Privacy, Economics, and Regulation: A Note[1]

Alessandro Acquisti[2]
Carnegie Mellon University

Prepared for the Atlanta Fed's Financial Markets Conference

Draft, May 2019

1. Introduction

In the past two decades, far-reaching innovation in information technology has brought forward a lively scholarly and policy debate over the benefits accrued from the collection of consumer data and the costs and concerns associated with its possible misuse. Costs, benefits, and trade-offs are the domain of economists' expertise and research. It is therefore not surprising that the academic field of privacy economics has experienced intense growth over the same period of time. Articles that employ economic methodologies to understand the diverse trade-offs associated with personal data have become common across a diverse set of outlets: traditional economic journals; marketing, law, and information systems venues; and computer science conference proceedings.

While information technologies keep evolving, the field of the economics of privacy is not new. Its origins predate the commercial explosion of the Internet and the rise to dominance of digital technologies in our personal and professional lives. Already at the cusp between 1970s and 1980s, some economists were starting to consider the economic value of personal information and the economic consequences of its protection. Scholars such as Posner (1977, 1981), Stigler (1980) and Hirshleifer (1980) began looking at the economic incentives that may drive individuals to share, or protect, their personal information. A decade or so later, information economists and information systems scholars such as Varian, Noam, and Laudon started proposing economic interpretations vis-à-vis technologies for personal data sharing and protection (such as the

[1] This policy Note is partly based on Acquisti et al. (2016).
[2] Email: acquisti@andrew.cmu.edu. Please contact the author for latest version of this draft.

digitization of documents; encryption; and data warehouses). Varian (1997) pointed out the rational desire of consumers to share some information with marketers while protecting other information; Noam (1997) observed how Coasian arguments around the efficient allocation of resources and outputs in competitive markets may also apply to personal data, regardless of how rights over the protection or sharing of that data may be allocated to different stakeholders; and Laudon (1996) first proposed the notion of personal data markets, where consumers may trade rights over the collection and usage of their data.

While these – and many other – examples (which are covered in more extensive detail in Section 3) do not always explicitly mention the concept of *regulation* of personal information, they all ultimately have something valuable to say about how regulatory interventions into the way markets collect, share, and use data will impact the economic welfares of individuals and society at large. For Posner and Stigler, regulatory intervention in the privacy field may be at best unnecessary and at worst damaging. For Hirshleifer, policy-making efforts may be required to avoid wasteful over-investment in data collection. For Laudon, a clear and explicit framework of property rights may be necessary to stimulate markets for privacy and for personal data. In effect, one could say that the entire field of the economics of privacy, by virtue of considering the trade-offs associated with the collection/sharing or the protection/hiding of personal data, ends up implicitly being about (or, at least, having something to say on the topic of) the economic impact of privacy regulation, since regulation is often the fundamental way through which data protection can be achieved.

It is not a surprise, therefore, that as the privacy economics field started greatly expanding in the 2000s (after the advent of the commercial Internet, which brought about new channels for data collection, disclosure, and analysis), a substantial portion of the field began to focus *explicitly* on the economic impact of privacy laws. The goal of this Note is to provide a brief overview of that literature, as well as to offer a lens for understanding this space. The Note focuses on distilling and highlighting critical issues rather than exhaustively categorizing all related work. This space has become quite large, and remains complex and nuanced for multiple reasons. One is that privacy is an ambiguous concept that means something different to different people (including scholars; see Solove 2005). For the most part, this manuscript will focus on the informational dimension of privacy, interpreted as (some degree of) control over flows of personal data. Even so, when

scholars use economics to study privacy, they may actually be studying very different things – from identity theft to targeted advertising; from online price discrimination to electronic medical records and Health Information Exchanges. Another reason is (as will be discussed in Section 2) that not only can the term "regulation" in the context of privacy mean many different things, but regulation itself is not the only means of privacy protection – other ones being technology, self-regulation, consumer responsibility, and so forth. Thus, it is often difficult (and sometimes impossible) to provide generalizable conclusions about the merits or risks of policy-making interventions in the privacy field, as those interventions could take many different forms, each of them associated with slightly different, and nuanced, downstream consequences. For these reasons this Note explicitly does not attempt to provide a one-size-fits-all conclusion about the impact of privacy regulation on economic metrics such as welfare, efficiency, or innovation. In fact, it is hard to believe there is a single correct answer to that question.

The rest of this Note proceeds as follows. After this introduction, it presents a series of considerations and caveats worth keeping in mind when trying to interpret and assess the impact of privacy regulation (Section 2). It then surveys related work, starting with theoretically focused studies (Section 3) and moving on to empirical work (Section 4). Both Sections 3 and 4 will cover examples from a wide array of sub-fields and data-relevant scenarios – such as medical data, online tracking, or the impact of recent regulatory initiatives such as GDPR. Finally, the Note highlights some open issues and questions (Section 5). As is perhaps inevitable in this nuanced and evolving space, the Note concludes that the impact of privacy regulation is similarly nuanced and context-dependent. Claims that regulatory interventions aimed at protecting privacy by shielding data from collection or usage are always and univocally welfare-decreasing are probably incorrect; and claims that the online economy (which is ad-supported and increasingly reliant on the collection of personal data) would be in peril if privacy gets (more) regulated are probably also overblown. At the same time, regulation does create waves of changes, can impose new costs to firms, can decrease the value (and utility) of data, and can affect the balance of power among different data stakeholders (such as data "subjects" – that is, the subjects of the data; and data "holders" – that is, the organization collecting that data). Privacy is inherently redistributive. Whether those costs are acceptable in light of the importance of privacy is not just an economic question but also a question about a society's values and beliefs. This Note cannot resolve that question, but hopefully can provide some inputs on how to think about it.

2. Caveats by way of a framework

This section focuses on a few points to consider when thinking about the economic impact of privacy regulation. It is not a framework per se, in that this section does not propose a systematic scheme of analysis. Rather, it presents a series of considerations, or caveats, that may prove useful or even important to appreciate the nuance and complexity of the consequences of privacy protection.

