutions and their limited accuracy. Bajari et al. (2019) provide both theoretical guidance and empirical support for the countervailing force of diminishing return to data.

**3.2.2. Market power.** Although data flows help firms acquire consumers, too much data can reduce a firm's market power. Therefore, it can be beneficial for firms to maintain an intermediate level of privacy for consumers. Choe et al. (2018) consider a two-period model in which two firms compete dynamically through acquiring consumer information in their first-period purchases and offering personalized pricing in the second period. No matter whether product differentiation is exogenously or endogenously chosen, both firms end up worse off compared to when they use simpler pricing strategies or commit to substantial product differentiation. When the use of customer information is solely for pricing, more customer information is typically bad for competing firms because of the intensified competition in the first period of information gathering.

Data-enabled price targeting could intensify price competition, which may hurt the competing seller with better quality. A higher-quality firm can be worse off with personalized pricing (Choudhary et al. 2005). Casadesus-Masanell & Hervas-Drane (2015) demonstrate that the existence of a data market can also lead a low-quality firm to translate its competitive pressure to consumer data disclosure. When firms have two revenue sources—the sales revenue from products and the disclosure revenue from trading consumer data—the presence of the additional revenue stream from data sales harms the quality-improvement incentives. Consumer data soften the intensity of competition when consumers are heterogeneous and firms focus on differentiating their privacy policies.

# 4. EXTERNALITIES

Thus far, the focus of this review has been on the direct impacts of data flows and privacy protection on consumers and firms. However, data have externalities. Much of the more recent research on digital privacy has focused on these externalities (Choi et al. 2019, Acemoglu et al. 2022, Bergemann et al. 2022). Data externalities occur when an individual shares data and the data also reveal information about others. The externalities can be negative or positive.

## 4.1. Negative Externalities of Data

> Data is the pollution problem of the information age, and protecting privacy is the environment challenge.
>
> —(Schneier 2015, p. 238)

The negative externalities of data provide insights into answering the questions, Why do consumers tend to allow some forms of data collection even if they are fully aware of the potential for the data to cause harm? Under what circumstances will firms collect too much personal data, both for their customers and for themselves? What is the role of policy in regulating data flows?

One person's data provide information about others in three possible ways. The first is direct: A person's list of contacts includes information about that person (who their friends are) but also information about their contacts. Similarly, a social media feed contains information about the account owners' posts and likes but also about the posts and likes of their connections, and sometimes the connections of their connections. Thus, an individual's decision to share data may directly affect others. The second and third ways are indirect. The second way is that individual preferences and behaviors are correlated: Therefore, information about one individual provides probabilistic information about others. The third way lies not in the data themselves but in the data-generating process: By choosing to withhold information, consumers may reveal their types in the market activities as well. In this regard, individuals may provide information about the instrumental value of their data.

Choi et al. (2019) emphasize the second type of externality. They model a monopolist platform's data collection. The market equilibrium is characterized by the excessive collection of personal information, in particular, the excessive collection of sensitive information with negative externalities. The primary mechanism is that individuals do not take into account the spillover effects of their data sharing. Even if fully informed agents make individually optimal decisions, the outcome may not be socially efficient. Bergemann et al. (2022) demonstrate that this data externality means that each individual has little incentive to keep their data private and that selling the data to a data intermediary will yield near zero compensation, even if the individual knows the information can be sold to a firm that seeks to extract their surplus. Acemoglu et al. (2022) model a data market where a monopoly digital platform can trade users' data. They show that there exists an equilibrium in which too much data are shared, and the price of data is depressed. This equilibrium is directly due to data's negative externalities, which lead to excessive use of data by platforms and firms. Furthermore, the data prices will no longer reflect consumers' value of data and privacy. In addition to the overuse of data, the externalities also shift surplus from users to the platform and firms.

Ichihashi (2020) examines how a commitment to avoid price discrimination could make the seller better off and the consumer worse off. The mechanism is that under the commitment regime, certain consumers choose greater disclosure, which leads to higher prices and therefore lowers the welfare of other consumers. As consumers fail to internalize this negative effect, they opt for the highest level of disclosure, even though they could benefit from collectively withholding information. In related work, Ichihashi (2021) provides a model of how the firm and consumers divide the surplus created by data externalities. The impact of the data collection depends on what data externalities consumers impose on each other: It could be beneficial or harmful, depending on whether the allocation of data is substitutable or complementary. However, firms tend to collect too much data because of the spillover effect from one consumer to another.

Overall, the literature on emphasizing externalities from correlated behavior and preferences has emphasized that firms usually collect too much data relative to a welfare-maximizing benchmark.

