2014

From the Economics of Privacy to the Economics of Big Data

Alessandro Acquisti
Carnegie Mellon University, acquisti@andrew.cmu.edu

Published In

Privacy, Big Data, and the Public Good: Frameworks for Engagement; Stefan Bender, Julia Lane, Helen Nissenbaum, and Victoria Stodden (eds), 76-95.

From the Economics of Privacy to the Economics of Big Data

Alessandro Acquisti

Draft[1]

1 Introduction

Imagine a world in which consumers' preferences can be so precisely estimated by observing their online behavior, that firms are able to anticipate consumers' needs, offering the right product at exactly the right time. Imagine that same world, but now consider that extensive knowledge of consumers' preferences also allows precise inferences about their reservation prices (the maximum price each consumer will pay for a good), so that firms can charge different prices for the same product to each of their buyers, and absorb the entire surplus arising from an economic transaction.

Imagine a world in which the collection and analysis of individual health data allow researchers to discover the causes of rare deceases and the cures for common ones. Now, consider the same world, but imagine that employers are able to predict job candidates' future health conditions from few data points extracted from the latter's social network

[1] This manuscript is an early draft for a chapter on "The Economics and Behavioral Economics of Privacy" prepared for Lane, Julia, Victoria Stodden, Stefan Bender, and Helen Nissenbaum, eds. Privacy, Big Data, and the Public Good: Frameworks

profiles – and then, imagine those employers making hiring decision based on those predictions, without the candidate's consent or even awareness.

The economics of privacy attempts to study the costs and benefits associated with personal information – for the data subject, the data holder, and for society as a whole. As a field of research, it has been active for some decades. Progresses in data mining, business analytics, and so-called big data, have the potential for magnifying the size and augmenting the scope of economic benefits and dangers alike. This chapter overviews the growing body of theoretical and empirical research on the economics and behavioral economics of privacy, and discusses how these streams of research can be applied to the investigation of the implications of consumer data mining and business analytics. Among the many possible interpretations of privacy, this capture focuses on its informational aspects: the trade-offs arising from the protection or disclosure of personal data.

Since the second half of the last century, progresses in information technology and the transformation of advanced economies into service economies have made it possible for organizations to monitor, collect, store, and analyze increasing amounts of individual data. Those progresses have also raised significant, and in some cases novel, privacy concerns. Attempting to analyze privacy in the age of big data from an economic perspective does not imply the assumption that all modern privacy issues have explicit monetary dimensions. Rather, this type of analysis stems from the realization that, with or without individuals' awareness, decisions that data subjects and data holders make about personal data often carry complex trade-off. The mining of personal data can help increase welfare, lower search costs, and reduce economic inefficiencies; at the same

---

for Engagement. Cambridge University Press, 2014.

time, it can be source of losses, economic inequalities, and power imbalances between those who hold the data and those whose data is controlled. For instance, a firm may reduce its inventory costs by mining and analyzing the behavior of many individual consumers; however, the infrastructure needed to carry out analysis may require substantial investments, and if the analysis is conducted in manners that raise consumers' privacy concerns, those investments may backfire. Likewise, a consumer may benefit from contributing her data to a vast database of individuals' preferences (for instance, by sharing music interests with an online vendor, and receiving in turn targeted recommendations for new music to listen to); that same consumer, having lost control over that data, may end up suffering from identity theft, price discrimination, or stigma associated with the information unintended parties can acquire about her.

This chapter offers an overview of the lessons that economics (and behavioral economics) can tell us regarding privacy in the age of big data. The chapter suggests that the microeconomic theory of privacy has brought forward arguments both supporting the view that privacy protection may increase economic efficiency in a marketplace, and decreases it: personal information, when shared, can become a public good whose analysis can reduce inefficiencies and increase economic welfare; when abused, it can lead to transfer of economic wealth from data subjects to data holders. Similarly, empirical evidence has been offered of both the benefits and costs, for data subjects and data holders alike, of privacy protection. It is unlikely that economics can answer questions such as what is the "optimal" amount of privacy and disclosure for an individual and for society – but it can help us think about the trade-offs associated with personal information.