The first consideration is that it may prove useful in thinking about this space to separate two questions: To what degree should consumer privacy be protected? And, how do we achieve that degree of protection? Separating those questions can be useful because, while they are obviously related (and in fact, the reasoning and rationale necessary to address one often also apply to the other), confusion and misunderstanding arise when people who discuss privacy protection do not realize they may be talking over each other because they are actually addressing different questions. The first question is about understanding individual and societal trade-offs associated with protection or disclosure and collection of private data: What are those trade-offs? How do we calculate them? And so on. The second question is about implementation: Is there evidence that an optimal (or even just desirable) balance between collection and protection is currently already achieved in our economies and societies? If not, what mechanisms or approaches do we need to implement to steer markets in that direction?

Another reason why it may prove valuable to separate the two questions is that doing so almost immediately provides evidence for a point made at the end of Section 1: how profoundly nuanced, complex – and ultimately even un-answerable in purely economic terms – the question of the "right" amount of privacy (regulation) turns out to be.

To explain that claim, I will start from the "how much" question (to what degree should consumer privacy be protected). Essentially, this question is about understanding and measuring individual and societal trade-offs associated with protection or disclosure and collection of private data. Such understanding and measuring, however, presents a number of unique challenges.

A first challenge is defining the precise (economic) objective function the social planner is supposed to maximize: Aggregate economic welfare? Economic welfare of specific stakeholders, such as data subjects? Economic efficiency? Technological innovation? Economic growth? Equity

and economic equality, broadly defined? Each objective function may provide different conclusions regarding where to draw the balance between collection and protection of personal information. For instance, limits on the secondary use of medical data without individuals' explicit consent may reduce the likelihood of some patients with particular traits or socio-economic backgrounds being taken advantage of by unscrupulous providers; but may also reduce the ability of researchers to combine medical data from different databases to uncover promising novel health treatments, which may beneficially impact society at large.

A second challenge is that, even if the objective function was narrowly defined as (the increase of) aggregate welfare, both data collection and data protection are inherently redistributive, which raises the issue of privacy regulation (as well as, importantly, lack thereof) as, inevitably, being a tool of economic redistribution. As Section 3 will consider, the flows of data (as well as barriers to such flows) inevitably create economic winners and losers, because the economic interests of different stakeholders are rarely aligned when it comes to data incentives. Consider the prototypical example of the sales person trying to infer the reservation price of the buyer: if regulation prohibits the seller from scouring databases to find the prospective buyer's income bracket, the seller may not be able to push the price for the product or service he is selling close to the consumer's reservation price; thus a higher amount of surplus (relative to the case where the seller acquires precise willingness-to-pay information) from the resulting transaction will remain in the pocket of the buyer, and less in the pocket of the seller. Vice versa, if no regulation impedes the sale person from seeking and using the prospective buyer's personal data, the seller may be able to get closer and closer to the buyer's theoretical reservation price – pocketing away the entire surplus from the transaction. In short, different decisions on data protection will affect different stakeholders differently. Not only does maximizing aggregate welfare imply different welfare implications for different stakeholders (for instance, data subjects versus data holders), even the theoretical goal of achieving Pareto-optimal societal outcomes will inevitably force policy makers to face thorny questions regarding whose welfare they want to prioritize (after all, by way of analogy, both perfect competition and monopoly with first-degree price discrimination are Pareto-efficient, and yet they reflect opposite extremes of surplus distribution in an economy). In other words: even if we tried to be neutral social planners, trying to maximize objective welfare when choosing to what extent (and if at all) to protect individual privacy, we would soon realize that there is often no single correct solution to the problem.

Economics could provide us with guidance, but the issue would, ultimately, remain one of social policy (and politics).

A third challenge in answering the "how much" question is that both the protection and the disclosure of consumer data often create long-term effects that are hard or even impossible to precisely capture by rigorous economic analysis – the type of analysis that can identify causal relationships robust enough to be publishable in premier academic economic journals. For instance: regulation that limits services' ability to track online data and use it for targeted advertising may protect privacy but reduce advertising effectiveness (Section 4 considers such a scenario). In turn, in the short run, reduced advertising effectiveness may reduce merchants' sales and consumer spending. But in the long run, other diverse second-order effects may emerge, with different implications for different stakeholders. And reasonable theoretical arguments could lead to very opposite predictions about the ultimate downstream outcomes for those stakeholders – all of those outcomes being plausible based on traditional economic reasoning, but few of them being demonstrable on empirical grounds, given the challenges of causally linking these complex chains of events. For example, reduced targeting precision may lead to lower bids by advertisers in auctions for ad space to show their ads to certain audiences, and thus lower revenues for online publishers (the sites displaying the ads), and in turn reduced ability for the publishers to produce novel quality content. Or, the opposite may occur: reduced targeting precision may actually lead to *higher* bids by advertisers in ads auctions - because, when consumers cannot be hyper-precisely targeted, the number of different merchants interested in showing their ads to a given consumer may actually increase, relative to the case where hyper-targeting de facto restricts the set of interested merchants to a small group. As bids increase in value, revenues for online publishers increase, and so does publishers' ability to invest in new content. As a matter of fact, one could construct similarly contrasting scenarios for other stakeholders – such as consumers. Facing tracking and targeting restrictions, consumer search costs will rise (as ads will play a less than desirable role in matching consumers with products they like); this will depress sales. Or, in the long run, consumer confidence in browsing and buying online will increase (due to the feeling of protection afforded by the privacy regulation); sales will increase. In essence, the challenge is that both regulatory interventions in the handling of consumer data and the absence of said interventions have not just direct (and sometimes measurable) economic effects, but are likely to create many more downstream (and much harder to measure) economic effects. Unfortunately,

when it comes to privacy economics, many of us trying to investigate those effects are like the economist in the well-known joke who looks for her lost keys at night under the streetlamp not because that is where the keys were lost, but merely because that is where the light is.