There is also information in the choice not to provide data. For example, by not providing data, consumers may reveal their types in terms of advertising responsiveness and willingness to pay (Bergemann & Bonatti 2015). This effect depends on the proportion of consumers who withhold data because of their intrinsic value for privacy. If enough consumers have an intrinsic value of privacy, the firm cannot infer the instrumental value of an individual's data based on the decision to withhold those data.

Despite the extensive theory work, there is still a lack of empirical evidence on the data externalities that could guide the policy debate.[1] Empirically documenting the negative externalities from data highlighted in the large and growing theory literature is a promising area for future research.

## 4.2. Positive Externalities of Data

Data also have positive externalities. These can benefit consumers and firms, increasing overall welfare. For example, the benefits of Google search results come from the data flows from all users' search activities. Likewise, the recommendation functions on Spotify and the "frequently bought together" section on retail websites are available because of other users' data on consumption patterns.

**4.2.1. Productivity and data economy.** A growing macroeconomic literature emphasizes the data economy (see Veldkamp & Chung 2019 for a survey). Jones & Tonetti (2020) examine how property rights for data determine their use in the economy. Data serve as an input into the development of high-quality ideas, and the nonrival nature of data means that there are social gains to the wide use of data. If firms own property rights over data, then data might be hoarded, and the social gains will go unrealized. In contrast, if consumers control the property rights, the data may be used more broadly. While Jones & Tonetti's paper does not emphasize the negative externalities to consumers highlighted above, a key implication is that the direct benefits that consumers get from privacy, and the externalities that firms get from data, can be addressed through giving property rights to consumers.

Farboodi & Veldkamp (2022) emphasize the impact of data as a byproduct of economic activity on productivity and economic growth. Data serve to improve predictions and thereby optimize

[1] There is, however, empirical work on externalities of a different type. Goh et al. (2015) document a negative externality from the decision to withhold data on others in the context of the US Do Not Call registry. They show that when more consumers register to block marketing calls, the remaining consumers receive more calls.

business processes. In the short run, data may have increasing returns as firms with many customers gather data, which in turn improves productivity. This allows the firm to attract even more customers. In the long run, however, the data economy does not generate sustained growth. Data have diminishing returns to improving predictions. With respect to externalities from data, the core takeaway from the work of Farboodi & Veldkamp (2022) is that there are two opposing forces—increasing and decreasing returns to data—and that despite nonrivalry, increasing returns, and the production of data as a byproduct of economic activity in the short run, data production is efficient in equilibrium. With a natural bound on the prediction error, data have diminishing returns in the long run. Therefore, the positive externalities do not create incentives to subsidize data. Negative externalities, as expected, generate too much data in equilibrium.

**4.2.2. Socially beneficial behavior.** Information (data) disclosure is not always harmful to individuals. It may in some cases decrease information distortion, incentivize prosocial behaviors, and improve social welfare. Following Bénabou & Tirole (2006), who explore individuals' incentives in the pursuit of the prosocial activity, Daughety & Reinganum (2010) develop a model of the economics of privacy whereby individuals' actions generate externalities. Under the regime of privacy, agents choose their full information optimal actions, while under publicity, they distort to enhance others' perceptions of themselves. The trade-off arises between the expected disutility due to signaling and the increased contribution to the public good. The model includes three primary elements: an intrinsic value for the activity, esteem (and, by contrast, social disapproval), and the consumption of public goods that arise from the aggregate activity of all other individuals. When the disutility of distortion is low relative to the marginal utility of the public good, a policy of publicity is optimal. Of course, the use of data to incentivize prosocial behavior can have negative consequences, an issue related to what Tirole (2021) labels digital dystopia.

## 5. IMPLICATIONS

We organize the discussion of the implications along two perspectives: regulatory implications and nonregulatory privacy protections. The externalities highlighted above suggest a role for targeted government intervention. In addition, consumer preferences create incentives for firms to invest in nonregulatory privacy protections.

### 5.1. Models of Regulation

Privacy regulations restrict the flow of data. In the process, they can protect consumers from direct and indirect harms from data flows. They can also encourage firms to avoid competitive pressures that may degrade consumer privacy and security. Privacy regulations can also have negative consequences on market outcomes, particularly with respect to competition, innovation, and both producer and consumer surplus.