The rest of this chapter first provides a brief summary of some relevant results from the micro economic theory of privacy (Section 2). It then describes the potential trade-offs associated with privacy and disclosure in the age of big data (Sections 3 and 4), consumers' privacy valuations (Section 5), and behaviors (Section 6). It concludes by discussing the role of privacy enhancing technologies and market forces (Section 7) in balancing the value of data and the value of privacy.

2 Privacy and Economic Theory

Among the many heterogeneous dimensions of privacy (Solove 2006), formal economic analysis has predominantly (albeit not solely) focused on privacy as concealment of personal information – a form of information asymmetry (Akerlof 1970). For instance, before a consumer interacts with a seller, the seller may not have knowledge of the consumer's "type" (such as her preferences, or her reservation price for a product). After the consumer has interacted with the seller (for instance, she has completed a purchase of a certain product at a certain price), it is the consumer who may not know how the seller is going to use the information it acquired through the transaction. The former is a case of hidden information; the latter, of hidden action.

Some of the earliest explicit economic discussions of privacy appeared in the literature near the end of the 1970s and the beginning of the 1980s – thanks in particular to the work of scholars belonging to the so-called Chicago School. Among them, Stigler (1980) argued that the protection of privacy may lower the quality of information about economic agents available in the marketplace. Hence, excessive protection of privacy rights may end up being economically inefficient and redistributive, as it may deny to the

market the signals needed to allocate, compensate, and efficiently price productive factors. Similarly, Posner (1981) argued that concealing personal information may transfer costs from one party to another: for instance, the employer who cannot fully scrutinize the job candidate may end up paying the price of hiring an unsuitable employee. According to this view, legislative initiatives that favor privacy protection by restricting the activities of companies are likely to create inefficiencies, raise firm costs, and ultimately decrease economic welfare.

Hirshleifer (1980), however, took positions that may be considered alternative to Stigler and Posner. He noted that economic studies based on the assumption of neo-classically rational economic agents may not adequately capture the nuances of transactions that occur outside the logic of the market, such as those involving privacy. Earlier, Hirshleifer (1971) had also noted that investment in private information gathering may be inefficient: using private information may have redistributive effects, which leads to overinvestment in information gathering. A similar conclusion is found by Taylor (2004b), who finds that market forces alone may not guarantee efficient economic outcomes (under competition, firms have a private incentive to invest more than socially optimal into collecting larger amount of consumer data).

The contraposition of the results found by Stigler (1980) or Posner (1981) and those by Hirshleifer (1971) or Taylor (2004b) highlights a common theme in the economic literature on privacy: privacy costs and privacy benefits are inextricably related. Works by Varian (1996), Noam (1996), Taylor (2004a) and Acquisti and Varian (2005) offer further examples.

Varian (1996) noted that a general ban on the dissemination of personal data would not be in the interest of the consumer herself, as well as the firms she interacts with. A consumer may naturally be interested in disclosing certain personal traits to other firms (for example, her preferences and tastes, so as to receive services). However, the same consumer may have an interest in keeping other types of information private (for example, her reservation price for a particular good). Noam (1996), applying Ronald Coase's 'theorem' to the study of privacy, argued that in absence of transaction costs, the interaction between consumers interested in the protection of their data and firms interested in accessing will, under free market exchanges, lead to an equilibrium in which the agent with the greatest interest in either accessing the data or protecting it from access will be the one to actually achieve its goal – independently of the initial assignments of privacy rights or rights over access to consumer data. However, transaction costs and uncertainties regarding the initial assignment of rights over personal information are likely to be substantial in the interaction between consumers and firms, in which case it would no longer be guaranteed that market forces alone would produce the most efficient privacy outcomes. Similarly, both Taylor (2004a) and Acquisti and Varian (2005) studied the economic impact of tracking technologies that make customer identification possible. In these types of models (which typically analyze intertemporal interactions between consumers and merchants, and focus on consumers' reservation prices as private information the merchant is interested in inferring), when consumers are rational decision makers, a regulatory regime for privacy protection turns out not to be necessary. For instance, in Acquisti and Varian (2005), consumers who expect to be tracked can engage

in strategic behaviors that render tracking counterproductive; to avoid this, firms must use consumer information to offer personalized services that consumers will value.