A fourth challenge in addressing the "how much" question is that the consequences of privacy invasions (and, accordingly, the expected reduction in harm via privacy protection) take many different forms because – as discussed – many and diverse are the scenarios and markets in which various categories of personal information play an economic role. Thus, economic consequences may vary in terms of typology of harm (from financial loss associated with price discrimination to attention waste associated with unwanted advertising); in magnitude (from negligible nuisances associated with spam messages to catastrophic consequences of medical identity theft); and in likelihood (from low probability events, such as loss of employment following regrettable public disclosures on social media; to high frequency events, such as the opportunity costs associated with very slight delays in the loading of most modern webpages due to the presence, in their html codes, of numerous third-party trackers[3] capturing data about each visitor and exchanging it with other servers). The challenge here is that, in isolation, each of those typologies of harm is notionally measurable (at least in expected terms – that is, treating the harm as stochastic and estimating its impact across a population over time). But if the goal was to estimate the *overall* economic impact of broad, general-purpose (as opposed to specific and sectorial) privacy regulation (think the European General Data Protection Regulation, which applies broadly to consumers' data, as opposed to – say – the U.S. Video Privacy Protection Act, which applies specifically to video rentals), one would have to combine all of these heterogeneous typologies of harms, with their different profiles, likelihoods, and magnitudes, into some unified aggregate function. Clearly, opportunities for estimation noise, bias, and errors would abound in such patchwork.

A fifth challenge is that privacy regulation (or lack thereof) does not merely create economic consequences. There are numerous dimensions of privacy that are intangible, or even immeasurable, and yet no less important: the psychological harms associated with privacy invasion (Calo 2011); or the role of privacy in protecting individual autonomy (Cohen 2012), freedom

[3] See https://medium.freecodecamp.org/what-you-should-know-about-web-tracking-and-how-it-affects-your-online-privacy-42935355525.

(Westin 1968), and dignity (Schoeman 1992). How do we account for those dimensions in our economic assessment of the consequences of privacy regulation, or lack thereof?

Considering all the challenges just listed, one could even argue that much of the impact of privacy is "economic dark matter": we think it's there, but it's hard to capture or measure. As for the second question (the "how" question: Once we have set a certain target of privacy protection, how do we achieve it?), answering it is, no doubt, marred by methodological challenges as well. On the upside, however, we can abstract from the first question (the "how much" question) and simply assume that, via economic analysis, policy debate, elections, or other ethical discussions, we have determined some societal objective in terms of privacy protection. Then, the next step is to consider whether market forces alone have, already, realized and ensured that degree of protection, or whether some form of intervention is necessary. If so, we can then debate over the relative likelihood of different interventions to achieve the professed goal: from soft paternalistic nudges to incentives to firms in a fundamentally self-regulatory framework; from regulation focused on notice and consent mechanisms for consumers to stricter interventions requiring privacy by design or privacy by default. Section 5 will discuss how a combination of economics, behavioral economics, and information technology can help cast at least some light over the relative advantages and disadvantages of these different interventions.

3. Theoretical work

This section considers some foundational work on modeling the economic impact of the protection of personal data. Early seminal stream of work, dating back to the 1980s, identified in privacy protection the source of market distortions and economic inefficiencies. However, later work painted a more nuanced picture of privacy trade-offs, including the fact that unfettered collection and access to individual data is unlikely to maximize consumer, or even aggregate, welfare.

Intuitively, when privacy is construed as the "hiding" of consumer information which can negatively affect the economic interests of other parties, its effects on market transactions may be undesirable: the more information is eliminated from the market, the more distortions arise (as markets cannot play their natural matchmaking role between the incentives and interests of different stakeholders), leading to loss of economic efficiency and generally sub-optimal economic

equilibria. This, in a nutshell, was the position of seminal thinkers from the University of Chicago such as Posner and Stigler, who pioneered the economic analysis of privacy.

Posner (1977, 1981) used the example of a job seeker, who may have an incentive to hide or misrepresent certain parts of her background to job providers. If protection of information related to the job seeker that is valuable to the job provider is permitted (for instance, via employment privacy laws), this may reduce the ability of a hiring firm to choose the right candidate for a position. In essence, the protection of the job seeker's privacy diminishes the quality of the job market matching and imposes costs to the firm. Stigler (1980) pushes this argument further, by noting that individuals with favorable information may in fact want to share this information with others; not so individuals with unfavorable information. Consider the case of credit worthiness, and the role of diligence in paying back past debts: an individual who has been diligent in paying back her debts may indeed want her credit history to be known to potential lenders, whereas an individual who has a history of defaulting may want that information to stay hidden. So, again, regulatory intervention allowing the protection of certain types of data would be redistributive (it would cause downstream costs to the economic agents who do business with the subject without having sufficient information about her); but it would also be ultimately ineffective, in that individuals with favorable information may still make an effort to disclose it; from this, other parties may infer that an individual protecting her information is doing so because she is hiding *negative* information.

An updated version of Stigler's argument can be found, some years later, in a paper by Noam (1997). Noam argues that, ultimately, the sharing or collection of an individual's data will not depend on whether privacy regulation has, or has not, mandated the protection of data. The initial allocation of those rights is, in Coasian terms, irrelevant because the actual market outcome will be a function of the relative valuations of the different agents who have an interest in accessing the information or keeping it protected. For instance, assume that privacy regulation makes it illegal for Amazon to collect data from its Echo devices. If Amazon values that information more than the consumers value its protection, Amazon would engage in some form of contract negotiation with consumers, offering some payment to consumers to agree to trade away their privacy right and, in fact, allow data collection. Both Amazon and consumers would be better off (as long as the price they agree upon is less than the value Amazon expects to extract from the

data, and more than the privacy costs the consumer expects to bear from its release), and the ultimate outcome will be that the data is released (notwithstanding the initial regulatory assignment of rights, which gave consumers a right to their privacy). Consider, now, a situation where there is not such regulatory protection of data, but consumers do value their privacy more than Amazon values their data. Then, contract negotiations between Amazon and its consumers will lead to mutually beneficial outcomes where consumers "buy back" protection of their data from Amazon, and Amazon agrees not to collect data, in exchange for money.

