The Economics of Privacy

Alessandro Acquisti[*] Curtis Taylor[†] Liad Wagman[‡]

Abstract

This article summarizes and draws connections among diverse streams of empirical and theoretical research on the economics of privacy. Our focus is on the economic value and consequences of privacy and of personal information, and on consumers' understanding of and decisions about the costs and benefits associated with data protection and data sharing. We highlight how the economic analysis of privacy evolved through the decades, as, together with progress in information technology, more nuanced issues associated with the protection and sharing of personal information arose. We use three themes to connect insights from the literature. First, there are theoretical and empirical situations where the protection of privacy can both enhance and detract from economic surplus and allocative efficiency. Second, consumers' ability to make informed decisions about their privacy is severely hindered, because most of the time they are in a position of imperfect information regarding when their data is collected, with what purposes, and with what consequences. Third, specific heuristics can profoundly influence privacy decision-making. We conclude by highlighting some of the ongoing issues in the privacy debate.

1 Why an Economics of Privacy

The value and regulation of information assets have been some of the most important focuses of economic research since Hayek (1945)'s treatise on the use of knowledge in society. Contributions to what has become known as the field of *information economics* have been among the most influential, insightful, and intriguing in the profession. Seminal studies have investigated the informative role of prices in market economies (Stigler, 1961); the creation of knowledge and the incentives to innovate (Arrow, 1962); the prevalence of asymmetric information and adverse selection (Akerlof, 1970); the transmission of private information through signaling activity (Spence, 1973); and voluntary disclosures (Grossman, 1981; Milgrom, 1981). It may be proper, however, to think of information economics not as a single field, but as an amalgam of many related sub-fields. One such sub-field currently receiving growing attention by economists is the subject of this survey: the study of privacy.

Privacy is difficult to define. It means different things to different people (Solove, 2006). It has been described as the protection of someone's personal space and their right to be left alone (Warren and Brandeis, 1890); as the control over and safeguard of personal information (Westin,

[*] Heinz College, Carnegie Mellon University, Pittsburgh, PA; acquisti@andrew.cmu.edu

[†] Economics Department, Duke University, Durham, NC; crtaylor@econ.duke.edu

[‡] Stuart School of Business, Illinois Institute of Technology, Chicago, IL; lwagman@stuart.iit.edu

1967); or as an aspect of dignity, autonomy, and ultimately human freedom (Schoeman, 1992). While seemingly different, all of these definitions relate to the boundaries between the self and the others, between private and public. As individuals and as consumers, we constantly navigate those boundaries, and the decisions we make about them determine tangible and intangible benefits and costs, for ourselves and for society. Thus, at its core, the economics of privacy concerns the trade-offs associated with the balancing of public and private spheres between individuals, organizations, and governments. Within this broad scope, economists' interest in privacy has primarily focused on its informational dimension: the trade-offs arising from the protection or sharing of personal data. Other sub-fields of information economics therefore also relate, in a sense, to the topic of this review, because they pertain to the trade-offs arising from the public or private status of information. For instance, an auction may be structured in a way that its participants will reveal their true costs or valuations, or a tax mechanism may be designed so that the agents will truthfully reveal their types. However, whereas research on auctions and optimal taxation may pertain to the *private* information of abstract economic agents (which could be consumers, firms, or other entities), the field of *privacy economics*, which is our focus, pertains more specifically to *personal* information of actual individuals. As a consequence, of course, the field is often influenced by research in the other streams of information economics.

This article reviews the theoretical and empirical economic literature investigating individual and societal trade-offs associated with sharing and protecting personal data. In particular, it focuses on the flow and use of information about individuals by firms. In so doing, the article identifies a number of key themes. One theme is that a single unifying economic theory of privacy is arguably infeasible, because privacy issues of economic relevance arise in widely diverse contexts. Nevertheless, we are able, within a given context, to identify a number of robust theoretical insights present in the literature. A second key theme is that both economic theory and empirical analysis of privacy expose varying scenarios — in some, privacy protection can decrease individual and societal welfare; in others, privacy protection improves them. Thus, it is not possible to conclude unambiguously whether privacy protection entails a net 'positive' or 'negative' change in purely economic terms: its impact is context specific. A third key theme relates to the observation that consumers are rarely (if ever) completely aware about privacy threats and the consequences of sharing and protecting their personal information. Hence, market interactions involving personal data often take place in the absence of individuals' fully informed consent.

1.1 The Value of Personal Data and the Value of Privacy

Economists' interest in informational privacy, generally intended as the control or protection of personal information, can be readily understood: the protection and disclosure of personal data are likely to generate trade-offs with tangible economic dimensions. The transition of modern economies toward production of knowledge, and recent radical advancements in information technology (in particular, the rise of the Internet), have vastly enlarged the amount of individual information that can be collected, stored, analyzed, and re-purposed for new uses. The ascent of the so-called Web 2.0 (blogs, social media, online social networks) has rendered individuals no longer mere consumers of information, but public producers of often highly personal data. The spread of mobile computing and sensor technologies has blurred the distinctions between digital and physical, online and offline. All of this has led to services that simultaneously generate and capture digital trails of personal and professional activities — activities that were previously conducted in private and left little or no trace.[1] Simultaneously, the Internet has evolved from an architecture of decentralized and possibly anonymous interactions (Berners-Lee et al., 2000), to one where packets of data capturing all types of behaviors (from reading to searching, from relaxing to communicating) are uniquely (Bendrath and Mueller, 2011) and sometimes personally (Xie et al., 2009) identified. In this environment, a few "gatekeeper" firms are in a position to control the tracking and linking of those behaviors across platforms, online services, and sites, for billions of users. As a result, chronicles of peoples' actions, desires, interests, and mere intentions are collected by third parties, often without individuals' knowledge and explicit consent (Acquisti, 2004), with a scope, breadth, and detail that are arguably without precedent in human history.

These vast amounts of collected information have substantial economic value. Individuals' traits and attributes (such as their age, address, gender, income, preferences, and reservation prices — but also clickthroughs, comments posted online, photos uploaded to a social media site, etc) are increasingly regarded as business assets that can be used to target services or offers, to provide relevant advertising, or to trade with other parties. In an effort to leverage the value inherent in personal data, new services (such as search engines and recommender systems), companies (such as social networking sites and blogging platforms), and even markets have emerged (such as markets for "crowdsourcing" (Schenk and Guittard, 2011), or the complex online advertising ecosystem (Evans, 2009)). Existing services, such as travel agencies, record companies, and news media, have also been transformed.

The tools and products made possible by the increased availability of personal data have borne benefits for data subjects and data holders alike. Despite those benefits, public concerns over personal privacy have increased. With the advent of Internet and data analytics, issues surrounding the protection or sharing of personal data have emerged as crucial nexuses of economic and policy debate.[2] Over the years, national surveys have consistently found widespread evidence of significant privacy concerns among Internet users.[3] From the standpoint of self-interested individual behavior,

[1] For instance, the act of listening to music online using a streaming service, compared to buying a CD, can be captured by the streaming service: the songs to which the user listened, from where, for how long, or for how many times. This data can be combined with other information about the individual, and then used in various manners: to compile a profile of the listener; to infer his or her other interests and preferences; to present him or her with advertising; or to sell information to data aggregators or other parties.

[2] Consider, for instance, the 2013 White House's report on "Big Data: Seizing Opportunities, Preserving Values," available at `http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf`.

[3] For instance, a Pew Research Center survey of 1,002 adult users conducted in 2013 found that 86% had taken steps online to remove or mask their digital footprints, and 68% believe that current laws are not good enough in

the economic motive behind concerns for privacy is far from irrational, and nearly self-evident. If information is power, control over personal information affects the balance of economic power among parties. Thus, privacy can simultaneously grant protection from the economic leverage a data holder could hold over the data subject (if the merchant figures out how little you know about the product you are browsing, he may steer you towards merchandise or prices that serve his interests better than yours), as well as a tool the data subject may strategically use against the non-holder (if the salesperson cannot estimate your reservation price, you may be able to exploit this information asymmetry to cut a nice bargain).

For the individual, therefore, the potential benefits of protecting certain data, as well as the potential costs of revealing too much of it to the wrong parties (from price discrimination to other more odious forms of discrimination; from social stigma to blackmailing; from intangible nuisances to identity theft) are quite apparent. Equally apparent, however, are the costs others may incur when they find themselves in a position of information asymmetry relative to the subject. For instance, the firm that cannot conduct background checks on job applicants may end up hiring the wrong employees. As Posner (1981) points out, privacy is redistributive — as is, of course, the *lack* of privacy.

Beyond mere questions of redistribution, the trade-offs associated with protecting or sharing personal information are actually more nuanced for both the data subject and for the market (as well as society) as a whole, for a number of reasons.

First, individuals can directly benefit from sharing their data. It could be in the form of the pleasure of reconnecting with a long lost friend via an online social network; or in the form of coupons, discounts, or personalized services one receives after joining a merchant's loyalty program; or in the form of reduced search costs and increased accuracy of information retrieval one experiences when a search engine tracks them more closely. Those benefits turn into opportunity costs when the individual chooses not to reveal certain personal data.

Second, both positive and negative externalities may arise through the complex interplay of data creation and transmission. In particular, the benefits arising from individuals sharing their information, because of advances in data mining, may be enjoyed by society as a whole. For instance, aggregation of online searches can provide early alerts for epidemics (Dugas et al., 2012) or unveil unexpected interactions between pharmaceutical drugs (White et al., 2013). Conversely, other individuals' comfort with sharing data ("I have nothing to hide") may legitimize expansions of intrusive surveillance programs that affect the rest of society. Society may suffer when certain behaviors stay hidden (consider insider trading; or, consider social progress being delayed, and social norms failing to evolve, because of individuals' fears to disclose legitimate but fringe opinions); but society may also benefit when other information is suppressed (various jurisdictions allow certain juvenile criminal records to be expunged in the belief that unfettered re-integration of minors has positive social value). Similarly, an individual may personally benefit from other people's sharing

protecting online privacy (Rainie et al., 2013).

(for instance, collaborative filtering of other users' movie ratings may produce accurate viewing recommendations); conversely, an individual may pay a price when a merchant's business analytics tools permit the latter to accurately predict the reservation price of the former, based on the past behavior of other consumers. In fact, even an individual's costs (and ability) to protect her information may be a function of the disclosure choices made by others. That "anonymity loves crowds" is a common refrain in the literature on privacy-enhancing technologies, reflecting the observation that, online as offline, it is easier to hide as one among many who look alike. Conversely, protecting one's data becomes increasingly costly the more others reveal about themselves (for instance, the success of online social networks has encouraged other entities, such as online news sites, to require social media user IDs in order to enjoy some of their services, thus curtailing users who do not want to create social media accounts), or altogether infeasible (even if an individual chooses to protect certain data, that data may still be inferable through the analysis of similar individuals who did not choose to protect theirs; see, e.g., Jernigan and Mistree, 2009).

Analyzed as economic goods, privacy and personal information reveal other, peculiar characteristics.

First, when shared, personal information has some of the characteristics of other public goods, such as non-rivalry and non-excludability (in fact, a complex online advertising ecosystem engages in trades of Internet users' personal information). And yet one of the prime components of informational privacy is the ability to keep that information protected — that is, to exclude someone from knowing or using certain information. The value of keeping some personal information protected, and the value of it being known, are almost entirely context-dependent and contingent on essentially uncertain combinations of states of the world. Privacy sensitivities and attitudes are subjective and idiosyncratic, because what constitutes sensitive information differs across individuals.

Specifically, individuals differ in what they may experience if some private information were to be made public, as well as in their beliefs that the information may in fact be released. For instance, the healthy individual who just lost his job may flaunt his active lifestyle on social media, but hide his unemployment status to avoid shame; the reverse may be true for the affluent manager who was just diagnosed with an STD. Different pieces of information will matter differently to different people (your piano teacher may not be as interested in the schools you attended, unlike your potential employer). The value of information will change over time (an online advertiser may not be as interested in logs of your online activity from five years ago as in your activity right now). In fact, the value and sensitivity of one piece of personal information will change depending on what other pieces of data it can be combined with (your state of birth and your date of birth, alone, may not uniquely identify you; together, they may allow the prediction of your Social Security number with some accuracy; see, e.g., Acquisti and Gross, 2009).

Second, disclosing data often causes a reversal of informational asymmetries: beforehand, the data subject may know something the data holder does not (for instance, a customer's willingness to pay for a good); afterwards, the data subject may not know what the data holder will do with their data, and with what consequences (for instance, how the merchant will use the customer's information, including estimates of her reservation price, following a purchase). As a consequence, privacy trade-offs are also inherently inter-temporal: disclosing data often carries an immediate benefit, be it intangible (friends "liking" your online status updates) or tangible (a merchant offering you a discount). The costs of doing so are often uncertain, and are generally incurred at a more distant point in time (a future prospective employer will not like your updates as much as your friends did; a merchant will use your information for price discrimination).

Third, and as already implied, privacy trade-offs often mix the tangible (the discount I will receive from the merchant; the increase in premium I will pay to the insurer), with the intangible (the psychological discomfort I experience when something very personal is exposed without my consent), and the incommensurable (the effect on society of surveillance; the loss of autonomy we endure when others know so much about us).

Fourth, privacy has elements of both a final good (one valued for its own sake), and an intermediate good (one valued for instrumental purposes; see, e.g., Farrell, 2012). Attitudes towards privacy mainly capture subjective preferences; that is, people's valuations of privacy as a good in itself (privacy as a final good). But those valuations are separate, and sometimes even disjoint, from the actual trade-offs that arise following the protection or sharing of personal data (from price discrimination to identity theft; from coupons to personalized services) — that is, from the value of privacy as an intermediate good (for instance, regardless of whether an individual thinks "my life is an open book, I have nothing to hide," that individual will still suffer tangible harm if she is victim of identity theft).

Fifth, it is not always obvious how to properly value privacy and personal data. Should the reference point be the price one would accept to give away their data, or the amount they would pay to protect it? Is it the expected cost the data subject may suffer if her data is exposed, or the expected profit the data holder can generate from acquiring her personal information? For traditional products and services that economists study, the way to address these questions is generally self-evident: the market captures the accurate price of privacy and personal data, reflecting the reservation prices of different buyers (data holders) and sellers (data subjects). However, there is yet no open, recognized market for personal data in which data subjects themselves can participate. Personal data is continuously bought, sold, and traded among firms (from credit-reporting agencies to advertising companies to so-called "infomediaries" which buy, sell, and trade personal data), but consumers themselves do not have access to those markets: they cannot buy back their data, or offer their data for sale (even though the concept of personal-information markets for consumers, or individuals' markets for privacy, has been around since the mid-1990s; see, e.g., Laudon, 1996, and Section 2.1 of this survey). Moreover, issues associated with individuals' awareness of privacy challenges, solutions, and trade-offs cast doubts over market outcomes accurately capturing and revealing, by themselves, individuals' true privacy valuations. However, individuals do engage daily in transactions involving their personal data. With a query on a search engine, the searcher implicitly sells information about her current interests in exchange for finding relevant results. By using an online social network, members implicitly sell information about their interests, demographics, and networks of friends and acquaintances, in exchange for a new method of interacting with them. Applying the principle of revealed preference, we could infer people's valuations for their personal data by observing their usage of those tools. However, for service providers data trading is the essence of the transaction, whereas from the perspective of the data subject, the trade of personal data is a secondary, mostly inconspicuous, and often altogether an invisible aspect of a different, more salient transaction (having a question answered; interacting with peers online, etc).

1.2 Focus of The Survey

Information asymmetries regarding the usage and subsequent consequences of shared information, as well as heuristics studied by behavioral decision researchers, raise questions regarding individuals' abilities, as rational consumers, to optimally navigate privacy trade-offs. They raise questions about the extent to which individual responsibility, market competition, and government regulation, can steer the market towards a balance of disclosure and protection of personal data that best serves the interests of the different parties. Which brings up more questions: are there privacy "equilibria" that benefit both data holders and data subjects? What is the allocation of surplus gained from the usage of individuals' personal data? And how should it be allocated – based on market forces, treating privacy as another economic good, or based on regulation, treating privacy as a fundamental right? Should an allocation favor the data subject as the owner of the data, or the data holder who invested in collecting and analyzing the information?

The studies we review in the remainder of this article investigate these diverse issues. The review focuses on the economic value and consequences of privacy and of personal information, and on consumers' understanding of and decisions about the costs and benefits associated with data protection and data sharing. In investigating these issues, we focus more on microeconomic than macroeconomic analysis. We focus on scholarly work published in economic journals — although, due to the nature of the subject, we also draw from fields such as psychology, marketing, information systems, and computer science. We begin with a survey of the theoretical literature on privacy (Section 2). The survey highlights how the economic analysis of privacy evolved through the decades, as, together with progress in information technology, more nuanced issues associated with the protection and sharing of personal information began to arise. As mentioned above, a key theme emerging from the theoretical literature is that there is no unequivocal impact of privacy protection (or sharing of information) on welfare; depending on context and conditions, privacy can either increase or decrease individual as well as societal welfare. Next, we survey the empirical literature on privacy trade-offs, as well as what is known about consumers' attitudes and behaviors towards privacy (Section 3). The review of empirical work on privacy reveals various insights. First, it confirms the principal theme arising from the theoretical literature: empirical evidence exists both for scenarios in which the protection of privacy slows innovation or decreases economic growth, and scenarios in which the opposite is the case. A second insight highlights consumers' inability to make informed decisions about their privacy, because most of the time they are in a position of imperfect information regarding when their data is collected, with what purposes, and with what consequences. A third insight relates to the heuristics that can profoundly influence privacy decision-making, since privacy trade-offs are intertemporal in nature and often uncertain. Finally, we highlight current issues in the privacy debate that may be of interest to economists (Section 4).