**5.1.1. Consumer behavioral factors.** Most current privacy protection regulations are designed under the assumption that consumers are uncertain and vulnerable to firms' actions. An extensive stream of literature has established that trust can reduce privacy concerns. Kummer & Schulte (2019) detail that the privacy concerns of consumers in the app installation process are influenced by the reputation of the application developers. Similarly, Chen et al. (2021) show through an experiment that trust plays an important role in people's data-sharing intent on the Alipay platform. The opt-in rate dropped following the Alipay logo removal, which did not change any actual contracts but affected the perception of trust. The effectiveness of regulations varies due to different levels of consumer perception of control. Miller & Tucker (2018) focus on how different features of privacy regulation lead to various effects. Among three alternative approaches to protecting patient privacy, they find that notice-and-consent deters individuals from obtaining genetic tests, while an approach that grants users control over redisclosure encourages the spread of genetic testing. The positive effect of redisclosure on data flows stems from its ability to provide consumers with the perception of control over the use of consented data. Additionally, Baye & Sappington (2020) show that the impact of privacy policies can depend on consumer sophistication. If sophisticated customers know enough to make optimal decisions in the absence of regulatory protection, then regulations aimed at protecting unsophisticated customers may do so at the expense of sophisticated consumers.

### 5.1.2. Competition and innovation.

Privacy regulations could increase market concentration. Campbell et al. (2015) theoretically investigate the relationship between privacy protection and market structure. The results of the model suggest that the commonly used consent-based approach may disproportionately benefit generalist firms over specialist firms. Privacy regulation may be anticompetitive due to the nature of its transaction costs. This negative effect is strongest in industries with little price flexibility, such as the advertising-supported Internet.

Data are an input into innovation (Goldfarb & Tucker 2012a). In online advertising, health care, and a number of other fields, digital data generate better products and services and more efficient production. Therefore, restrictions on data flows will have an impact on the rate and direction of innovation. For instance, privacy protection of patients could discourage health-care IT adoption efforts and consequently lead to worse health outcomes (Adjerid et al. 2016, Derksen et al. 2021).

### 5.1.3. Mitigation of negative externalities.

As noted above, a growing literature examines how negative externalities may mean that even fully informed and rational consumers provide data to firms in excess of the welfare-maximizing amount (Choi et al. 2019, Acemoglu et al. 2022, Bergemann et al. 2022). Therefore, giving consumers control rights over their data, which is the spirit of many existing and proposed regulations, including Europe's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), is insufficient. However, outlawing the selling of data entirely would also be harmful to consumers (Jones & Tonetti 2020). The literature suggests a number of alternative tools. Speaking directly to the negative externality, Bergemann et al. (2022) call for consumer unions, at the segment level instead of the individual level, to internalize the data externality when bargaining with powerful digital platforms. Fainmesser et al. (2022) advocate a two-pronged regulatory policy, combining a minimal data protection requirement and a tax proportional to the data collected, to restore optimal efficiency. Jones & Tonetti (2020) emphasize the value of consumer property rights over data. Ali et al. (2023) highlight granting consumers granular control instead of an all-or-nothing form such as the opt-in option in GDPR. Consumers' selective disclosure of data can amplify competition or prompt a monopolist to lower price. Montes et al. (2019) suggest that regulators' focus should be on how information is transacted (e.g., the data contracts between data suppliers and users) and not on directly facilitating consumer privacy.

## 5.2. Empirical Impact of Regulation

A number of privacy regulations exist. The impact of these regulations has been examined across a variety of contexts. Goldfarb & Tucker (2011) examine an early digital privacy regulation, the EU ePrivacy Directive, which came into effect in 2004. The authors use the responses of 3 million survey takers who had been randomly exposed to nearly 10,000 online display advertising campaigns (i.e., banners) to explore how privacy regulation influenced the effectiveness of advertising. They document that following the ePrivacy Directive, banner ads experienced a reduction in effectiveness of over 65% in terms of changing the difference between treatment and control groups in stated purchase intent. Thus, the privacy regulation appears to have worked. Firms likely used less data, and therefore the ads became less effective. However, to the extent that advertising-supported software is an important industry, the regulation likely reduced the growth of that industry in Europe relative to the United States.

Many papers have focused on the impact of the GDPR. The empirical challenge in this work is that the GDPR was meant to have a global impact, and so there is no straightforward control group. As a consequence, the best papers in this stream use the heterogeneous impact of the regulation on different types of firms and different types of data flows to build a convincing argument. **Table 1** provides a summary of the various papers in the literature. Overall, the conclusion from these papers is that the GDPR led to an immediate reduction in web visits and revenue (Aridor et al. 2022, Goldberg et al. 2023) and a reduction in the efficiency of online search (Zhao et al. 2021). It also appears to have reduced the firms' ability to target advertising and track consumers (Godinho de Matos & Adjerid 2022, Peukert et al. 2022). Competition appears to have decreased in the online advertising market (Zhao et al. 2021, Johnson et al. 2023), and there was a decline in new firms, venture capital investment, and new apps (Jia et al. 2021, Janssen et al. 2022). In summary, the early evidence in the aftermath of the GDPR is that it worked, in the sense that firms were using less data in the year following the law's passing. This, however, had costs in terms of firm profits, the consumer online experience, innovation, and competition. There is some suggestive evidence that the impact has declined over time, with both less consumer protection and less impact on concentration (Johnson et al. 2023).