Clearly, in these two different scenarios, we face two very different outcomes in terms of allocation of surplus (which party benefits the most): the assignment of initial rights does have an impact on the allocation of value but, in Noam's reasoning, does not ultimately impact whether privacy will be protected or not. It is the agents' valuations that determine the latter. Noam however also states that the usual caveats apply in the context of such Coasian reasoning: transaction costs, or asymmetric information, may render such data negotiation tricky or in fact unfeasible.

A deeper analysis of the literature will show that the economic implications of privacy protection are even more nuanced than what economists may have originally believed. Whereas Posner and Stigler would conceive of privacy in binary terms (individuals have "negative" and "positive" information; they want to share the positive but protect the negative, thus causing externalities on other agents), Varian (1997) points out that consumers may rationally want some information about themselves to be known, but not other information – and the distinction may not be due to the information itself being intrinsically "positive" or "negative." Varian considers the case of telemarketers, who (used to, and sometimes still do) call individuals at home with various offers of products and services. Paradoxically, Varian points out, the annoyance of telemarketing is not due to lack of privacy (the telemarketer "invading" the privacy of a person's home by calling at unexpected or even undesirable times); instead, the problem lies in the telemarketer not knowing *enough* about the consumer preferences and traits, and therefore calling with offers for products and services the consumer may have very little interest in. This causes a nuisance. If, however, the telemarketers knew the consumer's type (that is, her preferences), it would be rational for them to use their calls to target offers for products and services consumers may actually want – making the transaction mutually beneficial, and removing the privacy concern. So far, this approach may sound like a revised and updated version of the Chicago School

arguments outlined earlier. However, Varian continues by pointing out that, while a consumer may want to rationally share information about her traits and preferences with marketers (so that she can get offers for products she is actually interested in), the same consumer may also (and quite rationally) want to hide from the marketer *how much* she is interested in those products – that is, the reservation price (or maximum willingness to pay) for products. If the telemarketer also had that information, it would not just contact consumers with offers they like, but would also be able to price its offers at levels that consumers are willing to accept. This would lead to extraction of consumer surplus from the buyer to the seller – as in theoretical models of first-degree price discrimination. In sum, Varian's point is that it is quite rational that consumers may want some information about themselves to be shared with others, and other information to be protected. Similarly to the Chicago School's approach, sharing some information is desirable (and prohibition against telemarketers collecting consumer data would be ultimately undesirable for the market). Unlike that approach, however, under Varian's line of reasoning protecting information does not mean hiding negative information; and does not necessarily create market distortions. And yet, it is worthwhile to point out that even under this notion both the protection or the sharing of personal information will have (re)distributive effects: if telemarketers can use databases to infer consumer income and willingness to pay, they will more likely be able to engage in price discrimination, and more of the surplus originating from market transactions will go to them (the sellers) rather than the buyers. If, instead, regulation was to pose limits and constraints to telemarketers' ability to so minutely target consumers, it is possible that more surplus from their transactions with consumers would remain with the latter. (Note that we are focusing here on the first-order effects of price discrimination; for a discussion of second-order effects, see Varian 1989).

It is interesting to point out that Varian's paper was published in the very early days of the commercial Internet. While it did not mention the latter (the paper was about the old school technology of marketing via landline phone calls), its arguments seem very relevant to the domain that emerged a few years later: online advertising and, in particular, targeted advertising. There is discussion later in this section of some of the recent theoretical literature on the economic implications of targeted advertising and its regulation or prohibition.

In Varian, we already find a subtle departure from the general claim that the protection of data has univocally distortive effects on economic equilibria. Some authors departed from that

notion even more forcefully. For instance, it was Hirshleifer (1971, 1980) who showed how the private benefit of information acquisition may outweigh its social benefit, leading potential data holders to over-invest in collecting information about data subjects. Murphy (1995) and Daughety and Reinganum (2010) show how, in the absence of privacy protection, individuals may choose actions to manipulate their public reputations, whereas under a privacy regime they would have chosen optimal levels of a certain behavior. Thus, depending on the context, keeping certain behaviors private (discussions with doctors; checking into rehab) would increase the amount of actions that benefit both individual and societal welfare (feeling comfortable sharing sensitive and accurate information with a physician; seeking treatment for alcohol addiction). And Taylor (2004) shows that, when sellers can infer consumers' preferences and willingness to pay for products, if consumers are not sufficiently sophisticated to anticipate price discrimination, in equilibrium their economic surplus will be captured by the seller. Thus, under that scenario, privacy regulation will protect consumer surplus. As a matter of fact, we collected in Acquisti et al. (2016) a series of modeling-grounded scenarios highlighting how, contrary to the conventional Chicago School wisdom, the unfettered collection of consumer data will not just decrease data subjects' welfare, but will also decrease aggregate welfare.

An interesting new arena of application of these economic models (and, as we will see in the next Section, of their empirical validation) relates to the value of online (targeted) advertising. With the advent of the Internet, new consumer tracking and targeting technologies have emerged. Nowadays, consumers' online behavior can be tracked and linked across a variety of websites and services, as well as across devices. In turn, this amassed information can be used by marketers to attempt to target advertising messages to specific consumers at the most beneficial time. The advertising industry has long posited that the "beneficial" qualifier applies to all stakeholders in the online advertising ecosystem. Online targeted advertising, the argument goes, is an economic win-win for different agents.[4] Merchants benefit from it because they can allocate their advertising budget to consumers most likely to be interested in their products and services; consumers benefit because highly personalized offers reduce their search costs; publishers (websites displaying the ads) benefit, as they can sell more valuable online real estate to advertisers; and, of course, the data

[4] See, for instance: Ad Exchanger, "If a Consumer Asked You, `Why Is Tracking Good?', What Would You Say?" October 2011, https://adexchanger.com/online-advertising/why-is-tracking-good/.

industry, by playing a matchmaking role via advertising platforms and ad exchanges, also benefits, via fees on the transactions completed by merchants.