This series of microeconomic results suggests that not only does privacy protection (or lack thereof) carry both potential costs and potential benefits for data subjects and data holders alike; but also that economic theory should not be expected to be answer the question "what is the economic impact of privacy (or lack thereof) on consumer and aggregate welfare?" in an unambiguous, unequivocal manner. Economic analysis certainly can help us carefully investigate local trade-offs associated with privacy, but the economic consequences of privacy are nuanced, and the evolution of technologies for data mining and business intelligence are more likely to emphasize, rather than resolve, those nuances, as we highlight in the following section.

3 Markets for Privacy

Due to the concurrent evolution of Internet technologies, online business models, and data mining and business analytics tools, economic transactions of privacy relevance occur nowadays in different types of market. We distinguish three markets for privacy in this section.

The first type of transactions that have privacy relevance actually occur in the market for ordinary, non-privacy goods: in the process of exchanging or acquiring other products or services, individuals often reveal personal information, which may be collected, analyzed, and then used by the counterpart in the transaction in a variety of ways. In this case, the exchange of personal data, and the privacy implications of such exchange, are a secondary aspect of a primary transaction involving a good which is not, per se, privacy

related. An example of a transaction happening in this market may be the purchase of a book completed on an online merchant's site.

The second type of privacy-related transactions occur in what may be called the market for personal data. This market itself includes a variety of exchanges. One form of exchange involves 'infomediaries' that trade consumer data among themselves or with other data-holding firms. For instance, firms like Acxiom or credit reporting agencies like Transunion both acquire from, and sell to, consumer data by interacting with other consumer-facing firms. The data subjects are not generally active agents in these transactions. A second form of exchange involves so-called free products or services provided to consumers in exchange for their data. This market includes search engines and online social networks. In these exchanges, consumers are directly involved in the transaction, although the exchange of their personal information is not always a *visible*, explicit component of the transaction: while the price for services in this type of exchanges may be nominally zero, the customer is effectively purchasing the service by selling her data.

A third form of privacy-related transactions occur in what may be called the market for privacy. In this market, consumers explicitly seek products and services to manage and protect their personal information. For instance, they may acquire a privacy enhancing technology to protect their communications or hide their browsing behavior. The business models associated with providing consumers with more protection over their data have evolved rapidly, also due to the attention paid to the potential benefits associated with the sharing and mining of consumer data. Indeed, some business models end up being a bridge between the market for privacy and the market for personal data, in

that they aim at giving consumers more ‘ownership’ over (exchanges involving) their personal information, including – sometimes – the potential ability to monetize it.

4 Privacy Trade-offs

The evolution and success of data mining and business analytics tools is and will keep affecting the markets described in the previous sections and their emerging trade-offs, especially in the form of both positive and negative externalities that arise when a consumers’ data are aggregated and analyzed together with the data of many other consumers.[1] As anticipated in Section 2, the resulting costs and benefits for data subjects, data holders, and society at large are complex and nuanced. On the one hand, expected benefits can emerge from disclosed data for both data holders and data subjects (as well as opportunity costs when information is not shared or collected), together with the expected costs of the investments necessary to collect and process that data. On the other hand, expected benefits can arise from *protecting* data and expected costs can arise from privacy intrusions; however, costs are also associated with the protection of personal data. While a complete analysis of such dual benefits and costs associated with either sharing or protecting data are outside the scope of this chapter, in this section we provide a few key examples that are especially relevant to the context of data mining and business analytics.