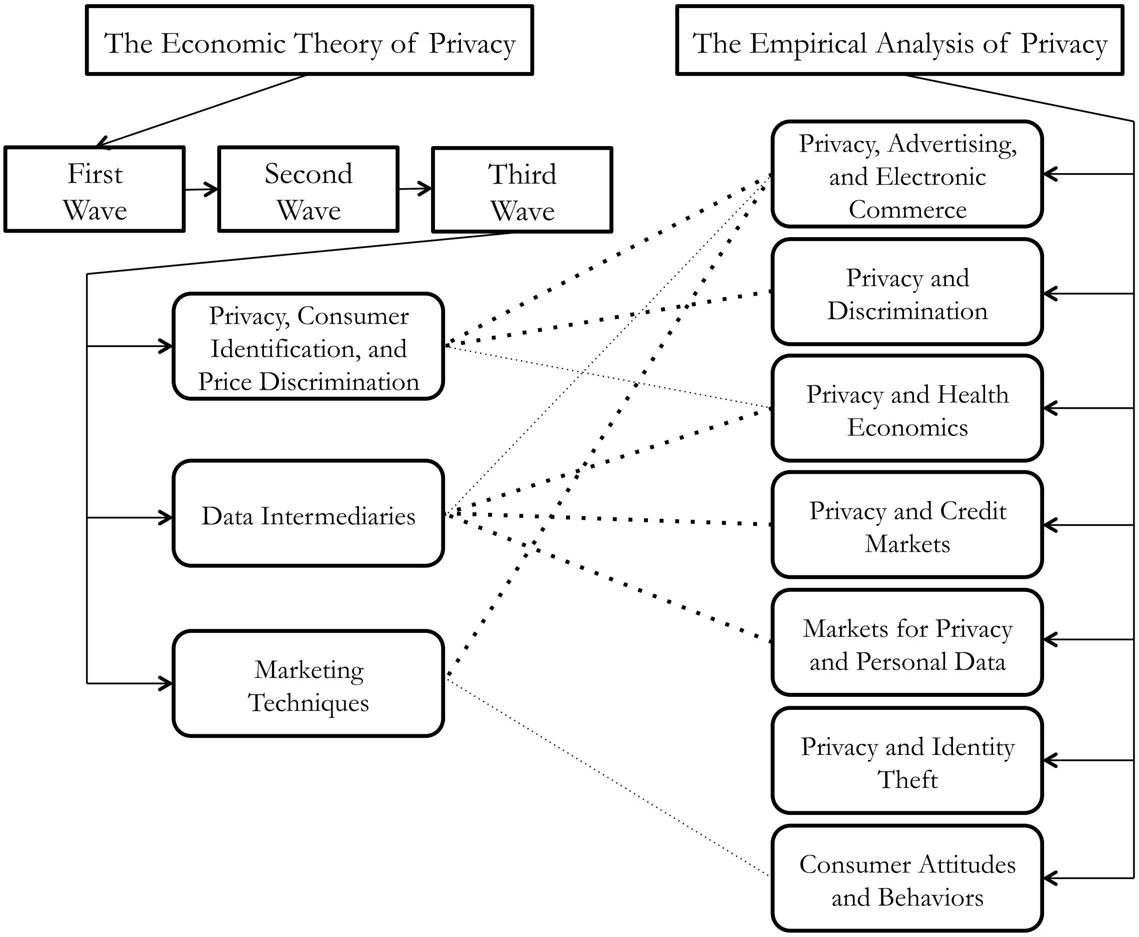
The Economic Theory of Privacy
The Empirical Analysis of Privacy
First Wave
Second Wave
Third Wave
Privacy, Consumer Identification, and Price Discrimination
Data Intermediaries
Marketing Techniques
Privacy, Advertising, and Electronic Commerce
Privacy and Discrimination
Privacy and Health Economics
Privacy and Credit Markets
Markets for Privacy and Personal Data
Privacy and Identity Theft
Consumer Attitudes and Behaviors

Figure 1: Depiction of primary and secondary connections between theory and empirics. Primary connections are dashed and bolded; secondary connections are dashed. Sections and subsections are listed in order matching that of the manuscript. We note that other connections exist (e.g., between Data Intermediaries and Privacy and Identity Theft), but we do not emphasize them in the manuscript and therefore do not emphasize them in the figure. The figure also helps illustrate that the study of Privacy and Identity Theft is somewhat distinct, crossing the threshold into the related field of Information Security.

Previous scholarship has distinguished different dimensions of privacy (such as seclusion, secrecy, solitude, anonymity, autonomy, freedom, and so forth; for a taxonomy, see Solove, 2007). As noted, this review focuses on informational privacy. Even under such a narrow focus, however, different dimensions and definitions of privacy emerge from the literature, such as privacy as *control over usage* of, versus privacy as *protection against access* of, personal information. Thus, this review covers studies as diverse as those that aim to capture individuals' willingness to pay to protect their data, and studies that capture the economic consequences of sharing or protecting data. Furthermore, when appropriate, the review touches upon other dimensions of informational privacy, such as the value of anonymity (which is a form of privacy of identity information); or the economic dimensions of spam or the do-not-call registry (which relate more to intrusions of a person's cyberspace made possible by knowledge of, rather than the flow, of their personal information); or the burgeoning literature on the economics of information security (which sometimes relates to privacy, for instance in studies of data breaches or identity theft that involve personal data, but more often relates to the protection of information infrastructures and other types of informational assets).

The diversity of privacy definitions and scenarios is reflected in the selection of manuscripts in this review. Figure 1 depicts some of the connections among the different areas of privacy that we categorize and survey in this manuscript. As we noted above, the reader should not hope to find a unified theory of privacy, or a single framework incorporating and connecting the diverse scholarly contributions we review. Privacy means different things to different people, and privacy issues with economic relevance arise in the most diverse contexts: from price discrimination to identity theft; from spam to targeted advertising. What connects these diverse definitions and scenarios is that they all involve the negotiation and management of the boundaries between private and public, and that those boundaries determine tangible and intangible trade-offs. Some of those privacy trade-offs may not just be intangible, but in fact immeasurable. The economics of privacy focuses on measurable, or at least assessable, privacy components. Some (perhaps, many) of the consequences of privacy protection, and its erosion, go beyond its economic dimensions — for instance, the intrinsic value of privacy as a human right, or individuals' innate desires for privacy regardless of its economic benefits or lack thereof. Using economics to study privacy does not imply the belief that such other, non-economic dimensions do not exist, or are not important. Quite the opposite: We acknowledge them, but do not focus on them, and we urge the reader to keep that in mind, when considering the broader policy implications of the economics of privacy.

2 The Economic Theory of Privacy

In this section, we discuss three waves of research in the economics of privacy: an early wave dating to the 1970s and early 80s; a middle wave active in the 1990s; and a more recent and growing third wave. For illustrative purposes, several simple and parsimonious algebraic examples appear throughout the discussion. Due to the many diverse scenarios in which issues of informational privacy arise, and their many dimensions, the examples we offer are not meant to represent any particular model or class of models, but rather to illustrate the complexity inherent in privacy trade-offs and in any potential regulation.

2.1 The First Wave

While privacy is far from a modern concept (Westin, 1967; Schoeman, 1984),[4] the extraordinary advances in information technology that have occurred since the second half of the twentieth century have brought it to the forefront of public debate. A first wave of economic research consists of seminal work produced between the 1970s and early 1980s by Chicago School scholars such as Stigler and Posner, and competing arguments by scholars such as Hirshleifer (1971, 1980). By and large, this initial wave of work did not consist of formal economic models, but rather general economic arguments around the value or the damage that individuals, and society, may incur when personal information is protected, thereby making potentially useful information unavailable to the marketplace.

Posner (1978, 1981) argues that the protection of privacy creates inefficiencies in the marketplace, since it conceals potentially relevant information from other economic agents. For instance, if a job seeker misrepresents her background and expertise to a hiring firm, protecting her personal information will negatively affect the firm's hiring decision. Therefore, the protection of the former's privacy comes at the cost of the latter's profitability, and removing an individual's personal information from the marketplace through privacy regulation ultimately transfers the cost of that person's possibly negative traits onto other market participants (see, also, Posner, 1993).

Similarly, Stigler (1980) argues that regulatory interference in the market for personal information is destined, at best, to remain ineffective. Because individuals have an interest in publicly disclosing favorable personal information and hiding negative traits, those who decide to protect their personal information (for instance, a debtor who does not want to reveal her credit history) are *de facto* signaling a negative trait. In this case, regulatory interventions blocking the flow of personal information would be redistributive and inefficient: economic resources and productive factors would end up being used inefficiently, or rewarded unfairly, because information about their quality had been removed from the marketplace.

However, Hirshleifer (1971, 1980) asserts that rational economic agents may end up inefficiently over-investing in collecting personal information about other parties, and that assumptions of rational behavior underlying the Chicago School's privacy models may fail to capture the complexity inherent in privacy decision-making by individuals and organizations. Hirshleifer shows that, given equilibrium prices, the private benefit of information acquisition may outweigh its social benefit (for more recent examples, see Burke et al., 2012; Wagman, 2014). In a pure exchange setting, information may have no social value at all, because it results only in a redistribution of wealth from ignorant to informed agents.

While not temporally belonging to the first wave of privacy literature, Daughety and Reinganum (2010) provide a very intriguing modern rebuttal to the Chicago-School view. The authors construct a model in which each individual cares about his reputation and in which his actions generate an

[4] Evidence of both a need and a desire for privacy, and a need and a desire for socializing and disclosing, can be found throughout history and across diverse societies; see, e.g., Acquisti et al. (2015).

externality (public good or bad). Under a regime of publicity, individuals distort their actions to enhance or preserve their reputations, whereas under privacy, they choose their individually optimal level of the activity. Thus, for example, whether an individual checks into a drug or alcohol rehab center should remain private because the stigma associated with doing so could otherwise deter him from seeking treatment, reducing both the private and public good. On the other hand, charitable contributions, typically, should be public, because contributing would raise the reputation of the giver and therefore increase the amount he would donate.

In a similar vein, but even more fundamentally, Spence (1973) can be viewed through a lens of privacy regulation. From this prospective, signaling activity may — as the Chicago scholars suggest — reveal payoff-relevant private information, but the aggregate cost of the signaling activity may nevertheless outweigh the benefits. Indeed, as is well-known, there are situations in which banning signaling behavior altogether (that is, enforcing a regime of complete privacy) may result in a Pareto improvement. This is illustrated in the following example inspired by Gottlieb and Smetters (2011), who report that nine of the fifteen most selective MBA programs in the US (including Chicago's own Booth School) do not disclose student grades to prospective employers.[5]

Example 1 (Signaling and Privacy)**.** suppose that an MBA student of privately known ability $\theta \in [0, 1]$ can earn grades of $g \geq 0$ by incurring effort cost g/θ. Upon graduating, her productivity on the job will be θ. Firms make competitive wage offers to each graduate. In particular, if the business school publicly reports its grades, then firm i makes an offer $w_i(g)$ to a graduate with grades g. On the other hand, if the business school keeps grades private, then firm i must make the same offer $\bar{w}_i$ to all graduates. The utility of a type θ student is given by

$$U(w, g; \theta) = w - \frac{g}{\theta}.$$

It is straightforward to verify that if grades are public, then in the unique outcome of a *least-cost separating equilibrium*, a student of ability θ will earn grades of $g^*(\theta) = \theta^2/2$ and receive a wage offer $w^*(g^*(\theta)) = \theta$. Her equilibrium payoff, therefore, will be $U^*(\theta) = \theta/2$. On the other hand, if the business school keeps grades private, then each student will earn a wage $\bar{w} = \mathrm{E}[\theta]$ and will 'waste' no effort on grade seeking. Thus, privacy of grades is a Pareto optimal policy iff $U^*(1) \leq \mathrm{E}[\theta]$ or $\mathrm{E}[\theta] \geq 1/2$.

2.2 The Second Wave

Following the first wave, economists, by and large, did not again exhibit particular interest in the economics of privacy for over a decade. This changed in the mid-1990s, arguably because

[5] See, also, the US Family Educational Rights and Privacy Act (FERPA), in connection to the privacy of student education records, `http://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html`, as well as the US Equal Employment Opportunity Commission (EEOC), which governs the federal legalities of information flows in hiring practices, `http://www.eeoc.gov/laws/practices/`.

of progress in digital information technologies on multiple fronts (the proliferation of electronic databases and personal computers, the advent of the Internet, the diffusion of electronic mail), which led to a new set of economic issues involving the usage of personal data. This second wave is similar to the first in terms of a preference for articulating economic arguments rather than formal models. However, it is differentiated from the first wave not just temporally, but also in terms of the specificity of the privacy scenarios considered, and an emergent awareness of the role of digital information technologies: Works produced in this wave began focusing on issues such as the role of cryptographic technologies in affecting economic trade-offs of data holders and data subjects, or the establishment of markets for personal data, as well as the economic implications of the secondary uses of personal information.

Among them, Varian (1997) observes that the development of low-cost technologies for data manipulation generates new concerns for personal information processing. Varian (1997) nonetheless recognizes that consumers may suffer privacy costs when too *little* personal information about them is being shared with third parties, rather than too much. The consumer, Varian notes, may rationally want certain information about herself known to other parties (for instance, a consumer may want her vacation preferences to be known by telemarketers in order to receive offers and deals from them that may actually interest her). The same consumer, however, may rationally not want *too much* information to be known by others — for instance, her willingness to pay for the deals in which she is interested.

This line of reasoning echoes Stigler's and Posner's approach, but Varian (1997) additionally notes that the secondary usage of personal data raises novel concerns. A consumer may rationally decide to share personal information with a firm because she expects to receive a net benefit from that transaction; however, she has little knowledge or control over how and by whom that data will later be used. The firm may sell the consumer's data to third parties, which may lead to spam and adverse price discrimination, among other concerns (Odlyzko, 2003). Such negative externalities may not be internalized by the consumer nor by the firm that distributes the information (Swire and Litan, 1998).

In accordance with the Coase Theorem (Coase, 1960), Noam (1997) argues that whether or not a consumer's data will remain protected does not depend on the initial allocation of rights on personal information protection (that is, it will not depend on the presence or lack of a privacy regulatory regime). Instead, whether data will eventually be disclosed or protected ultimately depends on the relative valuations of the parties interested in the information (what the presence or lack of a regulatory regime *will* affect, however, is which party — the data subjects, or the data holders — will pay the other for access to, or protection of, personal data). Coasian arguments in the analysis of privacy are also proposed by Kahn et al. (2000), but they depend on consumers being aware of, and internalizing, the costs and benefits of trading their private information.[6]

Laudon (1997) proposes the creation of information markets where individuals own their per-

[6] For analysis of the scope of the Coase Theorem in the presence of private information, see Farrell (1987).

sonal data and can transfer the rights to that data to others in exchange for some type of compensation. Similar to the view of the Chicago School scholars, Laudon argues that the mere legal protection of privacy is outdated, and a system based on property rights over personal information would better satisfy the interests of both consumers and firms. Clearly, a system of property rights over personal information would require appropriate legislation to define and assign those rights. This observation reveals that market-based and regulatory approaches to privacy are not binary opposites, but rather points on a spectrum. At one end of the spectrum, one would find regimes where no privacy legislation exists, and therefore the protection of data relies entirely on consumers' marketplace behavior (for instance, strategies such as avoiding interactions with firms that do not provide adequate protection of one's data, or adopting privacy-enhancing technologies to prevent the leakage of personal data), and on firms' self-regulated, competition-driven data-handling policies. On the opposite end of the spectrum, privacy regulation would establish strict default protection of, and limitation over the usage of, personal data. Somewhere in between, legislative initiatives may create a framework for property rights over personal data and for means to trade those rights across data subjects and potential data holders. While the assignment of property rights is generally welfare enhancing, granting consumers the right to sell their personal data may actually undermine consumer surplus, as illustrated in the following example.

Example 2 (A Market for Consumer Information)**.** Consider a market for a certain good, composed of a measure 1 of massless consumers. The consumers' valuations for the good are uniformly distributed on $[0, 1]$. The market is served by a monopolist with production cost normalized to zero. Absent a market for information, the monopolist would set its price at $p^M = \frac{1}{2}$; it would earn profit of $\frac{1}{4}$; and the top half of the market would earn aggregate consumer surplus of $\frac{1}{8}$.

Now suppose that each consumer possesses verifiable information (e.g., place of residence or employment) that correlates perfectly with her valuation for the good. The monopolist first makes an offer to pay $r \geq 0$ to any consumer who reveals her information. It then uses the information thus obtained to make personalized price offers $\hat{p}(v)$ to those consumers who sold their information and it posts a common price p to all those who did not.

It is straightforward to verify that in the unique Perfect Bayesian equilibrium of this game the following must hold. The monopolist offers $r = 0$ for information. Nevertheless, *all* consumers reveal their valuations. The monopolist sets $\hat{p}(v) = v$ and $p = 1$. The intuition here is similar to that in Grossman (1981) and Milgrom (1981). The marginal anonymous consumer makes no surplus and, therefore, is always willing to reveal her valuation for an arbitrarily small payment, but this means that there can be *no* marginal anonymous consumer in equilibrium. That is, the set of anonymous consumers unravels from the bottom.

This situation raises social surplus by $\frac{1}{8}$ and is allocatively efficient — the monopolist extracts all the social surplus of $\frac{1}{2}$. However, consumers are worse off: the unregulated market for information reduces consumer surplus from $\frac{1}{8}$ to 0, despite the fact that consumers initially owned property rights to their information.

2.3 The Third Wave

Following the commercial success of the Internet and the proliferation of databases containing consumer information, research on the economics of privacy dramatically increased at the start of the twenty-first century. Because so many transactions and activities, once private, are now conducted online, firms, governments, data aggregators, and other interested parties can observe, record, structure, and analyze data about consumer behavior at unprecedented levels of detail and computational speed (Varian, 2010). As a result, the digital economy is, to a degree, financed by the organization of large amounts of unstructured data to facilitate the targeting of product offerings by firms to individual consumers. For instance, search engines rely on data from repeat and past searches to improve search results, sellers rely on past purchases and browsing activities to make product recommendations, and social networks rely on selling data to marketers in order to generate revenues. This third wave, while temporally close to the second, is differentiated by the fact that studies are rooted in more formal economic models and empirical analyses, including lab experiments (we consider empirical analyses separately in Section 3). In addition, this third wave is more directly linked to the novel economic issues brought forth by developments in information technology, including search engines, behavioral targeting, and social media. Thus, this third wave is more fragmented than the previous two in terms of the focus of analysis.

While much of the third wave is focused on issues surrounding privacy as the protection of information about a consumer's preferences or type (hence a significant number of models examine the relationships between privacy and dynamic pricing), different dimensions to privacy (and even just informational privacy) exist, and economic trade-offs can arise from different angles of the same privacy scenarios. Consequently, other streams of work we consider in this section include related but disparate issues such as the rise of spam, the development of markets for privacy (Rust et al., 2002), behavioral targeting, the economic analysis of (personal) information security, and the relationship between public goods (resp. bads), social recognition (resp. pressure) and privacy (Daughety and Reinganum, 2010).