There is, indeed, plenty of theoretical research suggesting that the type of matching made possible by new online technologies is beneficial. For instance, Bergemann and Bonatti (2011) show that an increase in targeting increases the total number of consumer–product matches, and thus increases the societal value of advertising. However (and this, by now, should not surprise the reader) there are also numerous nuances that the theoretical literature has unveiled. Let us leave aside, for the time being, broad concerns over the creepiness of targeted advertising (Ur et al. 2012), generic concerns over its privacy-invasive nature, or more recent concerns over how the technological infrastructure of targeted advertising, combined with the viral properties of social media, may end up subtly influencing individual decision-making (see Kramer et al. 2014 on subtle emotional contagion via social media) and even elections (Persily 2017). Let us focus, instead, on purely economic factors. Johnson (2013) observes that increasingly sophisticated and accurate technologies to identify consumers and target them with personalized ads may not always increase consumer welfare, as consumers who prefer ads from mainstream sellers may end up seeing ads from niche ones. Hagiu and Jullien (2011) find that data intermediaries that use consumer data to influence the matching between buyers and sellers may not always, in fact, facilitate the best possible match. In fact, both De Corniere and De Nijs (2016) and Zhang and Katona (2012) find that intermediaries may not always have an incentive to share all consumer data with advertisers, or to improve targeting precision, as this may reduce their own profits. In other words, data intermediaries that control, via personal information, the matchmaking process between consumers and sellers may have incentives to strategically modulate the amount of information they share with other stakeholders in the advertising ecosystem to increase their private, as opposed to societal, welfare.

Similar conclusions are found in a recent paper by Marotta et al. (2016), in which I was a co-author. The manuscript proposes a theoretical framework for understanding the trade-offs that different stakeholders in the advertising ecosystem face under different regimes of consumer information disclosure and collection. The framework focuses on programmatic, targeted advertising, which increasingly relies on "Real-Time Bidding" – an auction-based process through which "inventory" (that is, ad space on publishers' sites) can be bid for and bought by advertising

merchants via so-called ad exchanges. The different regimes of information disclosure contrasted in the manuscript include a full information regime, where information about both consumer preferences for products and overall spending ability can be collected by the data intermediaries (the ad exchanges, in the model setup); a no information regime (where no consumer data can be collected and used for targeting); and two limited information regimes that protect one type or the other of consumer information (that is, either information about specific consumer preferences for products or their overall spending ability). The no information and the limited information regimes reflect scenarios where consumer data may be protected via an array of possible mechanisms, including technology (i.e., the usage of privacy enhancing technologies – see Section 5) or, in fact, regulation. The model, therefore, can be used to analyze how the economic welfare of specific stakeholders (consumers; merchants; a monopoly ad exchange), as well as aggregate welfare, vary depending on how much consumer data can be collected and used for targeting, and whether regulation is imposing limitations on that collection.

We find an interesting (and, again, nuanced) set of insights. First, the economic interests of the different agents are not necessarily always aligned. In other words, different agents may prefer different economic regimes, and this may in turn depend on the specific distribution of consumer preferences and spending abilities. Second, the ad exchange intermediary, in general, prefers scenarios where spending ability information is collected and shared with advertising firms, but specific consumer preferences are not shared (this leads to more competition across merchants to bid to show ads to consumers). Thus, an intermediary platform may strategically modulate how much information it is sharing with other agents of the ecosystem in order to drive up its profits. Third, the ultimate effect of targeting on consumer welfare is affected both by what type of information is available to the ad exchange and the merchants, as well as by the distribution of consumer preferences. In fact, the ultimate impact of targeting on consumers may be either positive or negative. It will depend on the degree of consumers heterogeneity in preferences, as well as on what information about them is used in the advertising ecosystem. Thus, depending on the conditions, regulation affecting the information regime faced by the advertisers may or may not help consumer (and aggregate) welfare.

To crystallize the main lessons from this section: from a theoretical perspective, it is not self-evident or obvious that more data sharing will be always welfare-increasing, and that

regulation limiting collection or use of data will be always welfare-decreasing. First, the welfare implications of the collection and use of data are nuanced and context-dependent. Second, perverse incentives may lead to over-collection and over-usage of personal data in some cases, and under-collection and under-usage in others. Third, the interests and incentives of different stakeholders will often be misaligned. In fact, both privacy regulation and – importantly – absence of it will have redistributive effects in the economy: whether or not we protect data often will inevitably create some economic winners and some losers; and there is no way out of this decision as even the decision *not* to intervene in the market place with regulation is itself a choice.

It is also worth pointing out that the vast majority (if not the totality) of studies covered in this section take a very narrow approach to the definition of privacy. They focus on clearly defined and theoretically measurable outcomes (prices paid, utility derived from products, and so forth). In other words, most of these studies do not account for or address the challenges in capturing the intricate, indirect, and less tangible implications of data protection or data sharing that Section 2 has discussed. This is something the reader may want to keep in mind in considering the overall balance of arguments for and against privacy regulation.

4. Empirical work

The nuanced outcomes espoused by the theoretical literature on privacy are reflected in a remarkable degree of diversity when outcomes are captured in empirical work. That work is covered in this section.

We start by noting that the works cited here cover a vast array of typologies of data and technologies, reflecting the multiple dimensions of privacy and the multiple scenarios in which privacy issues arise: from price discrimination to online advertising; from identity theft to medical records. There is no lack of studies focusing on the costs arising from the protection of data. Consider Goldfarb and Tucker (2011), for instance. The authors investigated how the EU ePrivacy Directive (a predecessor to the 2018 GDPR) impacted online visitors' purchase intentions. The ePrivacy Directive included rules meant to affect, and limit, the consumer data that advertisers could use. The authors used 3.3 million survey responses of online visitors who had been exposed to a display advertising campaign. They implemented a diff-in-diff-in-diff empirical approach,

looking at changes between control and treatment groups in hypothetical purchase intentions for advertised products before and after the enactment of the regulation, comparing European users to users in other countries. They found a statistically and economically significant decrease in display advertising after the EU laws were enacted.