We first consider some of the benefits of data sharing, as well as some of the costs associated with data protection.

Firms can capitalize in various ways on the data of current and potential customers. Detailed knowledge of a consumer’s preferences and behavior can help firms better target their products or ads, lowering advertising costs (Blattberg and Deighton 1991), providing

enhanced, personalized services (Acquisti and Varian 2005), increasing consumer retention and loyalty, but also enforcing profit-enhancing price discrimination (Varian 1985) (although the latter may not always be the case in presence of competition; Fudenberg and Tirole 2000). For instance, the granular targetability made possible by online advertising may increase revenues for marketers and merchants (according to Beales, 2010, the price of behaviorally targeted advertising is almost three times as much the price of untargeted advertising). Similarly, by aggregating consumer data, firms can forecast trends and predict individual preferences, leading to the ability to provide valuable product recommendations (Bennett and Lanning 2007), or improve or redesign services based on observed behavior. Furthermore, revenues from targeted consumers may allow firms to provide lower-cost versions of a product to other consumers, and support services provided to consumers at a price of zero – but in exchange for their data.

Indeed, some of the benefits data holders gain from data may get passed on, or shared with, data subjects themselves (Lenard and Rubin 2009; Goldfarb and Tucker 2010), in the form of free content or services (made possible by advertising or personal data trades), personalized services, reduced search costs, or more efficient interactions with merchants or their sites. Positive consumer externalities may also materialize. For instance, better consumer data may allow firms to target the right consumers, reducing the amount of marketing investment that gets wasted with consumers uninterested in the product, and potentially leading to lower product prices (see also Blattberg and Deighton 1991), or aggregation of web searches of many individuals could help detect disease outbreaks (Wilson and Brownstein 2009), or the aggregation of location data could be used to

improve traffic conditions and reduce road congestion. In other words, the aggregation of private data could create a public good, with societal benefits accruing from big data.

One should note, however, that many of the benefits consumers can enjoin from data sharing may also be obtained without the disclosure of *personally identified* data. In other words, repeating benefits of "big data" while protecting privacy may not necessarily be contradictory goals: as further discussed in Section 7, advancements in privacy enhancing technologies suggest that there exist many shades of grey between the polar extremes of absolute sharing and complete protection of personal data; rather, it is possible to selectively protect or disclose different types of personal information, and modulate their identifiability, in order to optimize privacy trade-offs for individuals and society as a whole. As a result, benefits from data may be gained also when data is protected, and the actual societal costs of privacy protection may turn to be limited. For instance, the privacy enhancing provisions of the Fair Credit Reporting Act did not raise as significant barriers to profitable uses of consumer data as critics of the Act feared before its passage (Gellman 2002); the possible reduction of ads effectiveness caused by regulation limiting behavioral targeting may simply be offset by using ads on sites with specific content, larger ads, or ads with interactive, video, or audio features (Goldfarb and Tucker 2010); and certain types of privacy regulation in healthcare may actually foster innovation in the form of higher probability of success of Health Innovation Exchanges (Adjerid et al. 2013).

As for the costs that come from privacy violations or disclosed data, they can be both tangible and intangible both data holders and data subjects alike.