2.3.1 Privacy, Consumer Identification, and Price Discrimination

Intended as the analysis of the relationships between personal data and dynamic pricing, the economics of privacy is closely connected to the vast stream of studies on intertemporal price discrimination based on consumer recognition. This literature solidifies the notion of consumer tracking and personalized pricing, but does not explicitly consider privacy issues in online environments. Chen (1997) studies discriminatory pricing when different consumers buy different brands, and Fudenberg and Tirole (1998) explore what happens when the ability to identify consumers varies across goods — they consider a model in which consumers may be anonymous or "semi-anonymous," depending on the good purchased. Villas-Boas (1999) and Fudenberg and Tirole (2000) analyze a duopoly model in which consumers have a choice between remaining loyal to a firm and defecting to a competitor, a phenomenon they refer to as "consumer poaching" (Asplund et al., 2008, demonstrate evidence of this sort of poaching in the Swedish newspaper industry). They show that a firm always has an incentive to offer discounts to a rival firm's customers who have revealed, through their prior purchases, their preferences for the rival firm's product. Such discounts initially tend to reduce consumer price sensitivity for a firm's product, as consumers rationally anticipate them; hence prices rise in later periods, thanks to anticipated customer poaching. Chen and Zhang (2009) study a "price for information" strategy, where firms price less aggressively in order to learn more about their customers. Jeong and Maruyama (2009) and Jing (2011) identify conditions under which a firm should discriminate against its first-time and repeat customers.

More specific to privacy, Taylor (2004) finds that, in the presence of tracking technologies that allow merchants to infer consumers' preferences and engage in price discrimination, the usefulness of privacy regulatory protection depends on consumers' level of sophistication. Naïve consumers do not anticipate a seller's ability to use any and every detail about their past interactions for price discrimination; consequently, in equilibrium, their surplus is captured by firms — unless privacy protection is enforced through regulation. Regulation, however, is not necessary *if* consumers are aware of how merchants may use their data and adapt their behaviors accordingly, because it is in a company's best interest to protect customers' data (even if there is no specific regulation that forces it to do so). This is an example of how consumers, with their choices, could make a company's privacy-intrusive strategies counterproductive (we discuss in Section 3 studies highlighting consumers' awareness and knowledge of tracking technologies and privacy trade-offs).

Similar conclusions are reached by Acquisti and Varian (2005), who study a two-period model in which merchants have access to "tracking" technologies and consumers have access to "anonymizing" technologies. Internet commerce offers an example: merchants can use cookies[7] to track consumer behavior (in particular, past purchases and browsing activities), and consumers can delete cookies, use anonymous browsing or payment tools, and so forth, to hide that behavior. Acquisti and Varian (2005) demonstrate that consumer tracking will raise a merchant's profits only if the tracking is also used to provide consumers with enhanced, personalized services.

Complementary to the above works, Villas-Boas (2004) shows how strategic consumers may make a firm worse off in the context of dynamic targeted pricing. The reason is that once consumers anticipate future prices, they may choose to skip a purchase today to avoid being identified as a past customer tomorrow — and thus have access to lower prices targeted at new consumers. This strategic "waiting" on the part of consumers can hurt a firm both through reducing sales and diminishing the benefit of price discrimination, and may push a firm to voluntarily adopt a privacy-friendly policy.

[7] Cookies refer to files that are stored on a user's device, which can be subsequently used to help recognize the user across different webpages, websites, and browsing sessions.

Calzolari and Pavan (2006) consider the exchange of information regarding customers between two companies that are interested in discovering consumers' willingness to pay. They find that the transmission of personal data from one company to another may in some cases reduce information distortions and enhance social welfare (see also Pavan and Calzolari, 2009; Kim and Choi, 2010; Kim and Wagman, 2015). Information disclosure is therefore not always harmful to the individual and may contribute to improving the welfare of all parties involved. Moreover, in line with Taylor (2004), companies may be inclined to develop their own privacy protection policies for profit-maximizing purposes, even without the intervention of a regulatory body.

Conitzer et al. (2012) confirm these findings in a model where strategic consumers can opt to remain anonymous towards sellers at some cost — a cost modeled as the monetary-equivalent burden of maintaining privacy. The authors show that consumer surplus and social welfare are non-monotonic in this cost, reaching their highest levels at an intermediate level of privacy. In contrast, consistent with the above works, Conitzer et al. (2012) show that a firm would prefer to give consumers the option to opt out of receiving targeted pricing.

In a complementary piece, Campbell et al. (2015) demonstrate that if privacy regulation only relied on enforcing opt-in consent, an unintended consequence may be the entrenching of monopolies. The authors show that consumers are more likely to grant their opt-in consent to large networks with a broad scope rather than to less established firms. Hence, if regulation focuses only on enforcing an opt-in approach, users may be less likely to try out services from less established firms and entrants, potentially creating barriers to entry by leading to a "natural monopoly" in which scale economics include privacy protection.

Armstrong and Zhou (2010) study a duopoly search model where consumers may choose not to purchase a product on their first visit — and sellers record this behavior. They show that, in equilibrium, firms set higher prices for returning consumers, whereby first-time visitors would pay discounted rates, and that such practices may lead consumers not to return. These types of pricing strategies can result in consumer backlash — akin to what took place with Amazon in 2001 (Anderson and Simester, 2010), which may lead firms to commit upfront not to engage in such practices. Indeed, one theme resonating throughout this line of research is that firms with market power often benefit from committing to privacy policies. This is illustrated in the following simple example.

Example 3 (Repeat Purchases and Customer Tracking)**.** Suppose a population of n individuals wishes to consume one unit of a good in each of two periods. Half of the individuals are high-valuation consumers who value the good at 1 in both periods and the other half are low-valuation consumers who value it at $\lambda \in \left(0, \frac{1}{2}\right)$ in both periods. Each consumer's valuation is privately known. The good is sold by a monopolist with production cost normalized to 0. The consumers and the firm are risk neutral and (for simplicity) do not discount the future. Also, it is common knowledge that the monopolist possesses a tracking technology (for instance, cookies, or browser fingerprints) with which it can recall whether a consumer purchased the good in the first period and what price he paid for it. Moreover, the monopolist may use this information to make personalized price offers to consumers in the second period.

It can be shown (see, e.g., Taylor, 2004; Acquisti and Varian, 2005) that on the path of play in any Perfect Bayesian Equilibrium of this game the following must hold: The monopolist makes first-period price offers $p_1 = 1$ to all consumers and second period offers $p_2 = 1$ to all consumers regardless of their purchase histories. A low valuation consumer never purchases the good. A high valuation consumer purchases with probability 1 in the second period but purchases with probability $\frac{1-2\lambda}{1-\lambda} < 1$ in the first period (leaving the monopolist just indifferent between $p_2 = 1$ and $p_2 = \lambda$ following a first period rejection).

If the monopolist could publicly commit not to use the tracking technology, then the price offers would be the same, $p_1 = p_2 = 1$, but high valuation consumers would accept with probability 1 in the first period because rejections could never induce lower second-period prices. Thus, the tracking technology leads to strategic first-period rejections by high-valuation consumers, a Pareto inferior outcome that reduces social surplus (in the form of monopoly profit) by $\frac{n\lambda}{1-\lambda}$.

2.3.2 Data Intermediaries

A number of works have incorporated questions regarding privacy into the study of two-sided markets. Such studies can help us understand the role of large data holders — companies such as Google, Facebook, and Amazon — which in part act as intermediaries, selling advertising space to advertisers on one end and providing services and products to users on the other.

Cornière (2011) shows that when consumers actively search for products, targeting leads to more intense competition. In a framework in which consumers search sequentially after having entered a query on a search engine, he shows that targeting reduces search costs, improves matches between consumers and firms, and intensifies price competition. However, a profit-maximizing search engine may choose to charge too high an advertising fee, which can negate the benefits of targeting. Hence, the optimal level of accuracy in terms of advertising matching solves a trade-off between consumer participation and the profit of the intermediary.

Hagiu and Jullien (2011) study how intermediaries can use information about consumer characteristics in order to affect matching between firms and consumers. They show that if an intermediary receives a fee each time a consumer visits an affiliated firm, the intermediary has an incentive to direct consumers towards firms that they would not have visited otherwise. Doing so, the intermediary manipulates the elasticity of the demands faced by its affiliated firms. Bergemann and Bonatti (2013) study the acquisition of user-pertinent information by an advertising platform and its subsequent sale to advertisers. In their model, a data provider sets the price of an information record (e.g., a cookie). Advertisers subsequently acquire information records from the data provider, form posterior beliefs about consumer types, and purchase advertising space.

The authors demonstrate situations where an improved precision of user information leads to fewer records being purchased. Consequently, a data provider may choose to restrict or cap advertisers' access to information about users (that is, constrain or reduce its precision) in order to sell more records and generate greater profits.

Gehrig and Stenbacka (2007) examine lenders in a repeated-interaction framework and consider the possibility of information sharing among lenders (for instance, via credit bureaus). The authors demonstrate that in the presence of information sharing, switching costs are essentially reduced, which relaxes competition for initial market shares and can end up reducing the welfare of borrowers. In other instances, firms may be reluctant to use first- or third-degree price discrimination, for fear of a public backlash (Anderson and Simester, 2010). De Cornière and Nijs (2014) rule out direct price discrimination based on consumers' personal information by focusing instead on firms' bidding strategies in auctions for more precise targeting of their advertisements. That is, given that consumers' private information provides a finer and finer segmentation of the population, firms can compete to advertise their non-discriminatory pricing over each of those consumer segments. By disclosing information about consumers, the platform ensures that consumers will see the most relevant advertisements, whereas when no information is disclosed under a complete privacy regime, ads are displayed randomly. They find that targeted advertising can lead to higher prices, and, in line with Levin and Milgrom (2010) and Bergemann and Bonatti (2013), that improving match quality by disclosing consumer information to firms might be too costly to an intermediary — because of the informational rent that is passed on to firms. Given a relationship between the match quality of advertising and consumer demand, it is then possible to specify conditions under which *some* privacy or some limits to disclosure are optimal for an intermediary (see, also, Cowan, 2007).

Zhang (2011) follows an approach that does not require the direct use of an intermediary but yields similar findings. He studies competitive markets with endogenous product design, and demonstrates that in an effort to avoid more aggressive pricing from competitors, market leaders may choose to introduce mainstream products that appeal to the broader segment of the population. By doing so, rather than pursuing an approach of product differentiation, firms can limit consumers' strategic release of preference information — similar to what an intermediary would do — in order to dampen competition and facilitate product entry.

Another approach to limit the release of information by consumers is explored in the study of intermediary gatekeepers (Baye and Morgan, 2001; Wathieu, 2002; Pancras and Sudhir, 2007) — a third party that provides consumers with access to some degree of anonymity, possibly at a cost. Consistent with the above works, Conitzer et al. (2012) show that it can be profit maximizing for both firms and a gatekeeper to reach agreements for granting users the ability to freely anonymize.[8]

[8] Kearns et al. (2014) study the design of mechanisms that satisfy the computer science criterion of differential privacy (Dwork, 2006) — put simply, the notion of being able to distinguish one agent (a consumer) from another in a dataset of consumer characteristics with only a low probability. They show that mechanisms can be designed to satisfy a variant of this criterion when there are large numbers of agents, and any agent's action affects another

At the same time, Taylor and Wagman (2014) demonstrate that the effects of firms' ability to target individual consumers on consumer surplus, profits, and overall welfare is context dependent, whereby any conclusions drawn from a given model must be understood within its specific market setting.

A common lesson arising from this literature is that firms — be they advertisers or data intermediaries — seldom possess socially optimal incentives to match consumers with products. This is illustrated with the following example suggested to us by Alessandro Bonatti.

Example 4 (Buying and Selling Consumer-Level Information)**.** An advertiser faces a continuum of heterogeneous consumers and a monopolist data provider. The match value v_i between consumer i and the advertiser's product is uniformly distributed on $V = [0, 1]$. The advertising technology is summarized by the *matching function* $m(x) = \frac{cx^2}{2}$ that represents the expenditure by the advertiser required to generate a contact of intensity x. The complete-information profits from generating a contact of intensity x with a consumer of value v are $\pi(v, x) = vx - \frac{cx^2}{2}$.

A data provider knows the match value v_i of each consumer i, and sells this data to the advertiser at a constant price per individual p. Hence, if the advertiser acquires information about consumer i, it is able to tailor its choice of contact intensity to the match value, i.e., $x^*(v_i) = v_i/c$, and obtain profits of $\pi^*(v_i) = \frac{v_i^2}{2c}$. In contrast, the advertiser must choose a constant intensity level $\bar{x}$ for all other consumers. Because the constant intensity level $\bar{x}$ depends on the composition of the *residual set* of consumers, the advertiser's information-acquisition problem can be formulated as the choice of a *targeted set* of consumers $A \subset V$.

The demand for information about specific consumers can be traced back to two sources of *mismatch risk*: excessive vs. insufficient advertising. Specifically, the profit-maximizing *residual* set for the advertiser in this case is a nonempty interval, $A^{\mathrm{C}} = \left[\frac{1}{2} - 2\sqrt{cp}, \frac{1}{2} + 2\sqrt{cp}\right]$. In other words, the advertiser purchases information about both very high- and very low-value consumers. Note that $\mathrm{E}\left[v_i | v_i \in A^{\mathrm{C}}\right] = \frac{1}{2}$, which implies $\bar{x} = \frac{1}{2c}$. This is *too much* advertising for the consumers in the bottom half of the residual set and *too little* advertising for the consumers in the top half.

Finally, if the data provider incurs no *marginal* cost of supplying information, then it chooses p to maximize $p(1 - 4\sqrt{cp})$; i.e., it sets $p^* = \frac{1}{36c}$. The advertiser thus purchases data only on the bottom and top sixths of the market, treating all other consumers as if they had match value $\frac{1}{2}$.

2.3.3 Marketing Techniques

Some studies expand the analysis of privacy to include the costs of intrusions into an individual's personal sphere, such as unsolicited mail or spamming, as in Hann et al. (2008), and personal pref-

agent's payoff by at most a small amount. Other related mechanism design issues have been studied, such as the issue of limiting "exposure," where agents internalize being exposed to the realized types and chosen actions of a subset of other agents (Gradwohl and Reingold, 2010) or to the party responsible for implementing the mechanism (Gradwohl, 2015), and the issue of "anonymity," where agents may seek to participate in a mechanism multiple times when anonymizing is too easy (Wagman and Conitzer, 2014).

erences over privacy, as in Tang et al. (2008). Here, the theoretical study of privacy connects with the marketing literature on couponing, market segmentation, and consumer addressability (Blattberg and Deighton, 1991). Work by Shaffer and Zhang (1995, 2002), Chen et al. (2001), Chen and Iyer (2002), and Conitzer et al. (2012) obtain complementary results. These authors show that when a firm has control over consumers' privacy, it chooses to segment the population optimally for pricing purposes. Their findings demonstrate that price discrimination can lead to intensified price competition, where firms may possess incentives to (i) decrease the level of accuracy of targeted promotions, (ii) differentially invest in customer addressability, and (iii) seek commitment mechanisms not to price discriminate.

Hann et al. (2008) study a competitive market with heterogeneous consumers, some who draw no benefit from unsolicited marketing and some who are interested in receiving information about new products. They show that attempts to use technologies that prevent unsolicited marketing on one side, and sellers' efforts to use direct marketing on the other, constitute strategic complements: the higher the attempts of consumers to protect themselves from unsolicited marketing, the higher the use of direct marketing by sellers. In related work, Chellappa and Shivendu (2010) examine the trade-offs vendors and consumers face between privacy concerns and the personalization of services and products that the sharing of data may make possible, and Hui and Png (2006) consider the use of private information for unsolicited marketing — in person, via telephone, mail or email — which competes with the marketing efforts of other companies and may inconvenience individuals.

Anderson and de Palma (2012) look at spamming as a problem of competition among senders (of messages) for the receivers' attention, which is a limited resource. Their model considers the costs that both parties have to incur in order to arrive at a transaction. These costs endogenously determine the number of messages sent by the sender and the number of messages read by receivers. If the cost of sending messages is too low, there will be a congestion problem, meaning that receivers will only read some of the messages sent (see, also, Alstyne, 2007). In this case, a welfare-enhancing solution may be to add a small tax on the transmission of a message. Such a tax may increase surplus, because senders who send messages of low quality will be crowded out (it would be too costly for them to send a message), fewer messages will be sent, and more will be read. Spiegel (2013) identifies conditions under which firms may choose to bundle new software with advertisements and distribute it for free as adware. While adware is more affordable to consumers and may contain advertisements that help improve their purchasing decisions, it also entails a loss of privacy.

Linking the above works to the study of privacy, Zandt (2004), Armstrong et al. (2009), Anderson and de Palma (2012), and Johnson (2013) also investigate the topic of congestion due to consumers having limited attention. In their models, consumers can choose to "opt out" from receiving sellers' marketing. The result is a form of a Prisoner's Dilemma situation: while each consumer has a private incentive to opt out of intrusive marketing, when all consumers do this, price competition is relaxed and consumers are harmed. Targeted ads, however, can also be counterproductive, if they trigger the recipient's privacy concerns, or her worries regarding the level of control over her private information (Tucker, 2014). In this sense, targeted advertising is a form of unsolicited marketing. While spamming involves the indiscriminate sending of advertisements, targeted advertising (or behavioral targeting), as the name suggests, consists of contacting a select group of recipients who, according to the information available to the sender about their previous behaviors or preferences, may be particularly interested in the advertised product or service.

Hoffmann et al. (2013) study targeted communications, a practice they refer to as hypertargeting, in the context of marketing and political campaigns. In a departure from the earlier literature on strategic disclosures (Grossman and Shapiro, 1984; Milgrom, 1981; Milgrom and Roberts, 1986), they assume that firms must be selective when choosing the amount of information they communicate to consumers (e.g., due to space or time constraints). Since consumers differ in their preferences, firms may wish to market different product attributes to different consumers. They model hypertargeting as the selective disclosure of information to a specific audience, and characterize the private incentives and welfare impact of hypertargeting. They demonstrate that a privacy policy that hinders hypertargeting by, for instance, banning the collection of personally identifiable data, is beneficial when consumers are naïve, competition is limited, and firms are able to segment the market to price discriminate. Otherwise, privacy regulation may backfire, because a policy that, for instance, requires consumer consent, can allow firms to commit to abstain from selective targeting — even when doing so would benefit consumers.

The following example demonstrates that the common wisdom that imposing a tax on messages will fall more heavily on spammers and thereby improve the average quality of contacts need not necessarily be true.