Next, consider the case of healthcare – a sector where privacy may be particularly important, given the sensitivity of the data, but where innovation in information technology relying on patients' information may also prove critical to providing better health outcomes. In a series of papers, Miller and Tucker (2009, 2011, 2012) found that privacy laws did significantly reduce hospitals' adoption of novel electronic medical records systems (by over 24%). In turn, this may have had adverse effects on patients (as one of the cited articles estimated that even a 10% increase in adoption of EMRs could significantly decrease infant mortality rates).

Or, consider the real estate market and mortgage applications. Kim and Wagman (2015) estimated the impact on mortgage denial rates (in the 2001–2006 period) of stricter financial privacy laws by exploiting variations in the adoption of financial privacy ordinances across different California Bay Area counties. They found that denial rates for home-purchase and refinancing loans decreased in counties with *opt-in* privacy ordinances (that is, stricter laws requiring financial institutions to obtain waivers from individuals before being able to share their information with other companies). In addition, the authors found that foreclosure rates increased a few years later in those same counties.

The examples cited above suggest that privacy regulation can indeed lead to adverse economic outcomes. However, as usual, the assessment of the overall economic impact of regulation is actually quite nuanced. For instance, in the above-mentioned study by Goldfarb and Tucker (2011) the loss of advertising effectiveness following the enactment of European privacy regulation was actually found only for a specific (and somewhat narrow) set of online ads, suggesting that other ads (namely, dynamic, larger, or contextually relevant to the content of a specific site) were, in fact, *not* affected by regulation – suggesting that there are ways to regulate privacy in that context without hurting ad effectiveness. Furthermore, scholars have also put forward evidence of beneficial economic effects of privacy regulation. Romanosky et al. (2011) investigated how state-level enactment of data breach disclosure laws affected identity theft rates in the United States. The rationale for those laws (which different states across the country enacted

at different points in time) is that, when companies are compelled to disclose the data breaches they suffer, they will respond by investing more, ex ante, in security of their databases, in order to avoid the costs associated with disclosure (both the public relations costs, and the tangible costs associated with informing large numbers of consumers). This ex ante increase in security investments should be expected to reduce, ex post, data breaches and thus identity theft originating from them. In addition, upon receiving disclosure of a breach, consumers may be able to take protective steps (such as freezing their credit accounts), thus also reducing the likelihood of falling victim to identity theft. Using panel data from the U.S. Federal Trade Commission, and taking advantage (as an identification strategy) of the heterogeneity in the timing of the passage of disclosure laws across the United States, the authors estimated the impact of those laws in the period covering 2002–2009, and found that adoption of such laws did reduce identity theft caused by data breaches by, on average, 6.1%.

A way to make sense of the apparently contradictory results for the effect of privacy regulation on economic outcomes is, first, to observe once again that the trade-offs associated with privacy are context-dependent (see Section 3), and, second, that binary metrics such as absence vs. presence of regulation may be too coarse to capture and understand the effects of regulatory interventions in this space. As Goldfarb and Tucker (2012) observe, privacy regulation may indeed affect innovation, competition, and market structure. However, those impacts may be quite heterogeneous, as a function of the specific characteristics and provisions of the regulation itself. As the authors put it, "the effects of policy are not uniform. While policies that simply restrict the use of data appear to have a substantial negative impact on the scope of data-using industries, policies that enable choice and facilitate trust may have a much more muted effect. Furthermore, these costs and benefits vary substantially across industries and contexts. The details of any privacy regulation matter a great deal in terms of the potential impact on innovation."

A specific example is readily provided by Adjerid et al. 2015. The authors investigated the impact of privacy regulation on the adoption and success of health information exchanges (HIEs), which are technology efforts meant to increase coordination of patient care across healthcare providers in the United States. In particular, the authors used state-level variation in regulation containing privacy requirements for sharing healthcare data and regulation containing incentives to promote the growth of HIEs. They found that, yes, privacy regulation alone can decrease the

number of planning and operational HIEs; however, they also found that privacy regulation with requirements for patient consent, when coupled with incentives, can actually improve the success of HIE efforts. These results are notable because they show how the impact of privacy regulation may be quite heterogeneous (both positive and negative, in purely economic terms), depending on the specific attributes of privacy laws.

This section concludes with a look at two timely issues: privacy and social media, and GDPR. Acquisti and Fong (2019) used a large-scale online field experiment to capture the extent to which U.S. employers use social media to find information about prospective job applicants, and the extent to which their hiring decisions may be affected by information that is illegal, or risky, for employers to enquire about during job interviews, but which is nowadays often available via social media – made available by the job candidates themselves. The authors found evidence of such search behavior, as well as evidence that some employers can in fact discriminate across job applicants based on the information they find online (namely, religious affiliation). The relevance of this study for the discussion on privacy and regulation is the following: both at the federal level, and at the state level, various forms of regulatory provisions attempt to protect job seekers from discrimination based on protected traits, by making those traits "private" from the perspective of the employers. And yet, modern communication technology (and, in particular, social media) can effectively bypass those regulations by making it easy for job seekers to reveal that information publicly (yet sometimes unintentionally), and easy for employers to find it. The tension between existing regulations and novel technologies highlights the challenges in designing viable policy-making approaches to privacy in a digital world in constant evolution.

As for GDPR: the General Data Protection Regulation, enacted in May 2018, is one of the most sweeping regulatory initiatives to date in the realm of consumer privacy. It contains broad provisions regulating the collection and processing of personal data, including an increased degree of control afforded to individuals over their privacy. Importantly, GDPR applies not just to all enterprises established in the European Economic Area, but also to all individual citizens of the European Union. This implies that, for instance, U.S. firms doing business with EU consumers would have to abide by the provisions of the law when dealing with the personal data of their EU consumer base; and conversely, U.S. consumers doing business with EU firms would also be protected under the law. Such a significant piece of regulation is naturally expected to produce an

array of direct and indirect consequences with economic implications. Several scholars across different domains (including economics, computer science, marketing, and so on) have in fact started looking for empirical evidence of those consequences (see, for instance, Jia et al. 2018, who focus on GDPR's effect on venture capital investments in emerging technologies). It will take probably a few years, however, before the dust settles and a clear picture of the economic impact of GDPR can emerge.