From the perspective of the data subject, Calo (2011) distinguishes between subjective and objective privacy harms: the former derive from unwanted perceptions of observation; the latter consist of the unanticipated or coerced use of information concerning a person against that person. Hence, the former relate to the anticipation, and the latter to the consequences, of losing control over personal information. Subjective harms may include anxiety, embarrassment, or fear; the psychological discomfort associated with feeling surveilled; the embarrassment associated when sensitive information is exposed publicly; or the chilling effects of fearing one's personal life will be intruded. These harms may be hard to capture and evaluate in economic terms, and usually are not recognized by U.S. courts as *actual* damage (Romanosky and Acquisti 2009). Objective harms could be both immediate and tangible, and indirect and intangible. They could include the damages caused by identity theft, the efforts spent deleting (or avoiding) junk mail; the time spent dealing with annoying telemarketing; the higher prices one pays due to (adverse) price discrimination; but also the effects of profiling, segmentation, and discrimination. For instance, profiling could be used to nudge consumers towards products that may not enhance their well-being,[2] and information revealed on a social network may lead to job market discrimination (Acquisti and Fong 2012). In more general terms, as an individual's data is shared with other parties, those parties may gain a bargaining advantage in future transactions with that individual. For instance, while a consumer, thanks to behavioral advertising, may receive targeted ads for products she is actually interested in, other entities (such as marketers and merchants) will accumulate data about the consumer that may permit the creation of a detailed dossier of her preferences and tastes, and the prediction of her future behavior.

As noted in Section 2, microeconomic models predict that, in presence of myopic customers, this information will affect the allocation of surplus of future transactions, increasing the share of the data holder over that of the data subject. Ultimately, the disclosure of personal data affects the balance of power between the data subject and the data holder.

Data holders can also, under certain conditions, bear costs from the misuse of consumers data. For instance, Romanosky and Acquisti (2009) note that, following a data breach, firms suffer in terms of consumer notification costs, fines, settlement costs, stock market losses, or loss of consumer's trust. However, it may often be the case that a firm could externalize the privacy costs of using consumer data, while internalizing much of the gains (Swire and Litan 1998).

Often, objective privacy harms are merely probabilistic: once data is revealed or intruded, it may or may not lead to the actual negative consequences we have just described. For instance, poor data handling practices by a consumer report firm may later cause a consumer's mortgage request to be wrongfully denied; or, the breach of a database containing consumers' credit cards may later lead to identity theft (Camp 2007). The metaphor of a 'blank check' has been used to refer to the uncertainty associated with privacy costs: disclosing personal information is like signing a blank check, that may never be cashed in – or perhaps cashed it at some unpredictable moment in time with an indeterminably low, or high, amount to pay. In economic terms, the damage from disclosed data are, in Knight (1921)'s terms, *ambiguous* and, up to a point, unknownable.[3] Consider, in fact, that some privacy costs are high-probability events with negligible individual impact (for instance, spam); other costs are low-probability events with very

significant adverse consequences (for instance, some of the costs associated with the more pernicious forms of identity theft). Because of this and their often intangible dimensions, privacy costs may be hard to assess and therefore also to act upon. Either because of low likelihood of occurrence, or limited perceived magnitude of damage, privacy costs may therefore dismissed as unimportant at the individual level – even when, in the aggregate, they may amount to significant societal damage, or a significant transfer of wealth from data subjects to others (including the data holders).

5 Do Consumers Value Privacy?

Farrell (2012) notes that privacy is both a final and an intermediate good: "[c]onsumers care about privacy in part for its own sake: many of us at least sometimes feel it's just icky to be watched and tracked. [...] Consumers also care about privacy in a more instrumental way. For instance, loss of privacy could identify a consumer as having a high willingness to pay for something, which can lead to being charged higher prices if the competitive and other conditions for price discrimination are present." In this section, we summarize a number of empirical investigations of consumers' privacy valuations. In the following section (Section 6), we examine the hurdles consumers face in making privacy decisions consistent with those valuations.

Numerous factors influence individuals' privacy concerns (Milberg et al. 1995), and therefore the mental 'privacy calculus' that individuals make when deciding whether to protect or disclose personal information (Laufer and Wolfe 1977; Culnan and Armstrong 1999; Dinev and Hart 2006). Researchers from diverse disciplines (such as economics, marketing, information systems, and computer science) have attempted to estimate

empirically the value that, in this calculus, individuals assign to privacy and their personal data. The resulting findings suggest that privacy valuations are significantly context dependent. Furthermore, willingness to pay, or reservation prices, may not adequately capture the value of privacy for those individuals who simply do not feel they should have to pay to protect their privacy.