Example 5 (Marketing and Spam)**.** Suppose there are two firms, a spammer (firm 0) and a retailer (firm 1). There is a consumer who has time to open exactly one email message. If she opens a message from the retail firm she receives a payoff of $v > 0$ and the retailer receives gross profit of b_1, and if she opens a message from the spammer she receives $-k < 0$ and the spammer receives $b_0 \in (0, b_1)$. Also assume $b_1 v > b_0 k$. The expected profit to firm i from sending the consumer m_i messages when its rival sends her m_j is

$$\Pi_i = \frac{m_i}{m_i + m_j} b_i - c m_i, \quad i \in \{0, 1\},\ i \neq j,$$

where c is the marginal cost of sending a message. The expected payoff to the consumer from opening a single message at random is

$$U = \frac{m_1 v - m_0 k}{m_1 + m_0}.$$

It is straightforward to verify that in the unique Perfect Bayesian equilibrium of this game the following must hold: Firm i sends

$$m_i^* = \frac{b_j b_i^2}{c(b_1 + b_0)^2}$$

messages and receives expected profit

$$\Pi_i^* = \frac{b_i^3}{(b_1 + b_0)^2},$$

and the consumer receives expected payoff

$$U^* = \frac{b_1 v - b_0 k}{b_1 + b_0}.$$

Observe that charging the firms a tax of t per message (resulting in marginal cost of $c+t$) reduces the number of messages sent by each firm, but has no impact on equilibrium payoffs. The firms would respond to a tax by sending proportionally fewer messages, reducing the absolute number received by the consumer, but not their composition. In contrast, a filter that correctly identifies a fraction ϕ of the messages sent by firm 0 as spam both reduces the number of messages sent by each firm in equilibrium and raises the consumer's expected payoff to

$$U^{**} = \frac{b_1 v - (1-\phi) b_0 k}{b_1 + (1-\phi) b_0}.$$

3 The Empirical Analysis of Privacy

If our perusal of the theoretical economic literature on privacy has revealed one robust lesson, it is that the economic consequences of less privacy and more information sharing for the parties involved (the data subject and the actual or potential data holder) can in some cases be welfare enhancing, while, in others, welfare diminishing. The various streams of research we covered highlighted that, in choosing the balance between sharing or hiding personal information (and in choosing the balance between exploiting or protecting individuals' data), both individuals and organizations face complex, often ambiguous, and sometimes intangible trade-offs. Individuals can benefit from protecting the security of their data to avoid the misuse of information they share with other entities. However, they also benefit from the sharing of information with peers and third parties that results in mutually satisfactory interactions. Organizations can increase their revenues by knowing more about the parties they interact with, tracking them across transactions. Yet, they can also bear costs by alienating those parties with policies that may be deemed too invasive. Intermediaries can increase their revenues by collecting more information about users, yet offering overly precise information to advertisers can backfire by reducing competition among sellers.

This section complements the previous one by surveying the empirical literature on the economics of privacy to highlight the costs and benefits of privacy protection and information sharing.

The market for personal data and the market for privacy are two sides of the same coin, wherein protected data may carry benefits and costs that mirror or are dual to the costs and benefits associated with disclosed data for both data subjects and data holders. For instance, *disclosed* personal information (or lack of data protection) can result in economic benefits for both data holders (savings, efficiency gains, surplus extraction, increased revenues through consumer tracking) and data subjects (personalization, targeted offers and promotions, etc). At the same time, such disclosures (or, the *lack* of protection of personal data) can be costly for both firms (costs incurred when data is breached or misused, or collected in ways that consumers deem too intrusive) and consumers (from tangible costs such as identity theft or (price) discrimination, to less tangible ones such as stigma or psychological discomfort; see, e.g., Stone and Stone, 1990). Furthermore, the *act* of collecting data can be costly for data holders (such as the investments necessary to establish Customer Relationship Management systems).

Similarly, protected data (or, *lack* of data disclosure) can be associated with both benefits and costs for data subjects and potential data holders; such benefits and costs are often dual (that is, the inverse) of the benefits and costs highlighted above. For instance, data subjects and data holders may incur opportunity costs when useful data is not disclosed (for instance, they may miss out on opportunities for increased efficiency or increased convenience), although both parties may also benefit in various ways (consumers, for example, by reducing the expected costs associated with identity theft; firms, for example, by exploiting privacy-friendly stances for competitive advantage). Furthermore, there are costs associated with the *act* of protecting data (investments necessary to encrypt data for the data holders to prevent *further* disclosures; costs of using privacy-enhancing technologies for the data subject, etc).

In short, there can be many dimensions to privacy harms (Ryan, 2011) and to the benefits arising from personal information. The rest of this section does not attempt to provide a comprehensive enumeration of those dimensions, but surveys the areas that have attracted more empirical economic analysis.

3.1 Privacy, Advertising, and Electronic Commerce

Online advertising is perhaps the most common example of how firms use the large amounts of data that they collect about users. The greater availability of personally identifiable data on the Internet in terms of scope, quantity, and the precision with which firms can target specific users challenges the traditional distinction between personal selling and remote communication. As a result, the targeting of advertisements affects marketing strategies and competition between online and offline media (see, for instance, Athey and Gans, 2010; Bergemann and Bonatti, 2011; Athey et al., 2013). Evans (2009) documents that 56 of the top 100 websites based on page views, accounting for 86% of page views for that group, presented some form of advertising in February 2008 and likely derived most of their revenues from doing so. According to IAB estimates, $36.6 billion was spent on digital ads in 2012, ahead of cable TV ($32.5 billion) and slightly below broadcast TV ($39.6 billion),[9] with its growth outpacing all other formats. The market capitalization of the major publicly traded newspaper businesses in the US declined by 42% between January 2004 and August 2008, compared to a 15.6% percent gain for the Dow Jones industrial average over that time period.

Key to the online collection of consumer information are the aspects of "targetability," the collection of data for the purpose of showing ads to specific subsets of users, and "measurability," the collection of data for the purpose of evaluating the efficacy of targeted ads. Data aggregators, advertising networks, and website operators establish relationships to enable them to track and target users across different websites and over time. Advertisers take advantage of the enhanced performance measurability of online advertising to experiment with different marketing messages before proceeding with a specific marketing campaign (cf. Lewis and Reiley, 2013). Advertisers and website operators can track user behavior using several techniques — from web bugs (also known as beacons),[10] to cookies, to browser and device fingerprinting.[11] Despite the large sums of money spent on targeted advertising, its effectiveness is unclear.[12] According to Farahat and Bailey (2012), targeted advertising generated, on average, twice the revenue per ad as non-targeted advertising. Beales (2010) documents that in 2009 the price of behaviorally-targeted advertising was 2.68 times the price of untargeted advertising. Lambrecht and Tucker (2013), however, find evidence indicating that personalized advertising may be ineffective, but becomes slightly less so when combined with data on consumers' mindsets. Blake, Nosko, and Tadelis (Blake et al.) reinforce their findings. They measure the effectiveness of paid search by running a series of large-scale field experiments on eBay, and find evidence that returns from paid search are a fraction of conventional non-experimental estimates (and can, in some cases, be negative). Targeted advertising, in principle, could provide consumers with information about products they want or are interested in, thereby reducing search costs and, theoretically, improving welfare. However, as the theoretical literature examined in Section 2.3.3 suggests, the effect of targeting can be rather complex and nuanced, and not necessarily always positive for consumers.

Along with being targeted with personalized offers, consumers may also face, for instance, price discrimination (tracking and measurability, in addition to websites' ability to dynamically update and personalize prices for each visitor, are bringing online markets closer to the theoretical scenario of first-degree price discrimination); or, they may be offered products inferior to the ones they

[9] http://www.iab.net/media/file/IAB_Internet_Advertising_Revenue_Report_FY_2012_rev.pdf

[10] Web beacons are small pieces of code placed on websites, videos, and in emails that can communicate information about a user's browser and device to a server. Beacons can be used, among other things, for website analytics or to deliver a cookie to a user's device.

[11] Fingerprinting refers to technologies that use details about a user's browser and device in order to identify the user's browser or device over time. Fingerprinting can be used for the same purposes as cookies, but does not require files to be stored on a user's device.

[12] In the US alone, the internet advertising industry recorded revenues totaling $36.6 billion in 2012, up 15.2% from 2011. Google alone registered $46 billion in global revenues in 2012, of which a reported $43.7, or 95%, were attributed to advertising. See http://www.iab.net/media/file/IAB_Internet_Advertising_Revenue_Report_FY_2012_rev.pdf and http://investor.google.com/financial/tables.html.

would have found otherwise, or even potentially damaging ones.[13] Additionally, concerns exist over the fact that tracking technologies are intentionally invisible to end-users (Smith, 1999), whereby a significant lack of awareness and misconception exists among consumers regarding the extent, nature, and depth of targeting techniques (McDonald and Cranor, 2010). Even sophisticated consumers may not be able to avoid being tracked, as the advertising and data industry has often found new ways of tracking and identifying users after consumers had learned about, and found measures to counter existing forms of tracking (Hoofnagle et al., 2012).

Web bugs allow advertisers to track consumers' browsing activities and gain insight into their interests. They are different from cookies because they are designed to be invisible to the user and are not stored on a user's computer. In particular, without inspecting a webpage's underlying code, a customer does not know that they are being tracked. Murray and Cowart (2001) found that 96% of websites that mentioned a top 50 brand (as determined by the 2000 Financial Times rankings) contained a web bug.

On the other hand, concerns exist that the introduction of strict privacy regimes may inhibit the development of electronic commerce (Swire and Litan, 1998). European countries have raised barriers to the collection and use of personally identifiable data, including a requirement that firms seek explicit consent from consumers to collect information about their past purchases and recent browsing behavior. The European ePrivacy Directive (2002/58/EC) predominantly addresses the telecommunications sector, but several of its provisions limit the ability of companies to track user behavior on the Internet.[14] The EU Privacy Directive (Recital 24) explicitly states that "[s]o-called spyware, web bugs, hidden identifiers, and other similar devices can enter the user's terminal without their knowledge in order to gain access to information, to store hidden information or to trace the activities of the user and may seriously intrude upon the privacy of these users. The use of such devices should be allowed only for legitimate purposes, with the knowledge of the users concerned." With regards to cookies, the Directive permits their use on the condition that users provide their opt-in consent, and the Directive does not explicitly address clickstream data.[15] The 2012 Do-Not-Track proposal by the Federal Trade Commission (FTC) suggests a set of guidelines

[13] For instance, data brokers sell lists of consumers to target individuals suffering from addictions such as alcoholism or gambling (see, e.g., `http://www.dmnews.com/media-one-gamblers-database/article/164172/`).

[14] See the Data Protection Directive (1995/46/EC) and the Privacy and Electronic Communications Directive (2002/58/EC), also known as the ePrivacy Directive, which regulates cookies and other similar devices through its amendments such as Directive 2009/136/EC, the so-called EU Cookie Directive, and the Privacy and Electronic Communications (EC Directive) (Amendment) Regulations 2011. The current prescription is that cookies or similar methods must not be used unless the subscriber or user of the relevant terminal equipment: (a) is provided with clear and comprehensive information about the purposes of the storage of, or access to, that information; and (b) has given his or her consent.

[15] Clickstream data describes the information a firm can collect about its customers based on basic browsing behavior, including the pages a user viewed at a website, how long they spent on each page, the visitor's browsing path, IP address, and page of origin. In contrast to web bugs and cookies, the EU Privacy Directive does not explicitly address restrictions on the use of clickstream data (Garrie and Wong, 2006). However, since such data may be used to identify users (particularly given the prevalence of static IP addresses), firms may be bound by existing laws to obtain consent (Baumer et al., 2004).

similar to the EU ePrivacy Directive for the US — giving consumers a way to opt out of having their data collected. However, as the FTC's proposal follows an opt-out rather than an opt-in approach, it may be less costly for US firms to continue collecting data and using targeted ads.[16]

Goldfarb and Tucker (2011b) examine the effects of the implementation of the EU Privacy Directive on hypothetical purchase intentions, and find some evidence that after the Privacy Directive was passed, advertising effectiveness decreased significantly. They use the responses of 3.3 million survey-takers who had been randomly exposed to 9,596 online display (banner) advertising campaigns to explore how strong privacy regulation in the European Union influenced the effectiveness of advertising. For each of the 9,596 campaigns, their data contains a treatment group exposed to the ads and a control group exposed to a public service ad. To measure ad effectiveness, they use a short survey conducted with both groups of users about their purchase intent towards an advertised product. They find that following the ePrivacy Directive, banner ads experienced a reduction in effectiveness of over 65% in terms of changing consumers' purchase intent. They see no similar change in ad effectiveness in non-European countries during a similar time frame. These findings raise a number of stimulating questions. One interpretation of the results is that privacy regulation can have a detrimental effect on the advertising industry. Moreover, as online advertising has become a primary source of revenue for many web-based businesses, the types of content and service provided on the Internet may shift as a result of privacy regulation. However, the decrease in hypothetical advertising effectiveness was only found within a subset of ads (static, content-specific, and small) whereas other types of ads were not at all affected (larger ads, dynamic ads, and ads consistent with the content of a website). This suggests a possible way forward for organizations — for instance, general interest websites may fine-tune ads based on the content of a specific webpage to make it easier to monetize, providing ads that are more contextually appropriate. Another interpretation of the results is that the legislation worked as intended (advertisers no longer could access or use certain consumer data), although no actual economic damage may have been ultimately incurred by consumers, or by merchants as a whole: if behavioral targeting's main value was about inducing consumers to buy from a given merchant, rather than another one (as opposed to informing consumers about products or services of which they would not have otherwise been aware), then holding budget constraints and preferences over goods constant, consumers may still buy products that have not been targeted to them.

Related work has addressed the question of which format of ads advertisers should and should not use. White et al. (2008) find that consumers may experience "personalization reactance" by negatively reacting to highly personalized messages when the fit between the targeted offer and consumers' personal characteristics is not explicitly justified. Goldfarb and Tucker (2011a) find that obtrusive targeted ads — targeted in the sense that they are matched to the content of a

[16] For instance, Johnson and Goldstein (2004) show that consumers tend to go with the default option chosen for them in the case of organ donation, despite heavy lobbying by organizations. That is, if the default approach is for consumers to be subscribed to targeted ads, then more consumers are likely to remain subscribed than under an opt-in system where consumers are by default unsubscribed.

website, and obtrusive in terms of visibility — are more likely to trigger privacy concerns among users in comparison to obtrusive but not targeted ads, or targeted but less obtrusive ads. These findings can help explain the enormous success of Google AdSense for Content, a service which provides contextually-targeted unobtrusive ads (AdSense accounts for about a third of Google's ad revenue, with the other two thirds coming from search advertising).[17] Moreover, it can help explain the apparent divide in online advertising between banner ads and unobtrusive targeted ads.

3.2 Privacy and Price Discrimination

Much of the literature surveyed in Section 2.3.1 focuses on merchants' ability to engage in forms of targeting that increasingly approach the textbook "ideal" of first-degree discrimination. Relative to the volume of theoretical analyses, empirical efforts to find evidence of Internet-based price discrimination have lagged behind. Valentino-Devries et al. (2012) suggest that certain online retailers may be engaging in dynamic pricing based on their ability to estimate visitors' locations, and, specifically, the (online) visitor's physical distance from a rival brick-and-mortar store. Mikians et al. (2012, 2013) find suggestive evidence of price discrimination based on information collected online about consumers, as well as evidence of "search discrimination" (steering consumers towards different sets of products, with different prices, following their searches for a certain product category). In particular, Mikians et al. (2013) suggest price differences of 10% to 30% for identical products based on the location and the characteristics (for instance, browser configurations) of different online visitors.

On the other hand, Vissers et al. (2014) find price *variation*, yet fail to uncover experimental evidence of consumer-based price discrimination in online airline tickets. In short, the evidence of systematic and diffuse individual online price discrimination is, currently, scarce. It is possible that firms may consider online price discrimination as not just challenging, but potentially risky,[18] yet anecdotal cases of firms selectively offering price discounts are ubiquitous (i.e., instead of raising prices to some consumers, firms may simply reframe their behavior by offering price discounts to others). It is also possible that the infrastructure for accurate price discrimination (and its detection) is yet underdeveloped — similarly to the case of behavioral ads (which, anecdotally, seem as likely to present consumers with offers of products they have already searched for or even bought, rather than undiscovered products in which they may be interested) — but still evolving.

3.3 Other Forms of Discrimination

Price discrimination is probably the least odious form of discrimination involving the use of personal information. In many other markets, nuanced trade-offs can arise as function of the amount of

[17] See `https://investor.google.com/earnings.html`. It is also worth noting that contextual targeting is also common in the offline world (e.g., magazines about fishing contain ads for fishing equipment).

[18] Consider the above-mentioned backlash following Amazon's attempts at price discrimination — which Amazon claims was only experimentation (Anderson and Simester, 2010).

personal information available to other parties, including scenarios where privacy protection will cause, or in fact hinder, discrimination.

Consider, for instance, hiring. Economists have long been interested in the role of information (Stigler, 1962) and signaling (Spence, 1973) in job market matching, and there exists, of course, a vast economic literature on discrimination in hiring or wages. Some of the experimental evidence shows that employers can infer candidates' personal traits from information available on their resumés (such as the candidates' race, from their names), and use that information to discriminate among prospective employees (Bertrand and Mullainathan, 2004). In fact, fairer job market outcomes may sometimes be achieved after *removing* information from the marketplace. Goldin and Rouse (2000), for instance, find evidence that "blind" auditions (in which screens conceal the identities of candidates, such as orchestra performers, from the jury) foster impartiality in hiring and increase the probability that a woman will be hired. On the other hand, Strahilevitz (2008) and Bushway (2004) point out a different dynamic: when employers are not able to retrieve pertinent information about a job applicant (for instance, their criminal records) due to privacy regulation or protection, the employer may end up increasingly reliant on statistical discrimination strategies, and an employer's animus or bias may disproportionately affect certain minorities. Under this scenario, expanding the information that is available to employers may generally lead to fairer and possibly more efficient outcomes.

Consider, also, online platforms that allow tenants to find landlords and vice versa, or platforms that enable property owners to "share" their houses with short-term renters, or platforms that enable car owners to share their vehicles with other drivers or passengers. These examples of IT-enabled "sharing economies" may increase efficiency by improving how resources such as housing or vehicles are used. However, when these platforms also expose members' personal information, they may inadvertently foster discrimination: using data from Airbnb (an online rental marketplace), Edelman and Luca (2014) find that New York City landlords who are not African American charge approximately 12% more than their African American counterparts for an equivalent rental (the race of an Airbnb landlord can often be inferred from profile photos on landlords' accounts).

Anther example where expanding the amount of information that is available to the marketplace may influence discrimination concerns employment opportunities for individuals with criminal records. Because of the stigmatizing effect associated with a criminal history, individuals with criminal records are more likely to experience job instability and wage decline (see, for instance, Waldfogel, 1994; Nagin and Waldfogel, 1995), and evidence suggests that employers do use criminal records to screen candidates (e.g., Bushway, 2004). Information technology has exacerbated the problem: large numbers of criminal histories are now computerized in state repositories and commercial databases. Thus, ex-offenders may be trailed by their crime histories wherever they may apply for jobs. This can occur despite a criminal record being, at some stage, "stale." This is in contrast to Blumstein and Nakamura (2009), who point out that the likelihood of recidivism, or a person's relapse into criminal behavior, declines with time spent without committing a crime, and at a certain point in time, an ex-offender who has remained "clean" can be regarded as providing no greater risk than a non-offender counterpart of the same age.