5. Some open issues

Section 3 highlighted, and separated, two distinct questions regarding privacy protection and its economic impact: the "how much" question, and the "how" question. In terms of how much privacy is desirable from an economic perspective, a key lesson emerging from both the theoretical (Section 3) and empirical (Section 4) economic literature on privacy is that the consequences of data regulation are context-dependent and nuanced. In fact, those consequences span short- and long-term windows, have multiple possible metrics of interest, and include both welfare-increasing and welfare-decreasing scenarios. It is now time to go back to the "how" question: How do we best achieve certain degrees of balance between protection and collection of data?

In tackling the "how" question, it is important to go back to one of the lessons of Section 4: privacy regulation is not a binary concept (contraposing absence to presence of regulation). Rather, we must look at policy-makers' interventions in the privacy realm along a spectrum. Along this spectrum, different approaches towards balancing and managing privacy and collection exist. Those different approaches (self-regulation; market forces; technologies for invasion and technologies for protection of data; consumer choice and responsibility; and so on) are not mutually exclusive (that is, they can in fact exist and be used simultaneously); play different roles depending on the type, intensity, and forcefulness of the regulatory intervention into the economy; and are, in fact, affected by the specific type of existing regulation (or lack thereof).

At one end of the spectrum, a society may entirely rely on firms' self-regulatory efforts to manage personal data. In absence of direct policy-making intervention, self-regulation (in the form of industry standards, best practices, firms' self-imposed policies around data handling, and so on) is influenced and driven by market forces. For instance, firms may differentiate in terms of their

privacy stance towards consumers, with some firms taking a more protective stance (using this as a form of competitive advantage to gain consumer trust and increase their customer base), and others focusing on the value they can instead extract from collection and usage of consumer data. To some extent (since the following companies actually operate within an environment with both self-regulatory and regulatory frameworks, rather than purely self-regulatory), Apple and Facebook may be considered two current examples of different stances over consumer data, driven less by regulation than by different business models. In a self-regulatory environment, consumer decision-making plays a crucial role. Consumers are expected to consider different options available to them in the market, and to engage in so-called privacy "calculus" (Acquisti et al. 2016) estimations to decide the extent to which to share or protect their personal information. Due to their different traits and subjective preferences over privacy, consumers would be expected to self-select into different segments of the corporate privacy spectrum, and would interact with companies that meet their respective demands for privacy. Technology – and in particular the existence of so-called "privacy enhancing technologies," or PETs – would play a crucial role in these interactions. More protective companies would adopt privacy technologies to better protect user data (for instance, stronger encryption safeguards for databases containing consumer data); and more privacy-conscious consumers would adopt more protective technologies to handle their own information (choosing, for instance, Duckduckgo as their search engine over Google, or the Signal messaging app over Whatsapp).

Virtually no western country has a purely self-regulatory approach to privacy, however. The reason is obvious: we know, by now, that too many things can go awry with each of the moderating factors examined above. Market forces lose part of their ability to restrain firms' data collection and usage practices if network effects (particularly strong in two-sided market platforms such as those common online: search engines, advertising networks, social media) lead to quasi-monopoly power that leave few outside options to privacy-seeking consumers (Acquisti et al. 2016). Consumer decision-making in the realm of privacy has been shown, time and again, to be confounded by a series of hurdles and challenges that make it hard or even impossible for privacy-conscious consumers to properly exercise their choices in the marketplace: from asymmetric information to bounded rationality; from resignation and learned helplessness to various cognitive and behavioral biases that affect, in particular, privacy valuations and decision making (Acquisti et al. 2015). Those hurdles and challenges make it ineffectual to rely on so-called "notice and

consent" frameworks (see Acquisti et al. 2013) for privacy protection (such as those relying on privacy policies, privacy settings, and – effectively – consumer "responsibilization"; see Giesler and Veresiu 2014). Furthermore, soft paternalistic approaches relying on nudges have been shown to be only partly effective (Acquisti et al. 2017). And finally, technology does not work in a vacuum: while tools to protect data while still analyzing and extracting value from it do already exist (from homomorphic encryption to differential privacy) their existence does not imply market success or adoption, if the economic incentives make it too desirable to collect and use data without the hindrance of data-protective and data-degrading tools.

For all the reasons above, most countries approach privacy with a mix of self-regulatory outlooks and regulatory interventions. But the regulatory interventions themselves vary greatly in terms of the burden they put on firms to protect data subjects' privacy, or the responsibility they place in the hands of consumers to take charge of the protection of their own data. And, as remarked repeatedly across this Note, different types of interventions will have different and nuanced effects on the economy; and so will, of course, the absence of any intervention.

It may seem disappointing, but is hopefully understandable, that this Note therefore ends without a single explicit recommendation or conclusion – a thumbs-up or thumbs-down gesture towards one specific form of intervention. By way of a summary, we can go back to the "how much" and "how" questions, and note that, while the analysis did not highlight a final answer to either question (probably because there is no single correct answer, but a multiplicity of answers), it has at least shown some answers to be *in*correct: the notion that privacy protection will unambiguously decrease or impair valuable economic metrics such as welfare generation, efficiency, or economic growth (in fact, it will not *univocally* negatively affect such metrics); and the notion that, left to their own devices, self-regulated markets will achieve the optimal degree of protection and disclosure (in fact, they will in all probability over- or undershoot that precious balance).

References

Acquisti, Alessandro, and Christina M. Fong. "An experiment in hiring discrimination via online social networks." *Management Science* (2019, forthcoming).

Acquisti, Alessandro, Idris Adjerid, and Laura Brandimarte. "Gone in 15 seconds: The limits of privacy transparency and control." *IEEE Security & Privacy* 11, no. 4 (2013): 72-74.

Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein. "Privacy and human behavior in the age of information." *Science* 347, no. 6221 (2015): 509-514.