Huberman et al. (2005) used a second-price auction to estimate the price at which individuals were willing to publicly reveal personal information such as their weight. Individuals whose weight was more deviant from the perceived norm for the rest of the group were more likely to exhibit higher valuations. Wathieu and Friedman (2005) found that survey participants were more acceptive of an organization sharing their personal information after having been explained the economic benefits of doing so. Cvrcek et al. (2006) reported large differentials across EU countries in the price EU citizens would accept to share mobile phone location data. Hann et al. (2007) focused on online privacy and, using a conjoint analysis, found that protection against errors, improper access, and secondary use of personal information was worth US$30.49-44.62 among U.S. subjects. Rose (2005) found that, although most participants in a survey self-reported being very sensitive to privacy issues, less than half of them would be willing to pay roughly $29 to have their privacy protected by means of property rights on personal information. Both Varian et al. (2005) and Png (2007) estimated U.S. consumers' implicit valuation of protection from telemarketers using data about the Do Not Call list adoptions. They found highly differing values, from a few cents to as much as $30. Tsai et al. (2011) found that, when information about various merchants' privacy policies was made available to them

in a compact and salient manner, subjects in an experiment were more likely to pay premia of roughly 50 cents to purchase products from more privacy protective merchants.

At the same time, various studies have highlighted a dichotomy between self professed privacy attitudes and actual self-revelatory behavior.

Tedeschi (2002) reported on a Jupiter Research study in which the overwhelming majority of surveyed online shoppers would give personal data to new shopping sites for the chance to win $100. Spiekermann et al. (2001) found that even participants in an experiment who could be classified as privacy conscious and concerned were willing to trade privacy for convenience and discounts: differences across individuals in terms of reported concerns did not predict differences in self-revelatory behavior. Similar findings were obtained in different settings by Acquisti and Grossklags (2005) and Acquisti and Gross (2006). Coupled with the observation that businesses focused on providing privacy enhancing applications have met difficulties in the marketplace (Brunk 2002), these results suggest a potential privacy paradox: people want privacy, but do not want to pay for it, and in fact are willing to disclose sensitive information for even small rewards (for an overview of this area, see Acquisti, 2004, and Acquisti and Grossklags, 2007). In fact, Acquisti et al. (2013) have recently presented an application of the endowment effect to the privacy domain: subjects who started an experiment from positions of greater privacy protection were found to be five times more likely than other subjects (who did not start with that protection) to forego money to preserve their privacy. These results illustrate the challenges with pinpointing exact valuations of personal data: consumers' privacy valuations are not only context dependent, but affected by numerous heuristics and biases (see Section 6), and so are individuals' decisions to share or to protect personal

information (John et al. 2011). In addition, awareness of privacy risks (and potential solutions to privacy threats) may also significantly affect consumers' privacy choices valuations – which is why revealed preferences arguments (that rely on observing consumer's choices in the marketplace – for instance, in the case of privacy, their propensity to share data online or to use protecting technology) may not necessarily provide the fuller or clearer picture of what privacy is ultimately worth to individuals. We discuss some of those hurdles that affect privacy decision making and valuations in the following section.

6 Hurdles in Privacy Behavior

A stream of research investigating the so-called privacy paradox has focused on the hurdles that hamper individuals' privacy-sensitive decision making. If consumers act myopically, or not fully rational (in the neo-classical economic sense of utility-maximizing, Bayesian-updates agents who make use of all the information consumers available to them), then market equilibria may not in fact guarantee privacy protection. In fact, in absence of regulatory protection of consumers' data, firms will tend to extract the surplus generated in transaction in which consumers' data is used for price discrimination (Acquisti and Varian 2005; Taylor 2004a).