In many countries, legislators are acutely interested in these problems, and finding the right balance of sharing and protection of personal information is a (thorny) matter of public policy. For instance, in the US, several states authorize courts to expunge or seal certain criminal records — but only for certain types of arrests and convictions. Similarly, in most of the US, an employer who asked about the religion of a job candidate would risk being sued under Equal Employment Opportunity (EEO) laws; however, different types of information enjoy different protections (for instance, information regarding religious affiliation cannot even be inquired about in interviews, whereas other types of personal information may, theoretically, be enquired about, but should not actually be used in decisions concerning hiring or wages).

Information technology, however, has once again created new challenges in this context: many job candidates nowadays publicly provide personal information through social-network profiles, including information such as sexual orientation or religious affiliation, which may be actually protected under state or federal laws. Employers are supposed not to ask about such information during the hiring process — but searching for it online significantly reduces the risk of detection. Acquisti and Fong (2013) have investigated the role of social media in the hiring behavior of US firms. In the authors' experiment, they create online social-media profiles for job candidates and then submit job applications on behalf of those candidates to a sample of over 4,000 US employers. If an employer were to search online for the name found in the resumè and application it received, it would find the social media profile of the candidate and be exposed to the experimental manipulation. Acquisti and Fong estimate that only a (sizable) minority of US employers likely searched online for the candidates' information, and that the overall effect of the experimental manipulations was small. However, they did find evidence of both search and discrimination among a self-selected set of employers. In this, as in other scenarios, it is still unclear the extent to which novel information channels will reduce market frictions and increase efficiency, or in fact promote new forms of discrimination.

3.4 Privacy and Health Economics

Privacy regulation may affect the extent and direction of data-based technological progress. For instance, trade-offs surrounding electronic privacy protection may arise in the context of technology adoption in the medical industry, as many new technologies depend on information exchange (Schwartz, 1997). The 1996 US Health Insurance Portability and Accountability Act (HIPAA) established some rules for privacy in healthcare. The 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act, part of the American Recovery and Reinvestment Act, devoted $19.2 billion to increasing the use of electronic medical records (EMRs) by healthcare providers. Innovations in digitizing health information lead to quality improvements in the health sector, because they make patient information easy to access and share.

Electronic medical records, for instance, allow medical providers to store and exchange patient information using computers rather than paper records. Hillestad et al. (2005) suggest that EMRs could reduce annual US healthcare costs by $34 billion through greater efficiency and safety, assuming a 15-year period and 90% EMR adoption. Miller and Tucker (2007, 2009, 2011a, 2014a) provide evidence quantifying the effect of state privacy protection on the diffusion of EMRs. Using variations in medical privacy laws across states and across time, they show that privacy regulations restricting a hospital's release of patient information significantly reduced the adoption of electronic medical records, primarily due to diminished network effects in adoption. Their analysis suggests that state privacy regulation restricting the release of health information reduces aggregate EMR adoption by more than 24%. They further estimate that a 10% increase in the adoption of such systems can reduce infant mortality by 16 deaths per 100,000 births. Although EMRs were invented in the 1970s, by 2005 only 41% of US hospitals had adopted a basic EMR system (Goldfarb and Tucker, 2012a). In the European Union, personal data recorded in EMRs must be collected, held, and processed in accordance with the Data Protection Directive (95/46/EC). In the US, hospitals may hesitate to adopt EMR systems if (i) they are concerned about their patients' response, and (ii) if regulation intended to protect patient privacy ends up hindering the adoption of such systems — because hospitals cannot properly utilize them by, for instance, exchanging patient information with other hospitals.

Miller and Tucker identify two schools of thought about the interplay of privacy concerns, regulation, and technological innovation. The first holds that regulatory protection inhibits technology diffusion by imposing costs upon the exchange of information. In addition to these trade-offs, hospitals are faced with complexities concerning state-specific regulation and information exchange across state lines. The second, instead, argues that explicit privacy protection promotes the use of information technology by reassuring potential adopters that their data will be safe. Consider a possible example of the latter dynamics, in the context of Health Information Exchanges (HIEs). HIEs are information-technology solutions that facilitate the sharing of patients' electronic medical records. They are expected to enhance information sharing capabilities among healthcare entities, with the aim of improving the quality of care. Their adoption, however, is said to have been hindered by privacy concerns, and it is unclear how privacy laws, such as legislation restricting the disclosure of health records, impact their adoption. In the US, state laws may incentivize HIE efforts, include specific privacy requirements for sharing healthcare data, or both. Adjerid et al. (2013) investigate the impact of state laws on the emergence of HIEs. They compare the adoption and success of HIEs in states with laws that limit information disclosure with states that do not have such laws. The authors find that the combination of adoption subsidies and stronger privacy protection (that is, legislation that included strict requirements for patients' consent in order to use their medical data) is associated with greater HIE adoption than either under privacy protection alone, or, importantly, under subsidies alone. Their results suggest that there can be important policy complementarities between privacy laws and other types of interventions (such as financial subsidies and technical assistance in the case of HIEs), and that different degrees of privacy regulation can have different effects on technology innovation and economic welfare. Regulators may thus find room for balancing meaningful privacy protection while incentivizing the adoption of information-technology efforts.

Oster et al. (2010) use data from a prospective cohort study of approximately 1000 individuals at risk for Huntington disease (HD), a degenerative neurological disorder with significant effects on morbidity, to estimate adverse selection in long-term care insurance. They find strong evidence of adverse selection: individuals who carry the HD genetic mutation are up to 5 times more likely than the general population to own long-term care insurance. Other genes, such as those associated with increased risks of breast cancer, colon cancer, Parkinson and Alzheimer diseases, among others, have also been identified, and testing for these genes is becoming more common and more precise (Burton et al. 2007). This testing, in turn, is likely to increase the amount of private information stored about individuals. On the one hand, this information may be useful in developing treatments, vaccines, and immunizations. On the other, while US laws limit an insurer's ability to observe an individual's specific genetic information, marketers (e.g., for certain drugs and treatments) and advertising platforms may certainly be interested in it. A number of companies aim at offering genetic testing to individuals at affordable rates.[19]

Using data on genetic testing for cancer risks, Miller and Tucker (2014b) examine how state genetic privacy laws affect the diffusion of personalized medicine. They identify three approaches taken by states to protect patients' genetic privacy: requiring informed consent; restricting discriminatory usage by employers, health care providers or insurance companies; and limiting *redisclosure* without consent. They show evidence that the redisclosure approach encourages the spread of genetic testing, in contrast to the informed consent approach, which deters it.

The results in both Miller and Tucker (2014b) and Adjerid et al. (2013) illustrate how privacy laws need to be tailored to take into account and balance specific and continually evolving trade-offs, and how rather than looking at privacy regulation in a binary, monotonic fashion, the effect of regulation on technology efforts can be heterogeneous — depending on the specific requirements included in the legislation. Genetic and genomic analyses may not only reveal information about an individual's current health, but also about future health risks, and this potential to reveal information is likely to expand. These analyses are useful for patients and health care providers because they facilitate the delivery of personalized medicine. At the same time, as personal genetic and genomic information becomes increasingly available, consumers face new privacy risks — for instance, if such information reaches the hands of advertising platforms and data aggregators, they may use it to construct risk profiles for individuals and their biological relatives such as children

[19] For instance, 23andme (`https://www.23andme.com/`) did so before the US Food and Drug Administration (FDA) directed them to cease while they undergo a regulatory review process. Prior to their primary operations being halted, they would store individuals' DNA, and offer updates on potential health issues as testing procedures advanced. Currently, they do so for a subset of of potential health conditions — those for which they received FDA approval.

and parents, combine it with other data, and improve their targeting of product offerings. Adding another angle to this is Miller and Tucker's finding that genetic privacy laws have distinct effects beyond standard health data privacy laws — in particular, different laws may alter individual behavior.[20] Yet another trade-off in this legislative balancing act is the fact that wider access to genetic and genomic analyses can lead to broader improvements in overall health care.

Another example of balancing health-related informational privacy with public benefit is the identification of infectious disease outbreaks — the reporting of these cases to state authorities is usually exempt from privacy restrictions. Rapidly identifying such outbreaks is critical for the effective initiation of public-health intervention measures, for preparation and readjustment of vaccines, and for the timely alerting of governmental agencies and the general public. Google Flu Trends, for instance, takes advantage of users' searches for influenza-related terms to provide both public-health professionals and the general population with real-time, geographically-specific view of influenza search activity in the US. Other monitoring services include the International Society for Infectious Diseases' Program for Monitoring Emerging Diseases and HealthMap. However, in comparison to data provided by users online, traditional surveillance methods are costlier and slower (Wilson and Brownstein, 2009). At the same time, the public availability of outbreak-related information may create challenges: effective risk communication by public officials, drawing the proper conclusions from the data, and privacy concerns. In the case of search engines, data is collected about individual users using a cookie, an IP address, and other methods. Associated with this profile are the search queries and subsequent clicks made by each user. Currently, Google is said to keep this information for 9 months (18 months in the case of cookies) and anonymize it afterwards. Microsoft is said to keep this information for 6 months. In practice, there is no real verification of whether data holders in fact delete or at least anonymize user information. Moreover, some regulatory bodies may require data holders to retain this user information for a period of up to 2 years for the purpose of law enforcement (see, for instance, the European Data Retention Directive 2006/24/EC).

3.5 Privacy and Credit Markets

In the US, credit reporting was not regulated at the federal level until 1970, when the Fair Credit Reporting Act (FCRA) was legislated—an act that was subsequently amended several times. Currently, the credit-reporting industry is among the most regulated in terms of data protection. The FCRA established permissible purposes of credit information disclosure, codified information flows along the lines that they had naturally developed in the market, introduced dispute settlement

[20] This finding is reminiscent of Johnson and Goldstein (2004), who, as previously mentioned, show that consumers tend to go with the default option chosen for them in the case of organ donation, despite heavy lobbying by organizations. That is, if the default approach is for consumers to disclose genetic information to the immediate health-care provider along the lines of the "redisclosure" approach, then more consumers are likely to accept the service than under the opt-in system of the "informed consent" approach.

mechanisms and data correction procedures, and assigned expiration dates to negative information such as bankruptcy and payment defaults. Several information flows, such as those among non-affiliates, were left unregulated at the federal level, although some states enacted their own regulations (Jentzsch, 2003). The 1990s brought major reforms in the US that were intended to strengthen financial privacy laws, in light of intensifying public debate about privacy erosion given advancements in information technology. The Consumer Credit Reporting Reform Act (CCRRA) of 1996, for the first time, introduced duties for financial information providers. In order to correct any inaccuracies in consumers' records, the CCRRA mandated a two-sided information flow to/from credit bureaus and providers, and formalized some information flows among affiliates.

The Gramm-Leach-Bliley (GLB) Act of 1999 extended the CCRRA by formally and legally allowing a variety of financial institutions to sell, trade, share, or give out nonpublic personal information about their customers to non-affiliates, unless their customers direct that such information not be disclosed by opting out. The GLB Act, while granting consumers the option to opt out, restricts it to non-affiliates. An affiliate is defined as any company that controls, is controlled by, or is under common control with another company. Consumers have limited if any power to restrict this kind of "corporate family" trading of personal information. There are also several other exemptions under the GLB Act that can permit information sharing despite a consumer's objection. For instance, if a financial institution wishes to engage the services of a separate company, they can transfer personal information to that company by arguing that the information is necessary to the services that the company will perform. A financial institution can transfer information to a marketing or sales company to sell new products or jointly offered products. Once this unaffiliated third party has a consumer's personal information, they can share it within their own "corporate family." However, they themselves cannot likewise transfer the information to further companies through this exemption. In addition, financial institutions can disclose users' information to credit reporting agencies to comply with any other laws or regulations.

For lenders, the extension of credit to borrowers depends on the acquisition and possibly the exchange of personal information among market participants. Jentzsch (2003) develops a Financial Privacy Index to quantify the extent of information protection across different regimes, demonstrating that the US grants less data protection than EU members. The primary concern is that more stringent data-protection regulations may lead to reduced access to credit, thus creating a trade-off with consumer privacy. In line with this work, Pagano and Jappelli (1993); Jappelli and Pagano (2002) predict that if banks share information about their customers, they would increase lending to safe borrowers, thereby decreasing default rates. Other empirical studies tend to focus on the effects of credit bureaus and creditor rights using data from a cross-section of countries (see, e.g., Djankov et al., 2007; Qian and Strahan, 2007).

States and local municipalities can enact legislation and local ordinances that exceed the protections in the GLB Act, to require, for instance, opt-in consent — as is the case in a subset of the Bay Area counties examined in Kim and Wagman (2015). In a study that directly bridges the theoretical and empirical analyses of privacy, Kim and Wagman incorporate information acquisition and privacy regulation — through restrictions on information trade — into a model of consumer screening. In their model, firms, such as mortgage lenders, compete in prices. Lower prices, however, entail more stringent screening of applicants. They show that, in equilibrium, consumers apply to obtain loans from firms posting the lowest prices, despite anticipating more stringent loan approval processes. They then demonstrate that enabling firms to sell applicant data to interested downstream parties, such as insurers, can lead to even lower prices, higher screening intensities, and higher rejection rates of applicants; however, social welfare, overall, increases.

One of the main criticisms of the GLB Act's privacy provisions has been that most consumers do not (and likely will not) take advantage of the *opt-out* option to request that a firm ceases trade in their information. In 2002, three out of five counties in the San Francisco-Oakland-Fremont, CA, Metropolitan Statistical Area enacted a local ordinance (effective January 1, 2003) that is more protective than then-current practices by pursuing an *opt-in* approach. Specifically, the local ordinance would require financial institutions to seek a written waiver before sharing consumer information with both affiliates and non-affiliates. The variation in the adoption of the ordinance — adopted in three of the five counties — led to simple policy differences in local financial-privacy statutes. Exploiting this variation, using Census tract-level and individual loan-level data on mortgage and refinancing applications, Kim and Wagman demonstrate that the opt-in ordinance had a statistically significant negative effect on loan denial rates (that is, approval rates increased), consistent with their theoretical model's predictions. They further provide some suggestive evidence that foreclosure start rates during the financial crisis of 2007–2008 were higher in the counties that adopted the privacy ordinance, possibly indicative of looser underwriting standards following the ordinance, also in line with their model's predictions.

3.6 Markets for Privacy and Personal Data

Databases of consumer data or consumer reports have existed throughout the twentieth century (Smith, 2000). The progress of information technology and the advent of the Internet have, however, vastly increased the scope and reach of those databases, ultimately giving rise to a market ecosystem of organizations that gather, merge, clean, analyze, buy, and sell consumer data. This ecosystem is complex and decentralized (Olejnik et al., 2014), although also dominated by a decreasing number of players (Krishnamurthy and Wills, 2009). There is no single, unified market for personal data. Rather, there are multiple markets in which data is traded, and multiple markets in which privacy is sought or purchased (cf. Lane et al., 2014). These include markets where data aggregators buy and sell data to other organizations (data subjects generally do not participate directly in these markets, and are in fact often unaware that reports on their names may exist); markets in which consumers exchange personal information for "free" products or services (for instance, search engines and online social networks); markets where consumers actively attempt to purchase protection for their data and/or against the negative consequences of privacy intrusions (for instance, identity theft, insurance services); and where consumers in fact attempt to explicitly trade their data in exchange for money (such as services provided by "personal data vault" firms, akin to the proposal for markets for privacy in Laudon, 1996).

One particular type of database that has attracted the attention of economists is the National Do-Not-Call Registry, a database established by the Do-Not-Call Implementation Act of 2003 to allow US residents, by registering themselves onto the list, to disallow telemarketers to call their phone numbers with promotional offers. Within 24 hours of its opening on June 27, 2003, over 10 million telephone numbers were registered; by February 2007, registrations exceeded 139 million. From the perspective of consumers, recent studies delineate the demographic characteristics of those who are likely to opt out using Do-Not-Call, and estimate consumers' benefit from doing so. Varian et al. (2005) calculate consumers' value for telemarketing privacy to range from $0.55 to $33.21 per household per year, while Png (2010), using state-level registries, estimates it to be between $13.19 to $98.33 per household per year.

Hui et al. (2014) use data from the federal registry and previously established state-level registries to demonstrate evidence of a certain externality associated with consumers opting out via Do-Not-Call. The decisions of some consumers to opt out — those consumers who prefer privacy to the benefits associated with targeted advertisements — end up reducing the pool of consumers available for sellers to solicit. In response, sellers redirect their marketing efforts to those consumers who are still available for solicitation. As these consumers experience an increase in solicitations, some of them respond by opting out as well. As more consumers opt out, sellers continue to adjust their solicitation. The conclusions of these works may extend to the FTC's proposed Do-Not-Track policy for online markets. In a sense, sellers face a form of a Prisoner's Dilemma situation under this mechanism. Individually, sellers wish to intensify their targeting of the pool of consumers who did not opt out; collectively, they would be better off holding back to keep this pool of consumers from shrinking further. Consumers, on the one hand, benefit from having the option of not being targeted with advertisements; on the other hand, consumers lose some of the benefits of targeted advertising, and those consumers who choose not to opt out are likely to be excessively targeted with advertisements.

3.7 Privacy and Information Security

While privacy and information security are distinct concepts, they overlap. Poor information security (by which we refer to the processes designed to protect data assets) can lead to what Solove (2006) refers to as "insecurity," or carelessness in protecting (personal) information from leaks and improper access. While the economics of information security has become a field in its own right, covering subjects as diverse as the optimal timing for patching operating systems or markets for software vulnerabilities,[21] a number of topics are of interest to both privacy and security researchers, such as spam, identity theft, and data breaches.

[21] For a review of the literature on the economics of information security, see Anderson and Moore (2006).

Two of the most commonly quoted consumer costs arising from the misuse of personal information are spam and identity theft. Spam refers to the indiscriminate use of electronic messaging systems for unsolicited advertisement to untargeted consumers. A recent study by Ferris Research estimates that in 2009, the cost of spam, accounting for decreased user productivity, was about $130 billion, with $42 billion in the US alone. While every Internet user receives spam, the cost per user is low, primarily due to users' reliance on filtering technologies. On the other hand, identity theft may affect fewer individuals, but at larger individual costs. The 1998 US Identity Theft and Assumption Deterrence Act (ITADA) defines identity theft as the knowing transfer, possession, or usage of any name or number that identifies another person, with the intent of committing, aiding or abetting a crime. Advances in information technology have allowed identity thieves to combine information taken from a variety of sources to open accounts in the names of others' identities (Cheney, 2005; Coggeshall, 2007). Anderson et al. (2008) report 30 mentions of "identity theft" in US newspapers in 1995; 2,000 in 2000; and 12,000 in 2005. The Bureau of Justice estimates that up to 16.6 million US residents ages 16 and older, or about 7% of the population in that age group, were victims of at least one incident of identity theft in 2012. It is estimated that identity theft resulted in corporate and consumer losses of $56 billion in 2005 and $61 billion in 2006, with 30% of known identity thefts caused by corporate data breaches. It is further estimated that 75% of recorded breaches between 2002 and 2007 were caused by hackers or external sources, with over 77% involving the theft of social security numbers — a piece of personal information whose loss may lead to identity theft (Synovate 2007; Schreft 2007; Anderson et al., 2008; Romanosky et al., 2011).