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. "The economics of privacy." *Journal of Economic Literature* 54, no. 2 (2016): 442-92.

Acquisti, Alessandro, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon et al. "Nudges for privacy and security: Understanding and assisting users' choices online." *ACM Computing Surveys (CSUR)* 50, no. 3 (2017): 44.

Adjerid, Idris, Alessandro Acquisti, Rahul Telang, Rema Padman, and Julia Adler-Milstein. "The impact of privacy regulation and technology incentives: the case of health information exchanges." *Management Science* 62, no. 4 (2015): 1042-1063.

Bergemann, Dirk, and Alessandro Bonatti. "Targeting in advertising markets: implications for offline versus online media." *The RAND Journal of Economics* 42, no. 3 (2011): 417-443.

Calo, Ryan. "The boundaries of privacy harm." *Ind. LJ* 86 (2011): 1131.

Cohen, Julie E. "What privacy is for." *Harv. L. Rev.* 126 (2012): 1904.

Daughety, Andrew F., and Jennifer F. Reinganum. "Public goods, social pressure, and the choice between privacy and publicity." *American Economic Journal: Microeconomics* 2, no. 2 (2010): 191-221.

De Corniere, Alexandre, and Romain De Nijs. "Online advertising and privacy." *The RAND Journal of Economics* 47, no. 1 (2016): 48-72.

Giesler, Markus, and Ela Veresiu. "Creating the responsible consumer: Moralistic governance regimes and consumer subjectivity." *Journal of Consumer Research* 41, no. 3 (2014): 840-857.

Goldfarb, Avi, and Catherine E. Tucker. "Privacy regulation and online advertising." *Management science* 57, no. 1 (2011): 57-71.

Goldfarb, Avi, and Catherine Tucker. "Privacy and innovation." *Innovation policy and the economy* 12, no. 1 (2012): 65-90.

Hagiu, Andrei, and Bruno Jullien. "Why do intermediaries divert search?." *The RAND Journal of Economics* 42, no. 2 (2011): 337-362.

Hirshleifer, Jack. "The private and social value of information and the reward to inventive activity." *American Economic Review* 61, no. 4 (1971): 561-574.

Hirshleifer, Jack. "Privacy: Its origin, function, and future." *The Journal of Legal Studies* 9, no. 4 (1980): 649-664.

Jia, Jian, Ginger Zhe Jin, and Liad Wagman. *The short-run effects of GDPR on technology venture investment*. No. w25248. National Bureau of Economic Research, (2018).

Johnson, Justin P. "Targeted advertising and advertising avoidance." *The RAND Journal of Economics* 44, no. 1 (2013): 128-144.

Kim, Jin-Hyuk, and Liad Wagman. "Screening incentives and privacy protection in financial markets: A theoretical and empirical analysis." *The RAND Journal of Economics* 46, no. 1 (2015): 1-22.

Kramer, Adam DI, Jamie E. Guillory, and Jeffrey T. Hancock. "Experimental evidence of massive-scale emotional contagion through social networks." *Proceedings of the National Academy of Sciences* 111, no. 24 (2014): 8788-8790.

Laudon, Kenneth C. "Markets and privacy." *Communications of the ACM* 39, no. 9 (1996): 92-104.

Marotta, Veronica, Kaifu Zhang, and Alessandro Acquisti. "The Welfare and Allocative Benefits of Targeted Advertising," PrivacyCon FTC Conference, Washington DC, (2016).

Miller, Amalia R., and Catherine Tucker. "Privacy protection and technology diffusion: The case of electronic medical records." *Management Science* 55, no. 7 (2009): 1077-1093.

Miller, Amalia R., and Catherine E. Tucker. "Can health care information technology save babies?" *Journal of Political Economy* 119, no. 2 (2011): 289-324.

Miller, Amalia R., and Catherine E. Tucker. "Electronic discovery and the adoption of information technology." *The Journal of Law, Economics, and Organization* 30, no. 2 (2012): 217-243.

Murphy, Richard S. "Property rights in personal information: An economic defense of privacy." *Geo. LJ* 84 (1995): 2381.

Noam, Eli M. "Privacy and self-regulation: Markets for electronic privacy." *Privacy and Self-Regulation in the Information Age*. US Department of Commerce (1997): 21-33.

Persily, Nathaniel. "The 2016 US election: Can democracy survive the Internet?" *Journal of democracy* 28, no. 2 (2017): 63-76.

Posner, Richard A. "The right of privacy." *Ga. L. Rev.* 12 (1977): 393.

Posner, Richard A. "The economics of privacy." *The American economic review* 71, no. 2 (1981): 405-409.

Romanosky, Sasha, Rahul Telang, and Alessandro Acquisti. "Do data breach disclosure laws reduce identity theft?" *Journal of Policy Analysis and Management* 30, no. 2 (2011): 256-286.

Schoeman, Ferdinand David. *Privacy and social freedom*. Cambridge university press, 1992.

Solove, Daniel J. "A taxonomy of privacy." *U. Pa. L. Rev.* 154 (2005): 477.

Stigler, George J. "An introduction to privacy in economics and politics." *The Journal of Legal Studies* 9, no. 4 (1980): 623-644.

Taylor, Curtis R. "Consumer privacy and the market for customer information." *RAND Journal of Economics* (2004): 631-650.

Ur, Blase, Pedro Giovanni Leon, Lorrie Faith Cranor, Richard Shay, and Yang Wang. "Smart, useful, scary, creepy: perceptions of online behavioral advertising." In *proceedings of the eighth symposium on usable privacy and security*, p. 4. ACM, 2012.

Varian, Hal R. "Price discrimination." *Handbook of industrial organization* 1 (1989): 597-654.

Varian, Hal R. "Economic aspects of personal privacy." In *Privacy and Self-regulation in the Information Age*. US Department of Commerce (1997).

Westin, Alan F. "Privacy and Freedom." (1968).

Zhang, Kaifu, and Zsolt Katona. "Contextual advertising." *Marketing Science* 31, no. 6 (2012): 980-994.