There is, indeed, evidence that consumers face known decision making hurdles when facing privacy trade-offs, such as (a) incomplete information, (b) bounded cognitive ability to process the available information, and (c) a number heuristics (or cognitive and behavioral biases) which lead to systematic deviations from theoretically rational decision

making (sometimes, various combinations of these factors affect consumer decision making at the same time).

Consider, first, the problem of incomplete information. In many scenarios – such as those associated with behavioral monitoring and targeting – the consumer may not even realize the extent at which her behavior is being monitored and exploited. Furthermore, after an individual has released control on her personal information, she is in a position of information asymmetry with respect to the party with whom she is transacting. In particular, the subject might not know if, when, and how often the information she has provided will be used. For example, a customer might not know how the merchant will use the information that she has just provided to the merchant through a website.

Furthermore, the 'value' itself of the individual's information might be highly uncertain and variable. The subject and the parties she is interacting with may evaluate differently the same piece of information, and the specific environmental conditions or the nature of the transaction may affect the value of information in unpredictable ways. For example, a customer might not know what damage she will incur because of her personal information becoming known, she might not know how much profit others will make thanks to that information, or she might not know the benefits she will forego if her privacy is violated. To what, then, is the subject supposed to anchor the valuation of her personal data and its protection?

Second, findings from behavioral economics exhaustively document consumers' inability to exhaustively consider the possible outcomes and risks of data disclosures, due to bounded rationality. Furthermore, the individual will often find herself in a weaker bargaining position than other parties she is interacting with (for instance, merchants). In

many transactions, the individual is unable to negotiate a desired level of information protection; she rather faces take-it-or-leave-it offers of service in exchange for personal data.

Third, even if the consumer had access to complete information about all trade-offs associated with data sharing and data protection, she will suffer from cognitive and behavioral biases that are more intense in scenarios where preferences are more likely to be uncertain. One such example is that, if the expected negative payoff from privacy invasions could be estimated, some individuals might seek immediate gratification, discounting hyperbolically (Rabin and O'Donoghue 2000) future risks (for example of being subject to identity theft), and choosing to ignore the danger. Hence, because of asymmetric information, self-gratification bias, overconfidence, or various other forms of misrepresentation studied in the behavioral economic literature, individuals might choose not to protect their privacy *possibly* against their own best interest. They might be acting *myopically* when it comes to protecting their privacy even when they might be acting *strategically* (as rational agents) when bargaining for short-term advantages such as discounts (Acquisti 2004).

Consider, for instance, the case of data breaches. As discussed in Romanosky and Acquisti (2009), after being notified of a breach of her financial information, a consumer may not be able to identify the right course of action: should she, for instance, punish the financial firm that, due to faulty security controls, compromised her data, by changing to a competitor? While this may appear as a risk-reducing behavior, by doing so the consumer would have now disclosed her personal information to another firm – and actually materially increased the probability that another future breach will involve her

data. Furthermore, the cost of acting may be significant: calling the breached firm to obtain details about the breach and its consequences, notifying financial institutions of the occurred breach and of potentially compromised accounts, or subscribing to credit alert and insurance services, are all actions which carry perceived cognitive, transaction, and actual costs. Such costs may appear greater to the consumer than the perceived benefit from action. It could also be that, because of psychological habituation due to repeated instances of data breaches report in the media, the consumer may become desensitized to their effects – which counter the desired impact of notifications. Ultimately, the consumer may 'rationally' decide to remain 'ignorant' (following the Choicepoint breach, fewer than 10% of affected individuals availed themselves of the free credit protection and monitoring tools offered by Choicepoint; Romanosky and Acquisti 2009). This example suggests how nuanced and full of obstacles is the path that lead from consumer notification of privacy problem to her actually taking action to solve that problem.