Miller and Tucker (2011b) study the impact of data encryption laws on data breach incidences. Their study highlights the risks associated with internal security threats (e.g., those by employees authorized to access a corporate database): policies that focus on outside threats may paradoxically redirect efforts away from protecting against internal threats. However, as noted by Mann (2014), "information lost may not be information abused": the unknown probability distributions of data breaches and of actual abuse of breached information make it harder to devise (or agree upon) sound framework for data security policy. For instance, Roberds and Schreft (2009) argue that the loss of privacy due to identity theft is outweighed by gains from the relative ease of gaining access to available credit. Kahn and Roberds (2008) model the incidence of identity theft as a trade-off between the desire to avoid costly or invasive monitoring of individuals on the one hand, and the need to control transaction fraud on the other. They suggest that this trade-off will prevail despite any technological advances. Kahn et al. (2005) examine the role of money in its provision of privacy and anonymous transactions, wherein a credit purchase may identify the purchaser. In a simple trading economy with moral hazard, the authors compare the efficiency of money and credit, and find that money may indeed be useful as a means of preserving anonymity toward sellers. More recently, the emergence of Bitcoin has provided a vehicle for doing just that — facilitating increased anonymity when transacting online (see, e.g., Kroll et al., 2013).

In response to increasing concerns regarding identity theft, many US states have responded by adopting data-breach disclosure laws that require firms to notify consumers if their personal information has been lost or stolen. Romanosky et al. (2011) uses FTC data to estimate the impact of data-breach disclosure laws on identity theft over the years 2002 to 2007. They find that the adoption of data-breach disclosure laws has a marginal effect on the incidences of identity thefts and reduces their average rate by under 2%. At the same time, state disclosure laws may have other benefits, such as reducing an average victim's loss and improving firms' security and operational practices (Schwartz and Janger, 2007).

In some sense, whether or not state laws require firms to disclose information about data breaches could be interpreted as firms' own level of privacy, which may pose its own set of trade-offs. Ideally, firms should be induced by strict disclosure laws to secure their customers' data. However, several studies that examine the financial impact of such disclosures on firms have come up with mixed and primarily mild results. Campbell et al. (2003), for instance, find "limited evidence of an overall negative stock market reaction to public announcements of information security breaches," although, they do find a significant and negative effect on stock price, specifically for breaches caused by "unauthorized access of confidential information." Considering a time window of one day before and one day after the announcement of a breach, they calculate a cumulative effect of -5.4%. Cavusoglu et al. (2004) find that the disclosure of a security breach results in the loss of 2.1% of a firm's market valuation over two days (the day of the announcement and the day after). Telang and Wattal (2007) find that software vendors' stock prices suffered when vulnerability information about their products is announced. Acquisti et al. (2006) use an event study to investigate the impact on stock market prices of firms that incur privacy breaches and find a negative and significant, but temporary, reduction of 0.6% in the stock market price on the day of the breach. Ko and Dorantes (2006) find that while a firm's overall performance is lower in the four quarters following a breach, the breached firm's sales increase significantly relative to firms that incurred no breach. These findings suggest that a strict disclosure policy alone may not be, by itself, the solution to aligning the interests of firms, in terms of data security, with those of their customers.

3.8 Consumer Attitudes and Behaviors

Surveys repeatedly find that privacy is one of the most significant concerns of Internet users. Turow et al. (2009) find that 66% of Americans do not want marketers to tailor advertisements to their interests, and 86% of young adults do not want tailored advertising if it is the result of following their behavior across websites. A 2013 Pew Research Center survey finds that 68% of adults believe that current laws are insufficient in protecting individuals' online privacy (Rainie et al., 2013).[22] At the same time, most consumers maintain their use of information technologies that track and share their personal information with unknown third parties. In fact, the adoption of privacy-

[22] For analyses of privacy complaints submitted by consumers to the FTC, see `http://www.knowprivacy.org/complaints.html` and `http://www.knowprivacy.org/report/KnowPrivacy_Final_Report.pdf`.

enhancing technologies (for instance, Tor[23]) lags vastly behind the adoption of sharing technologies (for instance, online social networks).

This dichotomy between attitudes and behaviors has caught the attention of scholars (e.g., Berendt et al., 2005), leading to a debate over the so-called privacy paradox (Norberg et al., 2007), and the value of privacy. Is the dichotomy real or imaginary? Do people actually care about privacy? If they do, how much exactly do they value the protection of their personal data?

A first possible resolution to the paradox is that it does not actually exist — attitudes are often expressed generically (for instance, the seminal categorization by Harris and Westin (1992) of individuals into privacy pragmatists, fundamentalists, and unconcerned, relied on broad and generic survey questions), whereas behaviors (or behavioral intentions) are specific and contextual. Thus, it is not surprising that the former may not correlate with or predict the latter (Fishbein and Ajzen, 1975).

A second resolution is that people routinely and expertly engage in mental trade-offs of privacy concerns and privacy benefits (Milberg et al., 1995), or a so-called privacy calculus (Laufer and Wolfe, 1977; Culnan and Armstrong, 1999; Dinev and Hart, 2006), naturally leading to situations in which they will choose to protect their data — and other situations in which protection is seen as too costly or ineffective, and sharing is preferred.

Privacy is, after all, a process of negotiation between public and private, a modulation of what a person wants to protect and what she wants to share at any given moment and in any given context. Therefore, neither does the sharing of certain information with others imply, per se, a loss of privacy, nor the complete hiding of data is necessary for the protection of privacy. In fact, the observation that people seem not to protect their privacy online is far from arriving at the conclusion that they never do so. Tsai et al. (2011) find that consumers are sometimes willing to pay a price premium to purchase goods from more privacy-protective merchants; Goldfarb and Tucker (2012b) use surveys to measure respondents' implied concern for privacy by their willingness to disclose information about income, and find evidence of privacy concerns increasing over an 8-year period; Stutzman et al. (2013) find evidence of increasing privacy-seeking behavior among a sample of over 4,000 early Facebook members; Kang et al. (2013) document Internet users' attempts to maintain anonymity online; and Boyd and Marwick (2011) discuss various alternative strategies teenagers adopt to protect their privacy while engaging in online sharing.

However, evidence of dichotomies between specific attitudes or preferences and actual behaviors have been also uncovered. Consider, for instance, Acquisti and Gross (2006), who provide a social-networking site context; or consider the aforementioned survey by Turow et al. (2009), which finds that 66% of Americans do not wish for marketers to tailor advertisements to their interests — while the vast majority of them arguably use search engines and social-networking sites, which operate based on enabling advertisers to target advertisements.

Thus, it is likely that the highlighted dichotomy between apparent desires for privacy and

[23] See `www.torproject.org`.

apparent willingness to have one's personal information acquired by strangers is the result of many different factors. Among them, a role is played by various decision-making hurdles consumers face when dealing with privacy challenges, especially online. In particular, an awareness that a problem even exists (for instance, Internet users are substantially unaware of the extents of behavioral targeting; see, e.g., McDonald and Cranor, 2010); knowledge of possible solutions (such as experience with privacy-enhancing technologies); as well as behavioral heuristics and decision-making biases — such as an immediate-gratification bias or a status-quo bias — which may affect the behavior of even well-informed and privacy-sensitive subjects (Acquisti, 2004).

Because of these challenges, it is difficult to pinpoint an exact valuation that consumers may assign to their privacy or to their personal data. There is no shortage of studies that attempt to quantify the value of data for organizations (for instance, Olejnik et al. (2014) find that elements of users' browsing histories are being traded among Internet advertising companies for amounts lower than $0.0005) and for end-users (for instance, Hann et al. (2007) quantify the value that US subjects assign to protection against errors, improper access, and secondary uses of personal information online to be between $30.49 and $44.62; Savage and Waldman (2013) find that a consumer is willing to make a one-time payment for a smartphone application of $2.28 to conceal their browser history, $4.05 to conceal their contacts list, $1.19 to conceal their location, $1.75 to conceal their phone's identification number, $3.58 to conceal the contents of their text messages, and $2.12 to eliminate advertising). However, privacy outcomes are uncertain (Knight, 1921) and privacy concerns and expectations are context-dependent (Nissenbaum, 2004), whereby common behavioral heuristics are likely to play a significant role in affecting privacy valuations (for instance, applying the endowment effect to the study of privacy, Acquisti et al. (2013) identify a discrepancy of up to five times the value individuals assign to the protection of personal information merely depending on the framing of the trade-offs, as opposed to actual changes in the trade-offs), as well as privacy-related behaviors, such as the propensity to disclose one's personal information to others (John et al., 2011).[24]

4 The Evolving Privacy Debate

Both the theoretical and the empirical economic studies that we have examined in the previous sections of this manuscript suggest that the path towards an agreeable solution for striking a balance between privacy protection, the benefits from information sharing, and (national) security, is uncertain. However, those studies also make the following clear: (i) the relevant stakeholders, including businesses, consumers, and governments, each have different, multi-layered, and often conflicting objectives; (ii) information technologies, privacy concerns, and the economics of privacy are constantly evolving, and no single study or piece of regulation can fully account for future (and some present) concerns; and (iii) rather than a uniform piece of regulation to address the decline in privacy, a nuanced approach that is dynamic and individualized to specific markets, contexts,

[24] For a review of the literature on privacy behavior and decision making, see Acquisti et al. (2015).

and issues is necessary. In this section, we point out a number of privacy issues that have started, or continue, to attract a lively debate involving economics, technology, and policy, and we propose a number of directions for future research.

4.1 Regulation versus Self-Regulation

Empirical studies of privacy trade-offs have contributed to the debate on how to best "protect" privacy without halting the beneficial effects — for data subjects and third parties alike — of information sharing. On one side of the debate, Gellman (2002) estimates that in 2001, $18 billion were lost by companies in Internet retail sales due to buyers' privacy concerns, and appraises at even larger amounts the costs that consumers bear when their privacy is not protected (the costs include losses associated with identity theft, higher prices paid by high-value consumers, spam, and investments aimed at protecting data). This side of the debate advocates regulatory solutions to privacy problems (Solove, 2002). One of the highlighted advantages of such solutions would be the avoidance of the complexity, for data subjects, to interact with different entities, each with a different privacy policy (cf. Milberg et al., 2000).

On the opposite side of the debate, Rubin and Lenard (2001) suggest that the costs of privacy protection are much higher for both firms and consumers alike. For instance, targeted advertising gives consumers useful information, advertising revenues support new Internet services, and reducing the use of online information would ultimately be costly to consumers. This side of the debate advocates self-regulatory solutions. Self-regulatory solutions may work when concerns over adverse consumer response limit advertisers' usage of invasive targeting of ads (Lohr, 2010); when website operators choose to comply with their published policies rather than engage in spam (Jamal et al., 2003); and when firms refrain from engaging in certain forms of price discrimination so as not to antagonize consumers (Anderson and Simester, 2010).

The US and the EU have taken different positions in this debate. The EU has focused on regulatory solutions, establishing principles that govern use of data across multiple sectors, including the need for consumer consent to data collection and processing. By contrast, the US has taken a limited and more sectorial and ad-hoc regulatory approach, often opting to provide guidelines rather than enforcing principles. For instance, recommendations to the US Congress by the FTC (Federal Trade Commission, 2012), motivated by the delays with which companies have adopted appropriate privacy rules, included the introduction of the aforementioned Do-Not-Track mechanism, similar to the Do-Not-Call list that became law in 2003. Such a mechanism would be built into websites and web browsers, and would presumably allow consumers to prohibit data collection processes about their online behavior with one click; however, limitations with respect to verification and enforcement would undoubtedly exist. Currently available services that allow consumers to opt out of advertising networks (such as the Self-Regulatory Program for Online Behavioral Advertising, and Google's opt-out settings) prevent users from receiving certain types of targeted ads but they do not stop advertisers or sites from collecting data.

Self-regulatory solutions often rely on transparency and control — and therefore are predicated around individuals' ability to be informed about, and properly manage, privacy settings and privacy concerns. However, numerous empirical studies have highlighted the limitations of transparency mechanisms. These include the failure of privacy policies to properly inform consumers about how their data will be used (Jensen and Potts, 2004); the large opportunity costs associated with frameworks that rely on consumers reading privacy policies (McDonald and Cranor, 2008); and how the same policy can be used to induce variable amounts of individual disclosure via simple changes in the way the policy itself is presented to users (Adjerid et al., 2013) and in control mechanisms. For instance, while research in information systems has suggested that providing users with control over their information can reduce privacy concerns (Culnan and Armstrong, 1999; Malhotra et al., 2004; Tucker, 2014), the protection afforded by control may be illusory: Brandimarte et al. (2013) highlight how the mere provision of more perceived control over personal information can paradoxically lead users to take more risks with their personal information, increasing their willingness to share sensitive data with another party. As a result, doubts are being expressed about the ability of self-regulatory "notice and consent" and transparency and control mechanisms to adequately protect consumers' privacy (Acquisti et al., 2013; Solove, 2013).

An alternative approach to privacy protection relies on the propertization (Laudon, 1996; Varian, 1997; Schwartz, 2004) or licensing (Samuelson, 2000) of personal information. As noted in Section 2.2, the idea is of markets where individuals can trade (rights over) their personal information. With the advent of Web 2.0, a number of startups began offering similar services. However, it is not clear that such markets for personal data could ever be successful. First, in dealing with services that offer trade and protection for their data, consumers face similar hurdles in dealing with transparency and consent to those they face with traditional privacy policies, including the hurdle of estimating the fair value of their personal information. Second, in absence of regulatory frameworks that would enforce protection of traded data, the possibility of secondary usage of personal information, once the subject has traded it to another party, would run counter to the very idea of protecting consumer data. Third, much of consumer data that is of value to advertisers is non-static information that is dynamically generated as part of the interaction of the individual with other online services — such as search engines or online social networks, and these services would be unlikely to relinquish control over the personal information their services help generate. Thus, an alternative that has been proposed in the literature is the reliance on soft paternalistic solutions (designed by governments, organizations, or data subjects themselves as self-control mechanisms) to "nudge" individuals towards personal information practices they have claimed to prefer (Wang et al., 2013).

As noted in Section 2.2, market-based solutions and regulatory approaches to privacy protection are not polar opposites; they are better perceived as points on a spectrum of solutions — from regimes that rely entirely on firms' self-regulation and consumers' responsibility (even in absence of clearly defined and assigned property rights over personal data), to regimes with strict regulatory protection of data. Similarly, as pointed out in Section 3.4, an understanding is emerging that the economic impact of privacy regulation is heterogeneous and context-dependent (with both positive and negative effects on economic growth and efficiency arising from regulation), depending on the specific *attributes* of privacy laws. Thus, a potentially worthwhile direction of future research will aim at focusing on the specific features and and degrees of regulation (and their differential effects on economic outcomes), rather than on simpler binary models contrasting regulation with its absence.

4.2 From Big Data to Privacy-Enhancing Technologies

As the amount of personal information produced and gathered about individuals continues to increase, so does the ability to utilize data mining to infer more sensitive information about individuals. For instance, Sweeney (1997) has highlighted the possibility of re-identifying supposedly anonymous health data, and Acquisti and Gross (2009) have shown how apparently innocuous self-revelations made on the Internet — such as making available one's date and state of birth in a social network profile — may have serious consequences in terms of privacy intrusion. With progress in data mining and business analytics applied to larger sets of personal data (the so called "big data" phenomenon), and with new technologies – from facial recognition to activity and health trackers to the so-called "Internet of Things" – the portions of our personal and professional lives that are not monitored and quantified are further reduced. On one hand, granular personal data may be used to provide increasingly targeted services and to ensure that advertising is shown only to those consumers who stand to gain most from it (Tucker, 2012). On the other hand, opportunities for abuse may abound. For instance, algorithmic discrimination may take subtle forms (Sweeney, 2013, documents cases in which advertising technologies employed by a search engine can expose racial bias), and personal data may be used to influence individual decision making in subtle, targeted, and hidden manners (Calo, 2014), raising questions over the limits of a person's autonomy and self-determination in a world where so much personal information can be gathered and used to influence the individual (for instance, Kramer et al. (2014) show that it is possible to influence the emotional states of users of a social networking site in the form of an emotional "contagion" by suppressing information containing positive or, alternatively, negative emotions).

The obvious trade-offs arising from the intersection of big data and privacy suggest several fruitful directions for research. For instance: To what extent the combination of sophisticated analytics and massive amounts of consumer data will lead to an increase in aggregate welfare, and to what extent will it lead to mere changes in the allocation of wealth? A related open question concerns the role of privacy-enhancing technologies (Goldberg, 2003) in affecting how personal information will be used and with what economic consequences. Privacy-enhancing technologies, or PETs, can allow the protection of sensitive consumer data without disrupting commercially valuable flows of consumer information. They do, however, reduce the granularity of individual information available to others (consider, for instance, differential privacy, as in Dwork, 2006), and therefore its economic value. Thus, how costly (in terms of opportunity costs of missed data) will the protection of information be by using techniques such as differential privacy or tools like PETs, relative to the benefits they could afford in terms of privacy? And, importantly, who will bear those costs — the data subjects or data holders? Finally, the contrast between the potential value of (big) data, and its privacy costs, raises questions about optimal retention policies: for instance, do larger quantities of consumer historical data provide competitive advantages to Internet search firms (Chiou and Tucker, 2014)? And will the so called "right to be forgotten," promoted by European regulators (Rosen, 2012), support individuals' privacy rights without hampering the societal benefits of data sharing?

4.3 Open Data, Government Records, and Surveillance

The topic of this survey is the economics of privacy, and we have, therefore, naturally focused on the commercial acquisition and exploitation of personal data. It would be remiss of us, however, not to mention a no–less important facet of the privacy debate, one with potentially even greater impact on individuals and societies — namely the role and value of open access to data, governmental records, but also the covert governmental collection of personal information.