## 5.3. Non-Regulatory Privacy Protection

Firms have incentives to protect consumer privacy, even in the absence of regulation. For example, Jullien et al. (2020) investigate the equilibrium privacy policy of websites that generate revenue from charging the third parties on user information. Therefore, customer retention incentivizes the website to be mindful in its monetization efforts or to invest resources in screening third parties. Furthermore, a firm's privacy protection choice can work as a competition-mitigation strategy (Lee et al. 2011). While empirical work is still nascent, recent changes at Apple and Google are likely to lead to a richer understanding of firm incentives for privacy protection. Specifically, Apple and Google have restricted certain types of data flows from their devices and operating systems to third parties (Bergen 2021). Apple's App Tracking Transparency (ATT) feature asks users' permission to be tracked from advertisers for every app they download on their iPhone. Similarly, Google introduced its privacy initiative—Privacy Sandbox—against cookies. These restrictions protect customers from data-related harms from third parties (whether intrinsic or instrumental) but may have negative consequences on data-driven innovation at other firms. Measuring these effects remains an open question.

Privacy technology solutions may complement the regulatory process. All-or-nothing forms of consumer control—such as track/do-not-track options—need richer and more sophisticated technologies to benefit consumers (Ali et al. 2023). A great deal of engineering effort has gone into enabling data-driven innovation while restricting the flow of personally identifiable data. One development is the use of so-called differential privacy, which preserves anonymity in data while trying to ensure the data can be used for statistical analysis (see Dwork & Roth 2014 for a review and Abowd & Schmutte 2019 for an example in economics). Another development includes decentralized data management through a distributed ledger (Zyskind et al. 2015) or through anonymous transactions (Böhme et al. 2015). Innovations in privacy-preserving machine learning

### Table 1 Empirical evidence of GDPR's effects

| Authors | Main findings | Implications | Data setting |
|---|---|---|---|
| Jia et al. (2021) | The study finds negative short-term effects of the GDPR on investment in technology ventures. The effect is particularly pronounced in the period immediately after the GDPR's rollout and for newer, data-related, and consumer-facing ventures. | GDPR had a disproportionally negative impact on venture capital investment into technology firms. | Venture capital investment |
| Godinho de Matos & Adjerid (2022) | Consumer's consent for different data types improved when GDPR-compliant consent was obtained, leading to an increase in sales because of more effective targeted advertising. | GDPR may be effective for enhancing consumer privacy protection while at the same time enabling companies to improve products that rely on consumers personal data. | A large telecommunication provider with operations in Europe |
| Zhuo et al. (2021) | All estimates show economically small or zero effects of GDPR: the number of observed agreements, the agreement types, the number of observed interconnection points per agreement, the entry, and the observed number of customers of networks. | GDPR had no visible short-run impact on the Internet interconnection layer. | Internet interconnection |
| Peukert et al. (2022) | Websites reduce the number of third-party web technology providers they use, including websites not legally bound by the GDPR. The changes are disproportionally pronounced among less popular websites. | All firms experience losses. However, the vendor leader, Google, incurs relatively smaller losses and greatly expands its market share in crucial markets like advertising and analytics. | Web technology industry |
| Aridor et al. (2022) | The opt-in requirement of GDPR led to a 12.5% decrease in the consumer amount. However, the remaining consumers are trackable and predictable for a longer period of time. Their rising value to advertisers offsets part of the losses. | GDPR-enabled opt-out option increases the trackability of the opt-in consumers who choose to reveal their data, imposing an externality. | Online travel intermediary |
| Goldberg et al. (2023) | The study finds a reduction of approximately 12% in both website page views and e-commerce revenue among EU users, as recorded by the Adobe's analytics platform after the GDPR's enforcement deadline. | GDPR both reduced data recording and harmed real economic outcomes. | Adobe's website analytics platform |
| Janssen et al. (2022) | GDPR induced approximately one-third of the available apps to exit and decreased the entry rate of new apps in the market by half. | GDPR reduced beneficial innovation. | Apps on Google Play store |
| Johnson et al. (2023) | After GDPR's enforcement deadline, the website use of web technology vendors decreased 15% among EU residents. At the same time, the concentration of vendor market increased by 17%, since websites are more likely to drop smaller vendors. | GDPR increased market concentration among technology vendors in a business-to-business context. | Web technology vendors |
| Zhao et al. (2021) | GDPR impacts consumers' online browsing and search behavior; the authors find a panelist exposed to GDPR has 21.6% more search terms for information and 16.3% more pages browsed for goods and services access, indicating higher friction. | GDPR increased friction in online search; the increased friction is heterogeneous across firms, where smaller e-commerce firms are hurt more. | Consumer online browsing, app usage, and search activities |

GDPR, General Data Protection Regulation.

solutions are growing, in which consumer's privacy is protected while some valuable information can still be extracted in order to improve products and services (Sutanto et al. 2013, Zhou et al. 2020).