Access to personal data (from governmental administrative records, to researchers' results arising from experiments and surveys, to firms' collections of consumers' data) is of great importance to empirical economists and social scientists. And once again trade-offs arise, between the utility of sharing publicly (or with other researchers) personal records and files, and the privacy risks associated with granting access to third parties. Essentially, data utility and risks of disclosure are correlated (Duncan and Stokes, 2004), and even statistical techniques meant to protect data (such as the technique of differential privacy, which attempts to minimize the risks of re-identification of records in a statistical database while maximizing the accuracy of queries from such database) still face risk/utility trade-offs (Fienberg et al., 2010), not just for firms but also for researchers (Komarova et al., 2014). Furthermore, even protected (or anonymized, or de-identified) data may still be exposed (Ohm, 2010; Heffetz and Ligett, 2013): for instance, a portion of anonymized movie ratings data made available by Netflix as part of a competition to improve its ratings algorithms could be re-identified using Internet Movie Database (IMDB) data (Narayanan and Shmatikov, 2008). These trade-offs apply both to government databases (the Census uses a variety of mechanisms and procedures to balance researchers' needs to access Census data with considerations for Census respondents' privacy) and to the private sector.[25] How to best balance researchers' and society's needs to access granular data with the need to protect individuals' records is a question that simultaneously involves economists and scholars in other disciplines, such as statisticians and computer scientists.

[25] Similarly, the problems of data breaches and identity theft discussed in Section 3.7 do not arise only in the context of firms' databases, but also governmental ones: in the last few years, a number of large-scale data breaches involved governmental data, such as the loss of 26.5M records of veterans, their spouses, and active-duty military personnel in 2006.

As for the topic of government surveillance, the US PATRIOT Act enacted in 2001 and extended in 2011 facilitated the US government's collection of more information, from a greater number of sources, than had previously been authorized in criminal or foreign intelligence investigations, superseding, among others, the Foreign Intelligence Surveillance Act, the Electronic Communications Privacy Act, and the National Security Letter statutes. The PATRIOT Act enabled greater access to records showing an individual's spending and communications, including e-mail and telephone conversations (Congress CRS Report, 2011). Beginning in June 2013, a series of disclosures by former CIA employee and contractor Edward Snowden of thousands of classified documents involving data collection by the National Security Agency triggered a massive wave of public concern about privacy and governmental overreach. The report of the US Administration's Review Group on Intelligence and Communication Technologies (2013) states that "excessive surveillance and unjustified secrecy can threaten civil liberties, public trust, and the core processes of democratic self-government," whereas in "an era increasingly dominated by technological advances in communication technologies, the United States must continue to collect signals intelligence globally in order to assure the safety of our citizens." Together, the report states, the US government "must protect, at once, two different forms of security: national security and personal privacy." The report concludes that the "government should base its decisions on a careful analysis of consequences, including both benefits and costs (to the extent feasible)." How to devise strategies that balance privacy protection, benefits from information sharing, and national security will likely remain a thorny and yet vital subject of research for years to come.

4.4 Conclusions

Personal information has both private and commercial value, and often (though not always) exploiting its commercial value entails a reduction in private utility and sometimes even in social welfare overall. Consumers have good reasons to be concerned about unauthorized commercial application of their private information. Use of individual data may subject an individual to a variety of personally costly practices, including price discrimination in retail markets, quantity discrimination in insurance and credit markets, spam, and risk of identity theft, in addition to the disutility inherent in just not knowing who knows what or how they will use it in the future. Personal data — like all information after all — is easily stored, replicated, and transferred, and regulating its acquisition and dissemination is a challenging undertaking for individuals and governments alike.

Given the fundamentally sensitive nature of personal data, it is not surprising that advancements in information and communication technologies and increased globalization of trade, investment, information flows, and security threats have brought concerns over the erosion of personal privacy to the forefront of public debate, and not without justification. Numerous Internet firms have collected large amounts of data from their users and either sell this data or use it to enable advertisers to target and personalize ads. While consumers may benefit from targeted product recommendations (Anand and Shachar, 2009), they may also incur substantial monetary costs and disutility from a violation of their privacy (Stone, 2010). Such concerns have led to new regulations across world governments, some protecting privacy (e.g., the EU's Data Protection Directive, the US Children's Online Privacy Protection Act), some legalizing its erosion (for instance, by allowing trade in personal information under certain circumstances; see, e.g., the US Gramm-Leach-Bliley Act of 1999), and some suggesting the implementation of additional opt-in and opt-out controls for users (e.g., the US Federal Trade Commission's online privacy guidelines, 2012).

With regulations struggling to keep pace, industry competition has been behind both new privacy-enhancing and privacy-invasive technologies. New search engines, social networks, ecommerce websites, web browsers, and individualized controls for privacy-conscious consumers have emerged. Concurrently, social media services have facilitated a culture of disclosure: a disclosure of one's activities, location, emotions, work history, and political opinions. While, overall, these technologies seemingly leave privacy choices in the hands of consumers, many (if not most) consumers, in practice, lack the awareness and technical sophistication required to protect and regulate the multiple dimensions of their personal information. Privacy-invasive technological services have become integral to every-day communications, job searches, and general consumption. At the same time, privacy-protecting services require additional levels of user effort and knowhow, limiting their efficacy, especially within some of the most vulnerable segments of the population.

We have, in this article, attempted to survey and rationalize the extant research on the economics of privacy. Because privacy is a multi-faceted concept, our survey has delved into numerous literatures across a variety of disciplines and fields, from marketing to economics to computer science. While this study is certainly not exhaustive, we believe it highlights some of the most relevant historical and current research on the topic. It is, however, abundantly evident that protection of personal privacy is rapidly emerging as one of the most significant current public policy issues, and research on the economics of privacy will, therefore, undoubtedly continue to expand and evolve in coming years.

References

Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the 5th ACM conference on Electronic commerce*, pp. 21–29.

Acquisti, A., L. Brandimarte, and G. Loewenstein (2015). Privacy and human behavior in the age of information. *Science 347*(6221), 509–514.

Acquisti, A. and C. M. Fong (2013). An experiment in hiring discrimination via online social networks. *Available at SSRN 2031979*.

Acquisti, A., A. Friedman, and R. Telang (2006). Is there a cost to privacy breaches? An event study. In *Twenty Seventh International Conference on Information Systems, Milwaukee 2006 and Workshop on the Economics of Information Security 2006*.

Acquisti, A. and R. Gross (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *Privacy enhancing technologies*, pp. 36–58. Springer.

Acquisti, A. and R. Gross (2009). Predicting social security numbers from public data. *Proceedings of the National academy of sciences 106*(27), 10975–10980.

Acquisti, A., L. K. John, and G. Loewenstein (2013). What is privacy worth? *The Journal of Legal Studies 42*(2), 249–274.

Acquisti, A. and H. R. Varian (2005). Conditioning prices on purchase history. *Marketing Science 24*(3), 367–381.

Adjerid, I., A. Acquisti, L. Brandimarte, and G. Loewenstein (2013). Sleights of privacy: Framing, disclosures, and the limits of transparency. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, pp. 1–11.

Adjerid, I., A. Acquisti, R. Telang, R. Padman, and J. Adler-Milstein (2013). Impact of health disclosure laws on health information exchanges. In *NBER Workshop on the Economics of Digitization*.

Akerlof, G. (1970). The market for 'lemons': Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics 84*(3), 488–500.

Alstyne, M. V. (2007). Curing spam: Rights, signals & screens. *The Economists' Voice 4*(2), 1–4.

Anand, B. and R. Shachar (2009). Targeted advertising as a signal. *Quantitative Marketing and Economics 7*(3), 237–266.

Anderson, E. and D. Simester (2010). Price stickiness and customer antagonism. *Quarterly Journal of Economics 125*(2), 729–765.

Anderson, K., E. Durbin, and M. Salinger (2008). Identity theft. *Journal of Economic Perspectives 22*(2), 171–192.

Anderson, R. and T. Moore (2006). The economics of information security. *Science 314*(5799), 610–613.

Anderson, S. and A. de Palma (2012). Competition for attention in the information (overload) age. *RAND Journal of Economics 43*(1), 1–25.

Armstrong, M., J. Vickers, and J. Zhou (2009). Consumer protection and the incentive to become informed. *Journal of the European Economic Association 7*(2-3), 399–410.

Armstrong, M. and J. Zhou (2010). Conditioning prices on search behaviour. Technical Report 19985, MPRA Paper, University Library of Munich.

Arrow, K. (1962). The economic implications of learning by doing. *The Review of Economic Studies 29*(3), 155–173.

Asplund, M., R. Eriksson, and N. Strand (2008). Price discrimination in oligopoly: Evidence from regional newspapers. *The Journal of Industrial Economics 56*(2), 333–346.

Athey, S., E. Calvano, and J. Gans (2013). The impact of the internet on advertising markets for news media. Working Paper.

Athey, S. and J. Gans (2010). The impact of targeting advertising technology on advertising markets and media competition. *American Economic Review 100*(2), 608–613.

Baumer, D., J. Earp, and J. Poindexter (2004). Internet privacy law: a comparison between the United States and the European Union. *Computers & Security 23*(5), 400–412.

Baye, M. and J. Morgan (2001). Information gatekeepers on the internet and the competitiveness of homogeneous product markets. *American Economic Review 91*(3).

Beales, H. (2010). The value of behavioral targeting. Working Paper.

Bendrath, R. and M. Mueller (2011). The end of the net as we know it? Deep packet inspection and Internet governance. *New Media & Society 13*(7), 1142–1160.

Berendt, B., O. Gunther, and S. Spiekermann (2005). Privacy in e-commerce: Stated preferences vs. actual behavior. *Communications of the ACM 48*(4), 101–106.

Bergemann, D. and A. Bonatti (2011). Targeting in advertising markets: implications for offline vs. online media. *RAND Journal of Economics 42*(3), 417–443.

Bergemann, D. and A. Bonatti (2013). Selling cookies. Working Paper.

Berners-Lee, T., M. Fischetti, and M. L. Dertouzos (2000). *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor*. HarperInformation.

Bertrand, M. and S. Mullainathan (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review 94*(4), 991–1013.

Blake, T., C. Nosko, and S. Tadelis. Consumer heterogeneity and paid search effectiveness: A large scale field experiment. *Econometrica*. forthcoming.

Blattberg, R. C. and J. Deighton (1991). Interactive marketing: Exploiting the age of addressability. *Sloan Management Review 33*(1), 5–14.

Blumstein, A. and K. Nakamura (2009). Redemption in the presence of widespread criminal background checks. *Criminology 47*(2), 327–359.

Boyd, D. and A. Marwick (2011). Social steganography: Privacy in networked publics. *International Communication Association, Boston, MA*.

Brandimarte, L., A. Acquisti, and G. Loewenstein (2013). Misplaced confidences privacy and the control paradox. *Social Psychological and Personality Science 4*(3), 340–347.

Burke, J., C. Taylor, and L. Wagman (2012). Information acquisition in competitive markets: An application to the us mortgage market. *American Economic Journal: Microeconomics 4*(4), 65–106.

Bushway, S. D. (2004). Labor market effects of permitting employer access to criminal history records. *Journal of Contemporary Criminal Justice 20*(3), 276–291.

Calo, R. (2014). Digital market manipulation. *George Washington Law Review 82*(4), 995–1051.

Calzolari, G. and A. Pavan (2006). On the optimality of privacy in sequential contracting. *Journal of Economic Theory 130*(1), 168–204.

Campbell, J., A. Goldfarb, and C. Tucker (2015). Privacy regulation and market structure. *Journal of Economics & Management Strategy 24*(1), 47–73.

Campbell, K., L. A. Gordon, M. P. Loeb, and L. Zhou (2003). The economic cost of publicly announced information security breaches: Empirical evidence from the stock market. *Journal of Computer Security 11*(3), 431–448.

Cavusoglu, H., B. Mishra, and S. Raghunathan (2004). The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. *International Journal of Electronic Commerce 9*(1), 69–104.

Chellappa, R. K. and S. Shivendu (2010). Mechanism design for "free" but "no free disposal" services: The economics of personalization under privacy concerns. *Management Science 56*(10), 1766–1780.

Chen, Y. (1997). Paying customers to switch. *Journal of Economics and Management Strategy 6*(4), 877–897.

Chen, Y. and G. Iyer (2002). Consumer addressability and customized pricing. *Marketing Science 22*(2), 197–208.

Chen, Y., C. Narasimhan, and Z. J. Zhang (2001). Individual marketing with imperfect targetability. *Marketing Science 20*(1), 23–41.

Chen, Y. and Z. J. Zhang (2009). Dynamic targeted pricing with strategic consumers. *International Journal of Industrial Organization 27*(1), 43–50.

Cheney, J. S. (2005). Identity theft: Do definitions still matter? Technical Report 2005-10, Federal Reserve Bank of Philadelphia.

Chiou, L. and C. Tucker (2014). Search engines and data retention: Implications for privacy and antitrust. Technical report.

Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics 3*(1), 1–44.

Coggeshall, S. (2007, April 13). ID theft knows no boundaries. *E-Commerce Times*.

Conitzer, V., C. Taylor, and L. Wagman (2012). Hide and seek: Costly consumer privacy in a market with repeat purchases. *Marketing Science 31*(2), 277–292.

Cornière, A. D. (2011). Search advertising. Working Paper.

Cowan, S. (2007). The welfare effects of third-degree price discrimination with non-linear demand functions. *RAND Journal of Economics 38*(2), 419–428.

Culnan, M. J. and P. K. Armstrong (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science 10*(1), 104–115.

Daughety, A. and J. Reinganum (2010). Public goods, social pressure, and the choice between privacy and publicity. *American Economic Journal: Microeconomics 2*(2), 191–221.

De Cornière, A. and R. D. Nijs (2014). Online advertising and privacy. *Available at SSRN 2191124*.

Dinev, T. and P. Hart (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research 17*(1), 61–80.

Djankov, S., C. McLiesh, and A. Shleifer (2007). Private credit in 129 countries. *Journal of Financial Economics 84*(2), 299–329.

Dugas, A. F., Y.-H. Hsieh, S. R. Levin, J. M. Pines, D. P. Mareiniss, A. Mohareb, C. A. Gaydos, T. M. Perl, and R. E. Rothman (2012). Google flu trends: Correlation with emergency department influenza rates and crowding metrics. *Clinical infectious diseases 54*(4), 463–469.

Duncan, G. T. and S. L. Stokes (2004). Disclosure risk vs. data utility: The RU confidentiality map as applied to topcoding. *Chance 17*(3), 16–20.

Dwork, C. (2006). Differential privacy. In *Automata, languages and programming*, pp. 1–12. Springer.

Edelman, B. G. and M. Luca (2014). Digital discrimination: The case of airbnb.com. *Harvard Business School NOM Unit Working Paper* (14-054).

Evans, D. S. (2009). The online advertising industry: Economics, evolution, and privacy. *Journal of Economic Perspectives 23*(3), 37–60.

Farahat, A. and M. Bailey (2012). How effective is targeted advertising? In *Proceedings of the 21st International Conference on WWW, 2012*.

Farrell, J. (1987). Information and the Coase theorem. *Journal of Economic Perspectives 1*(2), 113–129.

Farrell, J. (2012). Can privacy be just another good? *Journal on Telecommunications and High Technology Law 10*, 251.

Federal Trade Commission (2012). Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policy Makers. Report, http://www.ftc.gov/os/2012/03/120326privacyreport.pdf.

Fienberg, S. E., A. Rinaldo, and X. Yang (2010). Differential privacy and the risk-utility tradeoff for multi-dimensional contingency tables. In *Privacy in Statistical Databases*, pp. 187–199. Springer.

Fishbein, M. and I. Ajzen (1975). *Belief, attitude, intention and behavior: An introduction to theory and research.* Addison-Wesley.

Fudenberg, D. and J. Tirole (1998). Upgrades, trade-ins, and buy-backs. *RAND Journal of Economics 29*(2), 238–258.

Fudenberg, D. and J. Tirole (2000). Customer poaching and brand switching. *RAND Journal of Economics 31*(4), 634–657.

Garrie, D. and R. Wong (2006). Demystifying clickstream data: A Eutopian and U.S. perspective. *Emory International Law Review 20*(2), 563–589.

Gehrig, T. and R. Stenbacka (2007). Information sharing and lending market competition with switching costs and poaching. *European Economic Review 51*(1), 77–99.

Gellman, R. (2002). Privacy, consumers, and costs - How the lack of privacy costs consumers and why business studies of privacy costs are biased and incomplete. `http://www.epic.org/reports/dmfprivacy.html`.

Goldberg, I. (2003). Privacy-enhancing technologies for the internet, II: Five years later. In *Second International Workshop on Privacy Enhancing Technologies.*

Goldfarb, A. and C. Tucker (2011a). Online display advertising: Targeting and obtrusiveness. *Marketing Science 30*(3), 389–404.

Goldfarb, A. and C. Tucker (2011b). Privacy regulation and online advertising. *Management Science 57*(1), 57–71.

Goldfarb, A. and C. Tucker (2012a). Privacy and innovation. *Innovation Policy and the Economy 12*(1), 65–90.

Goldfarb, A. and C. Tucker (2012b). Shifts in privacy concerns. *American Economic Review: Papers and Proceedings 102.*

Goldin, C. and C. Rouse (2000). Orchestrating impartiality: The impact of" blind" auditions on female musicians. *The American Economic Review 90*(4), 715–741.

Gottlieb, D. and K. Smetters (2011). Grade non-disclosure. Available at NBER: http://www.nber.org/papers/w17465.

Gradwohl, R. (2015). Privacy in implementation. Working Paper.

Gradwohl, R. and O. Reingold (2010). Partial exposure in large games. *Games and Economic Behavior 68*(2), 602–613.

Grossman, G. and C. Shapiro (1984). Informative advertising with differentiated products. *Review of Economic Studies 51*(1), 63–81.

Grossman, S. (1981). The informational role of warranties and private disclosure about product quality. *Journal of Law & Economics. 24*(3), 461–483.

Hagiu, A. and B. Jullien (2011). Why do intermediaries divert search. *RAND Journal of Economics 42*(2), 337–362.

Hann, I.-H., K.-L. Hui, S.-Y. T. Lee, and I. P. Png (2007). Overcoming online information privacy concerns: An information-processing theory approach. *Journal of Management Information Systems 24*(2), 13–42.

Hann, I.-H., K.-L. Hui, T. S. Lee, and I. P. Png (2008). Consumer privacy and marketing avoidance: A static model. *Management Science 54*(6), 1094–1103.

Harris, L. and A. Westin (1992). The Equifax Canada Report on Consumers and Privacy in the Information Age. Technical report.

Hayek, F. (1945). The use of knowledge in society. *The American Economic Review 35*(4), 519–530.

Heffetz, O. and K. Ligett (2013). Privacy and data-based research. Technical report, National Bureau of Economic Research.

Hillestad, R., J. Bigelow, A. Bower, F. Girosi, R. Meili, R. Scoville, and RogerTaylor (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs 24*(5), 1103–1117.

Hirshleifer, J. (1971). The private and social value of information and the reward to inventive activity. *The American Economic Review 61*(4), 561–574.

Hirshleifer, J. (1980). Privacy: Its origins, function and future. *Journal of Legal Studies 9*(4), 649–664.

Hoffmann, F., R. Inderst, and M. Ottavniani (2013). Hypertargeting, limited attention, and privacy: Implications for marketing and campaigning. Working Paper.

Hoofnagle, C. J., A. Soltani, N. Good, and D. J. Wambach (2012). Behavioral advertising: The offer you can't refuse. *Harvard Law & Policy Review 6*, 273.

Hui, K., I. Png, and K. Goh (2014). Privacy externalities and 'opt out': Theory and evidence from do not call. Available at: http://pubsonline.informs.org/doi/10.1287/mnsc.2014.2051.

Hui, K.-L. and I. Png (2006). The economics of privacy. In T. Hendershott (Ed.), *Handbook on Economics and Information Systems*. Amsterdam, North-Holland.

Jamal, K., M. Maier, and S. Sunder (2003). Privacy in e-commerce: Development of reporting standards, disclosure, and assurance services in an unregulated market. *Journal of Accounting Research 41*(2), 285–309.

Jappelli, T. and M. Pagano (2002). Information sharing, lending and defaults: Cross-country evidence. *Journal of Banking and Finance 26*(10), 2017–2045.

Jensen, C. and C. Potts (2004). Privacy policies as decision-making tools: an evaluation of online privacy notices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 471–478.

Jentzsch, N. (2003). *The Regulation of Financial Privacy: The United States vs. Europe*. European Credit Research Institute.

Jeong, Y. and M. Maruyama (2009). Commitment to a strategy of uniform pricing in a two-period duopoly with switching costs. *Journal of Economics 98*(1), 45–66.

Jernigan, C. and B. F. Mistree (2009). Gaydar: Facebook friendships expose sexual orientation. *First Monday 14*(10).

Jing, B. (2011). Pricing experience goods: The effects of customer recognition and commitment. *Journal of Economics & Management Strategy 20*(2), 451–473.

John, L. K., A. Acquisti, and G. Loewenstein (2011). Strangers on the plane: Context-dependent willingness to divulge sensitive information. *Journal of Consumer Research 37*(5), 858–873.

Johnson, E. and D. Goldstein (2004). Defaults and donation decisions. *Transplantation 78*(12), 1713–1716.

Johnson, J. (2013). Targeted advertising and advertising avoidance. *RAND Journal of Economics 44*(1), 128–144.

Kahn, C., J. McAndrews, and W. Roberds (2000). A theory of transactions privacy. Technical Report 2000-22, Federal Reserve Bank of Atlanta.

Kahn, C., J. McAndrews, and W. Roberds (2005). Money is privacy. *International Economic Review 46*(2), 377–399.

Kahn, C. and W. Roberds (2008). Credit and identity theft. *Journal of Monetary Economics 55*(2), 251–264.

Kang, R., S. Brown, and S. Kiesler (2013). Why do people seek anonymity on the internet? Informing policy and design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2657–2666.

Kearns, M., M. Pai, A. Roth, and J. Ullman (2014). Mechanism design in large games: Incentives and privacy. *American Economic Review: Papers and Proceedings 104*(5), 431–435.

Kim, B. and J. Choi (2010). Customer information sharing: Strategic incentives and new implications. *Journal of Economics & Management Strategy 19*(2), 403–433.

Kim, J.-H. and L. Wagman (2015). Screening incentives and privacy protection in financial markets: A theoretical and empirical analysis. *RAND Journal of Economics 46*(1), 1–22.

Knight, F. H. (1921). Risk, uncertainty and profit. *New York: Hart, Schaffner and Marx*.

Ko, M. and C. Dorantes (2006). The impact of information security breaches on financial performance of the breached firms: An empirical investigation. *Journal of Information Technology Management 17*(2).

Komarova, T., D. Nekipelov, and E. Yakovlev (2014). Estimation of treatment effects from combined data: Identification versus data security. In *Economics of Digitization*. University of Chicago Press.

Kramer, A. D., J. E. Guillory, and J. T. Hancock (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences 111*(24), 8788–8790.

Krishnamurthy, B. and C. Wills (2009). Privacy diffusion on the web: A longitudinal perspective. In *Proceedings of the 18th international conference on World wide web*, pp. 541–550.

Kroll, J., I. Davey, and E. Felten (2013). The economics of bitcoin mining, or bitcoin in the presence of adversaries. In *Proceedings of WEIS, 2013*.

Lambrecht, A. and C. Tucker (2013). When does retargeting work? information specificity in online advertising. *Journal of Marketing Research 50*(5), 561–576.

Lane, J., V. Stodden, S. Bender, and H. Nissenbaum (2014). *Privacy, Big Data, and the Public Good: Frameworks for Engagement.* Cambridge University Press.

Laudon, K. (1997, January). Extensions to the theory of markets and privacy: Mechanics of pricing information. Stern School of Business - New York University - Working Papers.

Laudon, K. C. (1996). Markets and privacy. *Communications of the ACM 39*(9), 92–104.

Laufer, R. S. and M. Wolfe (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of Social Issues 33*(3), 22–42.

Levin, J. and P. Milgrom (2010). Online advertising: Heterogeneity and conflation in market design. *American Economic Review, Papers and Proceedings 100*(2), 603–607.

Lewis, R. and D. Reiley (2013). Online ads and offline sales: Measuring the effects of online advertising via a controlled experiment on Yahoo!. Working Paper.

Lohr, S. (2010). Privacy concerns limit online ads, study says. New York Times.

Malhotra, N. K., sung S.Kim, and J. Agarwal (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research 15*(4), 336–355.

Mann, C. (2014). Information lost: Will the "paradise" that information promises, to both consumer and firm, be "lost" on account of data breaches? the epic is playing out. In *Economic Analysis of the Digital Economy.* National Bureau of Economic Research, Inc.

McDonald, A. M. and L. F. Cranor (2008). The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society 4*, 540–565.

McDonald, A. M. and L. F. Cranor (2010). Beliefs and behaviors: Internet users' understanding of behavioral advertising. In *Proceedings of the 2010 Research Conference on Communication, Information and Internet Policy.*

Mikians, J., L. Gyarmati, V. Erramilli, and N. Laoutaris (2012). Detecting price and search discrimination on the Internet. In *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, pp. 79–84.

Mikians, J., L. Gyarmati, V. Erramilli, and N. Laoutaris (2013). Crowd-assisted search for price discrimination in e-commerce: First results. In *Proceedings of the Ninth ACM Conference on Emerging Networking Experiments and Technologies*, pp. 1–6.

Milberg, S. J., S. J. Burke, H. J. Smith, and E. A. Kallman (1995). Values, personal information privacy, and regulatory approaches. *Communications of the ACM 38*(12), 65–74.

Milberg, S. J., H. J. Smith, and S. J. Burke (2000). Information privacy: Corporate management and national regulation. *Organization Science 11*(1), 35–57.

Milgrom, P. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics 12*(2), 380–391.

Milgrom, P. and J. Roberts (1986). Relying on information of interested parties. *RAND Journal of Economics 17*(1), 18–32.

Miller, A. and C. Tucker (2007). Privacy, Networks Effects and Electronic Medical Record Technology Adoption. In *Proceedings of WEIS, 2007.*

Miller, A. and C. Tucker (2009). Privacy Protection and Technology Diffusion: The Case of Electronic Medical Records. *Management Science 55*(7), 1077–1093.

Miller, A. and C. Tucker (2011a). Can healthcare information technology save babies? *Journal of Political Economy 119*(2), 289–324.

Miller, A. and C. Tucker (2014a). Electronic discovery and the adoption of information technology. *Journal of Law, Economics, and Organization 30*, 217–243.

Miller, A. R. and C. Tucker (2014b). Privacy protection, personalized medicine and genetic testing. Available at SSRN: http://ssrn.com/abstract=2411230.

Miller, A. R. and C. E. Tucker (2011b). Encryption and the loss of patient data. *Journal of Policy Analysis and Management 30*(3), 534–556.

Murray, B. and J. Cowart (2001). A study of the presence and growth rate of web bugs on the internet. Technical report, Cyveillance, Inc.

Nagin, D. and J. Waldfogel (1995). The effects of criminality and conviction on the labor market status of young british offenders. *International Review of Law and Economics 15*(1), 109–126.

Narayanan, A. and V. Shmatikov (2008). Robust de-anonymization of large sparse datasets. In *Security and Privacy, 2008. SP 2008. IEEE Symposium on*, pp. 111–125. IEEE.

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review 79*, 101–139.

Noam, E. M. (1997). Privacy and self-regulation: Markets for electronic privacy. In *Privacy and Self-regulation in the Information Age*. US Department of Commerce.

Norberg, P. A., D. R. Horne, and D. A. Horne (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs 41*(1), 100–126.

Odlyzko, A. (2003). Privacy, economics, and price discrimination on the Internet. In L. J. Camp and S. Lewis (Eds.), *Economics of Information Security*. Springer-Kluwer.

Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review 57*, 1701.

Olejnik, L., T. Minh-Dung, C. Castelluccia, et al. (2014). Selling off privacy at auction. In *ISOC Network and Distributed System Security Symposium.*

Oster, E., I. Shoulson, K. Quaid, and E. R. Dorsey (2010). Genetic adverse selection: Evidence from long-term care insurance and Huntington disease. *Journal of Public Economics 94*(11–12), 1041–1050.

Pagano, M. and T. Jappelli (1993). Information sharing in credit markets. *The Journal of Finance 48*(5), 1693–1718.

Pancras, J. and K. Sudhir (2007). Optimal marketing strategies for a customer data intermediary. *Journal of Marketing Research 44*(4), 560–578.

Pavan, A. and G. Calzolari (2009). Sequential contracting with multiple principals. *Journal of Economic Theory 144*(2), 503–531.

Png, I. (2010). On the Value of Privacy from Telemarketing: Evidence from the 'Do Not Call' Registry. Working paper.

Posner, R. A. (1978). The right of privacy. *Georgia Law Review 12*(3), 393–422.

Posner, R. A. (1981). The economics of privacy. *The American Economic Review 71*(2), 405–409.

Posner, R. A. (1993). Blackmail, privacy, and freedom of contact. *University of Pennsylvania Law Review 141*(5), 1817–1847.

Qian, J. and P. Strahan (2007). How laws and institutions shape financial contracts: The case of bank loans. *The Journal of Finance 62*(6), 2803–2834.

Rainie, L., S. Kiesler, R. Kang, M. Madden, M. Duggan, S. Brown, and L. Dabbish (2013). Anonymity, privacy, and security online. *Pew Research Center*.

Roberds, W. and S. L. Schreft (2009). Data security, privacy, and identity theft:the economics behind the policy debates. *Economic Perspectives* (1Q), 22–30.

Romanosky, S., R. Telang, and A. Acquisti (2011). Do Data Breach Disclosure Laws Reduce Identity Theft? *Journal of Policy Analysis and Management 30*(2), 256–286.

Rosen, J. (2012). The right to be forgotten. *Stanford law review online 64*, 88.

Rubin, P. H. and T. M. Lenard (2001). *Privacy and the Commercial Use of Personal Information.* Kluwer Academic Publishers.

Rust, R., P. Kannan, and N. Peng (2002). The customer economics of Internet privacy. *Journal of the Academy of Marketing Science 30*(4), 455–464.

Ryan, C. (2011). The boundaries of privacy harm. *Indiana Law Journal 86*, 1131–1162.

Samuelson, P. (2000). Privacy as intellectual property. *Stanford Law Review 52*(1125).

Savage, S. and D. Waldman (2013). The value of online privacy. Working Paper.

Schenk, E. and C. Guittard (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics & Management 7*(1), 93–107.

Schoeman, F. D. (1984). *Philosophical dimensions of privacy: An anthology.* Cambridge University Press.

Schoeman, F. D. (1992). *Privacy and social freedom.* Cambridge university press.

Schwartz, P. (1997). Privacy and the economics of personal health care information. *Texas Law Review 76*, 1.

Schwartz, P. (2004). Property, Privacy, and Personal Data. *Harvard Law Review*, 2056–2128.

Schwartz, P. and E. Janger (2007). Notification of data security breaches. *Michigan Law Review 105*, 913–984.

Shaffer, G. and Z. J. Zhang (1995). Competitive coupon targeting. *Marketing Science 14*(4), 395–416.

Shaffer, G. and Z. J. Zhang (2002). Competitive one-to-one promotions. *Management Science 48*(9), 1143–1160.

Smith, R. (1999). The web bug FAQ. Technical report, Electronic Frontier Foundation.

Smith, R. E. (2000). *Ben Franklin's web site: Privacy and curiosity from Plymouth Rock to the Internet.* Privacy Journal.

Solove, D. J. (2002). Identity theft, privacy, and the architecture of vulnerability. *Hastings Law Journal 54*, 1227–1275.

Solove, D. J. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review 154*(3), 477.

Solove, D. J. (2007). *The Future of Reputation - Gossip, Rumor, and Privacy on the Internet.* New Haven and London: Yale University Press.

Solove, D. J. (2013). Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review 126*(7), 1880–1903.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics 87*(3), 355–374.

Spiegel, Y. (2013). Commercial software, adware, and consumer privacy. *International Journal of Industrial Organization 31*(6), 702–713.

Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy 69*(3), 213–225.

Stigler, G. J. (1962). Information in the labor market. *The Journal of Political Economy*, 94–105.

Stigler, G. J. (1980). An introduction to privacy in economics and politics. *The Journal of Legal Studies 9*(4), 623–44.

Stone, B. (2010, March 3rd). Ads posted on Facebook strike some as off-key. *The New York Times*.

Stone, E. and D. Stone (1990). Privacy in organizations: Theoretical issues, research findings, and protection mechanisms. *Research in personnel and human resources management 8*(3), 349–411.

Strahilevitz, L. J. (2008). Privacy versus antidiscrimination. *University of Chicago Law Review 75*(1), 363–381.

Stutzman, F., R. Gross, and A. Acquisti (2013). Silent listeners: The evolution of privacy and disclosure on Facebook. *Journal of Privacy and Confidentiality 4*(2), 2.

Sweeney, L. (1997). Weaving technology and policy together to maintain confidentiality. *The Journal of Law, Medicine & Ethics 25*(2-3), 98–110.

Sweeney, L. (2013). Discrimination in online ad delivery. *ACM Queue 11*(3), 10.

Swire, P. P. and R. E. Litan (1998). *None of Your Business - World Data Flows, Electronic Commerce, and the European Privacy Directive.* Washington, DC: Brookings Institution Press.

Tang, Z., Y. Hu, and M. Smith (2008). Gaining trust through online privacy protection: Self-regulation, mandatory standards, or caveat emptor. *Journal of Management Information Systems 24*(4), 153–173.

Taylor, C. and L. Wagman (2014). Consumer privacy in oligopolistic markets: Winners, losers, and welfare. *International Journal of Industrial Organization 34*, 80–84.

Taylor, C. R. (2004). Consumer privacy and the market for customer information. *RAND Journal of Economics 35*(4), 631–650.

Telang, R. and S. Wattal (2007). An empirical analysis of the impact of software vulnerability announcements on firm stock price. *IEEE Transactions on Software Engineering 33*(8), 544–557.

Tsai, J. Y., S. Egelman, L. Cranor, and A. Acquisti (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research 22*(2), 254–268.

Tucker, C. E. (2012). The economics of advertising and privacy. *International journal of Industrial organization 30*(3), 326–329.

Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. *Journal of Marketing Research 51*, 546–562.

Turow, J., J. King, C. Hoofnagle, A. Bleakley, and M. Hennessy (2009). Americans reject tailored advertising and three activities that enable it. Working paper.

Valentino-Devries, J., J. Singer-Vine, and A. Soltani (2012). Websites vary prices, deals based on users' information. *Wall Street Journal*.

Varian, H. (2010). Computer mediated transactions. *American Economic Review: Papers & Proceedings 100*(2), 1–10.

Varian, H., F. Wallenberg, and G. Woroch (2005). The demographics of the do-not-call list. *IEEE Security and Privacy 3*(1), 34–39.

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

Villas-Boas, J. M. (1999). Dynamic competition with customer recognition. *RAND Journal of Economics 30*(4), 604–631.

Villas-Boas, J. M. (2004). Price cycles in markets with customer recognition. *The Rand Journal of Economics 35*(3), 486–501.

Vissers, T., N. Nikiforakis, N. Bielova, and W. Joosen (2014). Crying wolf? On the price discrimination of online airline tickets. In *Proceedings of the 7th Workshop on Hot Topics in Privacy Enhancing Technologies.*

Wagman, L. (2014). Good news or bad news? Information acquisition and applicant screening in competitive labor markets. Working Paper.

Wagman, L. and V. Conitzer (2014). False-name-proof voting over two alternatives. *International Journal of Game Theory 43*(3), 599–618.

Waldfogel, J. (1994). The effect of criminal conviction on income and the trust "reposed in the workmen". *Journal of Human Resources*, 62–81.

Wang, Y., P. G. Leon, X. Chen, S. Komanduri, G. Norcie, K. Scott, A. Acquisti, L. F. Cranor, and N. Sadeh (2013). The second wave of global privacy protection: From Facebook regrets to Facebook privacy nudges. *Ohio State Law Journal 74*, 1307–1334.

Warren, S. and L. Brandeis (1890). The right to privacy. *Harvard Law Review 4*(5), 193–203.

Wathieu, L. (2002). Privacy, exposure and price discrimination. Technical Report 03-015, Harvard Business School, Division of Research.

Westin, A. (1967). *Privacy and Freedom.* New York: Atheneum Publishers.

White, R. W., N. P. Tatonetti, N. H. Shah, R. B. Altman, and E. Horvitz (2013). Web-scale pharmacovigilance: Listening to signals from the crowd. *Journal of the American Medical Informatics Association*, amiajnl–2012.

White, T. B., D. L. Zahay, H. Thorbjørnsen, and S. Shavitt (2008). Getting too personal: Reactance to highly personalized email solicitations. *Marketing Letters 19*(1), 39–50.

Wilson, K. and J. Brownstein (2009). Early detection of disease outbreaks using the Internet. *Canadian Medical Association Journal 180*(8), 829–831.

Xie, Y., F. Yu, and M. Abadi (2009). De-anonymizing the internet using unreliable IDs. In *ACM SIGCOMM Computer Communication Review*, Volume 39, pp. 75–86.

Zandt, T. V. (2004). Information overload in a network of targeted communication. *RAND Journal of Economics 35*(3), 542–560.

Zhang, J. (2011). The perils of behavior-based personalization. *Marketing Science 30*(1), 170–186.