## 6. CONCLUSION

The increasing concern for privacy is directly related to the increasing use of digital data. The digital privacy literature in economics has focused on the costs and benefits of restricting data flows. Data flows are useful. They allow firms to provide consumers with the products and services they want, at the time they want them. Data flows also have negative consequences. Many consumers attribute an intrinsic value to privacy and so are intrinsically hurt by data flows. Data flows can also be used in ways that hurt consumers, and so there is an instrumental value to privacy.

The recent theory literature has emphasized positive and negative externalities from data flows. The empirical literature, however, has focused largely on the direct impact of regulation on consumers and firms so far. Looking forward, a key open question is the empirical relevance of the various theories on data externalities in determining the nature and consequences of privacy regulations and the strategic benefits to firms of proactively restricting data flows in the absence of regulation.

Both the theoretical and empirical literatures have hinted at digital competition as central to our understanding of digital privacy. The earlier theoretical work provided reasons to be optimistic that increased competition may generate a welfare-maximizing level of privacy. The more recent work on externalities suggests that competition may be insufficient. The empirical work on the consequences of the GDPR broadly suggests a reduction of competition. It is not clear, however, the degree to which this is a short-term phenomenon or is driven by idiosyncratic aspects of the GDPR as a privacy regulation. There remain a variety of open questions for both theoretical and empirical work with respect to how privacy regulation will affect competition and how competition (and competition policy) will affect consumer privacy.

To conclude, the economics literature has emphasized that both data flows and privacy have benefits to consumers and firms. Privacy is not free, but it is valuable. It affects economic outcomes. As governments consider new privacy regulations, and as firms develop privacy strategies, we hope that the perspective of economists—emphasizing costs, benefits, externalities, and competition—will be central to the discussion.

## DISCLOSURE STATEMENT

A.G. has served as an expert on matters of competition and privacy in the ad tech space. His full disclosure statement is available at **https://www.avigoldfarb.com/disclosure**. V.F.Q. has nothing to disclose.

## ACKNOWLEDGMENTS

We thank the Social Sciences and Humanities Research Council of Canada and the Alfred P. Sloan Foundation for support.

## LITERATURE CITED

Abowd JM, Schmutte IM. 2019. An economic analysis of privacy protection and statistical accuracy as social choices. *Am. Econ. Rev.* 109(1):171–202

Acemoglu D, Makhdoumi A, Malekian A, Ozdaglar A. 2022. Too much data: prices and inefficiencies in data markets. *Am. Econ. J. Microecon.* 14(4):218–56

Acquisti A, Brandimarte L, Loewenstein G. 2015. Privacy and human behavior in the age of information. *Science* 347(6221):509–14

Acquisti A, John LK, Loewenstein G. 2013. What is privacy worth? *J. Legal Stud.* 42(2):249–74

Acquisti A, Taylor C, Wagman L. 2016. The economics of privacy. *J. Econ. Lit.* 54(2):442–92

Acquisti A, Varian HR. 2005. Conditioning prices on purchase history. *Mark. Sci.* 24(3):367–81

Adjerid I, Acquisti A, Telang R, Padman R, Adler-Milstein J. 2016. The impact of privacy regulation and technology incentives: the case of health information exchanges. *Manag. Sci.* 62(4):1042–63

Ali SN, Lewis G, Vasserman S. 2023. Voluntary disclosure and personalized pricing. *Rev. Econ. Stud.* 90(2):538–71

Altman I. 1975. *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding*. Pacific Grove, CA: Brooks/Cole

Ansari A, Mela CF. 2016. E-customization. *J. Mark. Res.* 40(2):131–45

Aridor G, Che Y-K, Salz T. 2022. The effect of privacy regulation on the data industry: empirical evidence from GDPR. *RAND J. Econ.* In press

Ascarza E. 2018. Retention futility: Targeting high-risk customers might be ineffective. *J. Mark. Res.* 55(1):80–98

Athey S, Catalini C, Tucker C. 2017. *The digital privacy paradox: small money, small costs, small talk*. NBER Work. Pap. 23488

Athey S, Gans JS. 2010. The impact of targeting technology on advertising markets and media competition. *Am. Econ. Rev.* 100(2):608–13

Bajari P, Chernozhukov V, Hortaçsu A, Suzuki J. 2019. The impact of big data on firm performance: an empirical investigation. *AEA Pap. Proc.* 109:33–37

Baye MR, Sappington DEM. 2020. Revealing transactions data to third parties: implications of privacy regimes for welfare in online markets. *J. Econ. Manag. Strategy* 29(2):260–75

Becker GS. 1980. Privacy and malfeasance: a comment. *J. Legal Stud.* 9(4):823–26

Bénabou R, Tirole J. 2006. Incentives and prosocial behavior. *Am. Econ. Rev.* 96(5):1652–78

Bergemann D, Bonatti A. 2011. Targeting in advertising markets: implications for offline versus online media. *RAND J. Econ.* 42(3):417–43

Bergemann D, Bonatti A. 2015. Selling cookies. *Am. Econ. J. Microecon.* 7(3):259–94

Bergemann D, Bonatti A. 2019. Markets for information: An introduction. *Annu. Rev. Econ.* 11:85–107

Bergemann D, Bonatti A, Gan T. 2022. The economics of social data. *RAND J. Econ.* 53(2):263–96

Bergemann D, Bonatti A, Smolin A. 2018. The design and price of information. *Am. Econ. Rev.* 108(1):1–48

Bergen M. 2021. Apple and Google are killing the (ad) cookie. Here's why. *Bloomberg*, April 26. **https://www.bloomberg.com/news/articles/2021-04-26/how-apple-google-are-killing-the-advertising-cookie-quicktake#xj4y7vzkg**

Bleier A, Goldfarb A, Tucker C. 2020. Consumer privacy and the future of data-based innovation and marketing. *Int. J. Res. Mark.* 37(3):466–80

Böhme R, Christin N, Edelman B, Moore T. 2015. Bitcoin: economics, technology, and governance. *J. Econ. Perspect.* 29(2):213–38

Brynjolfsson E, McElheran K. 2016. The rapid adoption of data-driven decision-making. *Am. Econ. Rev.* 106(5):133–39

Burtch G, Ghose A, Wattal S. 2015. The hidden cost of accommodating crowdfunder privacy preferences: a randomized field experiment. *Manag. Sci.* 61(5):949–62

Campbell J, Goldfarb A, Tucker C. 2015. Privacy regulation and market structure. *J. Econ. Manag. Strategy* 24(1):47–73

Casadesus-Masanell R, Hervas-Drane A. 2015. Competing with privacy. *Manag. Sci.* 61(1):229–46

Chan T, Hamdi N, Hui X, Jiang Z. 2022. The value of verified employment data for consumer lending: evidence from Equifax. *Mark. Sci.* 41(4):367–86

Chen L, Huang Y, Ouyang S, Xiong W. 2021. *The data privacy paradox and digital demand*. NBER Work. Pap. 28854

Chen Y, Iyer G. 2002. Research note: consumer addressability and customized pricing. *Mark. Sci.* 21(2):197–208

Chiou L, Tucker C. 2017. *Search engines and data retention: implications for privacy and antitrust*. NBER Work. Pap. 23815

Choe C, King S, Matsushima N. 2018. Pricing with cookies: behavior-based price discrimination and spatial competition. *Manag. Sci.* 64(12):5669–87

Choi JP, Jeon D-S, Kim B-C. 2019. Privacy and personal data collection with information externalities. *J. Public Econ.* 173:113–24

Choudhary V, Ghose A, Mukhopadhyay T, Rajan U. 2005. Personalized pricing and quality differentiation. *Manag. Sci.* 51(7):1120–30

Coase RH. 1972. Durability and monopoly. *J. Law Econ.* 15(1):143–49

Conitzer V, Taylor CR, Wagman L. 2012. Hide and seek: costly consumer privacy in a market with repeat purchases. *Mark. Sci.* 31(2):277–92

Daughety AF, Reinganum JF. 2010. Public goods, social pressure, and the choice between privacy and publicity. *Am. Econ. J. Microecon.* 2(2):191–221

De Corniere A, De Nijs R. 2016. Online advertising and privacy. *RAND J. Econ.* 47(1):48–72

Derksen L, McGahan A, Pongeluppe L. 2021. *Privacy at what cost? Using electronic medical records to recover lapsed patients into HIV care*. Work. Pap., Univ. Toronto, Toronto, Can.

Dwork C, Roth A. 2014. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.* 9(3–4):211–407

Erlich Y, Shor T, Pe'er I, Carmi S. 2018. Identity inference of genomic data using long-range familial searches. *Science* 362(6415):690–94

Fainmesser IP, Galeotti A, Momot R. 2022. Digital privacy. *Manag. Sci.* In press. **https://doi.org/10.1287/mnsc.2022.4513**

Farboodi M, Mihet R, Philippon T, Veldkamp L. 2019. Big data and firm dynamics. *AEA Pap. Proc.* 109:38–42

Farboodi M, Veldkamp L. 2022. *A model of the data economy*. Work. Pap., Columbia Univ., New York

Godinho de Matos M, Adjerid I. 2022. Consumer consent and firm targeting after GDPR: the case of a large telecom provider. *Manag. Sci.* 68(5):3330–78

Goh K-Y, Hui K-L, Png IPL. 2015. Privacy and marketing externalities: evidence from do not call. *Manag. Sci.* 61(12):2982–3000

Goldberg SG, Johnson GA, Shriver SK. 2023. Regulating privacy online: an economic evaluation of the GDPR. *Am. Econ. J. Econ. Policy.* In press

Goldfarb A. 2014. What is different about online advertising? *Rev. Ind. Organ.* 44(2):115–29

Goldfarb A, Tucker C. 2011. Privacy regulation and online advertising. *Manag. Sci.* 57(1):57–71

Goldfarb A, Tucker C. 2012a. Privacy and innovation. *Innov. Policy Econ.* 12(1):65–90

Goldfarb A, Tucker C. 2012b. Shifts in privacy concerns. *Am. Econ. Rev.* 102(3):349–53

Goldfarb A, Tucker C. 2019. Digital economics. *J. Econ. Lit.* 57(1):3–43

Hann I-H, Hui K-L, Lee S-YT, Png IPL. 2008. Consumer privacy and marketing avoidance: a static model. *Manag. Sci.* 54(6):1094–103

Hermalin BE, Katz ML. 2006. Privacy, property rights and efficiency: the economics of privacy as secrecy. *Quant. Mark. Econ.* 4(3):209–39

Ichihashi S. 2020. Online privacy and information disclosure by consumers. *Am. Econ. Rev.* 110(2):569–95

Ichihashi S. 2021. The economics of data externalities. *J. Econ. Theory* 196:105316

Iyer G, Soberman D, Villas-Boas JM. 2005. The targeting of advertising. *Mark. Sci.* 24(3):461–76

Janssen R, Kesler R, Kummer ME, Waldfogel J. 2022. *GDPR and the lost generation of innovative apps*. NBER Work. Pap. 30028

Jia J, Jin GZ, Wagman L. 2021. The short-run effects of the general data protection regulation on technology venture investment. *Mark. Sci.* 40(4):661–84

Jin Y, Vasserman S. 2021. *Buying data from consumers: the impact of monitoring programs in U.S. auto insurance*. NBER Work. Pap. 29096

Johnson GA, Shriver SK, Goldberg SG. 2023. Privacy and market concentration: intended and unintended consequences of the GDPR. *Manag. Sci.* In press

Johnson JP. 2013. Targeted advertising and advertising avoidance. *RAND J. Econ.* 44(1):128–44

Jones CI, Tonetti C. 2020. Nonrivalry and the economics of data. *Am. Econ. Rev.* 110(9):2819–58

Jullien B, Lefouili Y, Riordan MH. 2020. *Privacy protection, security, and consumer retention*. Unpublished manuscript, Toulouse Sch. Econ., Toulouse, Fr.

Kummer M, Schulte P. 2019. When private information settles the bill: money and privacy in Google's market for smartphone applications. *Manag. Sci.* 65(8):3470–94

Lee D-J, Ahn J-H, Bang Y. 2011. Managing consumer privacy concerns in personalization: a strategic analysis of privacy protection. *MIS Q.* 35(2):423–44

Lin T. 2022. Valuing intrinsic and instrumental preferences for privacy. *Mark. Sci.* 41(4):663–81

Loertscher S, Marx LM. 2020. Digital monopolies: privacy protection or price regulation? *Int. J. Ind. Organ.* 71:102623

Miklós-Thal J, Tucker CE. 2019. Collusion by algorithm: Does better demand prediction facilitate coordination between sellers? *Manag. Sci.* 65(4):1552–61

Miller AR, Tucker CE. 2011. Can health care information technology save babies? *J. Political Econ.* 119(2):289–324

Miller AR, Tucker CE. 2018. Privacy protection, personalized medicine, and genetic testing. *Manag. Sci.* 64(10):4648–68

Montes R, Sand-Zantman W, Valletti T. 2019. The value of personal information in online markets with endogenous privacy. *Manag. Sci.* 65(3):1342–62

Neumann N, Tucker CE, Whitfield T. 2019. Frontiers: How effective is third-party consumer profiling? Evidence from field studies. *Mark. Sci.* 38(6):918–26

Nissenbaum H. 2004. Privacy as contextual integrity. *Wash. Law Rev.* 79(1):119–57

Nissenbaum H. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA: Stanford Univ. Press

Norberg PA, Horne DR, Horne DA. 2007. The privacy paradox: personal information disclosure intentions versus behaviors. *J. Consum. Aff.* 41(1):100–26

Peukert C, Bechtold S, Batikas M, Kretschmer T. 2022. Regulatory spillovers and data governance: evidence from the GDPR. *Mark. Sci.* 41(4):746–68

Posner RA. 1981. The economics of privacy. *Am. Econ. Rev.* 71(2):405–9

Rafieian O, Yoganarasimhan H. 2021. Targeting and privacy in mobile advertising. *Mark. Sci.* 40(2):193–218

Reimers I, Shiller BR. 2019. The impacts of telematics on competition and consumer behavior in insurance. *J. Law Econ.* 62(4):613–32

Schneier B. 2015. *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. New York: W.W. Norton & Co.

Shapiro C, Varian HR. 1998. *Information Rules: A Strategic Guide to the Network Economy*. Cambridge, MA: Harvard Bus. Sch. Press

Shy O, Stenbacka R. 2016. Customer privacy and competition. *J. Econ. Manag. Strategy* 25(3):539–62

Smith MD, Bailey J, Brynjolfsson E. 2001. Understanding digital markets. In *Understanding the Digital Economy: Data, Tools, and Research*, ed. E Brynjolfsson, B Kahin, pp. 99–136. Cambridge, MA: MIT Press

Soleymanian M, Weinberg CB, Zhu T. 2019. Sensor data and behavioral tracking: Does usage-based auto insurance benefit drivers? *Mark. Sci.* 38(1):21–43

Solove DJ. 2008. *Understanding Privacy*. Cambridge, MA: Harvard Univ. Press

Solove DJ. 2021. The myth of the privacy paradox. *George Wash. Law Rev.* 89(1):1–51

Stigler GJ. 1980. An introduction to privacy in economics and politics. *J. Legal Stud.* 9(4):623–44

Sun T, Yuan Z, Li C, Zhang K, Xu J. 2021. *The value of personal data in Internet commerce: a high-stake field experiment on data regulation policy*. NET Inst. Work. Pap. 21-10, Net Inst., New York

Sutanto J, Palme E, Tan C-H, Phang CW. 2013. Addressing the personalization-privacy paradox: an empirical assessment from a field experiment on smartphone users. *MIS Q.* 37(4):1141–64

Tamir DI, Mitchell JP. 2012. Disclosing information about the self is intrinsically rewarding. *PNAS* 109(21):8038–43

Tang H. 2019. *The value of privacy: evidence from online borrowers*. Tech. Rep., HEC Paris, Paris

Taylor CR. 2004. Consumer privacy and the market for customer information. *RAND J. Econ.* 35(4):631–50

Tirole J. 2021. Digital dystopia. *Am. Econ. Rev.* 111(6):2007–48

Tucker CE. 2014. Social networks, personalized advertising, and privacy controls. *J. Mark. Res.* 51(5):546–62

Tucker CE. 2018. Privacy, algorithms, and artificial intelligence. In *The Economics of Artificial Intelligence: An Agenda*, ed. A Agrawal, J Gans, A Goldfarb, pp. 423–37. Chicago: Univ. Chicago Press

Veldkamp L, Chung C. 2019. Data and the aggregate economy. *J. Econ. Lit.* In press

Villas-Boas JM. 1999. Dynamic competition with customer recognition. *RAND J. Econ.* 30(4):604–31

Villas-Boas JM. 2004. Price cycles in markets with customer recognition. *RAND J. Econ.* 35(3):486–501

Warren SD, Brandeis LD. 1890. Right to privacy. *Harvard Law Rev*. 4:193–220

Westin AF. 1968. Privacy and freedom. *Wash. Lee Law Rev.* 25(1):166

Yang KH. 2022. Selling consumer data for profit: optimal market-segmentation design and its consequences. *Am. Econ. Rev.* 112(4):1364–93

Yoganarasimhan H. 2020. Search personalization using machine learning. *Manag. Sci.* 66(3):1045–70

Zhang J, Krishnamurthi L. 2004. Customizing promotions in online stores. *Mark. Sci.* 23(4):561–78

Zhao Y, Yildirim P, Chintagunta PK. 2021. *Privacy regulations and online search friction: evidence from GDPR*. Unpublished manuscript, Wharton Sch., Univ. Pa., Philadelphia

Zhou Y, Lu S, Ding M. 2020. Contour-as-face framework: a method to preserve privacy and perception. *J. Mark. Res.* 57(4):617–39

Zhuo R, Huffaker B, Claffy KC, Greenstein S. 2021. The impact of the General Data Protection Regulation on internet interconnection. *Telecommun. Policy* 45(2):102083

Zyskind G, Nathan O, Pentland A. 2015. Decentralizing privacy: using blockchain to protect personal data. In *2015 IEEE Security and Privacy Workshops*, pp. 180–84. New York: IEEE