Based on these hurdles, recent behavioral privacy research has questioned the validity and effectiveness of regimes based on transparency and control mechanism (also known as choice and notification; Brandimarte et al. 2013; Adjerid, Acquisti, Brandimarte, and Loewenstein, Adjerid et al.; Acquisti et al., 2013).[4]

An improved understanding of cognitive and behavioral biases that hamper privacy (and security) decision making, however, could also be exploited for normative purposes. Specifically, knowledge of those biases could be used to design technologies and policies that anticipate and counter those very biases (Acquisti 2009). Such technologies and policies would be informed by the growing body of behavioral economics research on soft or asymmetric paternalism (Loewenstein and Haisley 2008) as well as research on

privacy and security usability. They may help consumers and societies achieve their desired balance between information protection and information sharing.

7 Technology, Regulation, and Market Forces

As noted in Acquisti (2010), progresses in computer science, statistics, or data mining have not only produced potentially privacy-eroding business analytics tools or big data technologies; they have also led to the development, over the past few decades, of privacy enhancing technologies which allow the protection of (certain) individual data simultaneously to the sharing, or analysis, or aggregate, de-identified, or non-sensitive identified data. Online activities and transactions for which privacy preserving correspondents exist include electronic payments (Chaum 1983), online communications (Chaum 1985), Internet browsing (Dingledine et al. 2004), credentials (Camenisch and Lysyanskaya 2001), or even online recommendations (Canny 2002). One of the most interesting direction of research relates to executing calculations in encrypted spaces (Gentry 2009), and whether these types of computations will make it possible to have both privacy *and* big data, confidentiality *and* analytics. In the best-case scenario, the deployment of privacy enhancing technologies may result in a win–win for data holders and data subjects: certain data is protected (thereby avoiding costs associated with certain privacy intrusions), whereas other data gets shared, analyzed, and used (thereby enjoying the benefits and the value of data, big or small). Alternatively, the old economic adage that there is no free lunch may apply: whenever protection is applied to a dataset, the utility of that dataset is decreased (Duncan et al. 2001). The interesting economic question then becomes, whose utility will be adversely affected – or, in other words, who will bear

the costs if privacy enhancing technologies become more popular in the age of big data: data subjects (whose benefits from business analytics and big data would shrink with the amount if information they share), data holders (who may face increasing costs associated with collecting and handling consumer data), or both?

Attempting to answer the above question remains an open research question. An additional and related open question, however, is whether, even if privacy enhancing technologies were found to increase data subjects' welfare more than they would adversely affect data holders' welfare, market forces alone would lead to the deployment and success of those technologies. While there is no lack of evidence online of both disclosure/publicity seeking and privacy seeking behavior, privacy enhancing technologies (as opposed to security technologies such as anti-viruses or firewalls) have not gained widespread adoption. Several reasons may explain this situation: on the consumers' side, a first obvious explanation is low consumer demand for privacy; however, other, more nuanced (and non–mutually exclusive) explanations include users' difficulties and costs in using privacy technologies (see Whitten and Tygar 1999), switching costs, as well as biases such as immediate gratification, which reduce demands for those products even by privacy sensitive consumers. On the data holders' side, in absence of regulatory intervention, or of clear evidence that privacy protection can act as a distinctive source of competitive advantage for a firm, it is unlikely that firms will incur the costs to transition to technologies that may, in the short run, limit their access to consumer data relative to their competitors.

The debate over the comparative economic advantages of regulation and self-regulation of privacy remains intense to this date. On the one hand, Gellman (2002)

challenges the view that the unrestricted trafficking in personal information always benefits the consumer, and that privacy trade-offs may merely be evaluated on the basis of monetary costs and benefits. He concludes that an unregulated, privacy-invasive market in personal data can be costly for consumers. F. H. Cate (2002), Cate et al. (2003), Rubin and Lenard (2001), and Lenard and Rubin (2009), on the other hand, claim that legislative initiatives that restrict the amount of personal information available to business would actually penalize the consumers themselves: regulation should be undertaken only when a given market for data is not functioning properly, and when the benefits of new measures outweigh their costs.

It may not be possible to resolve this debate using purely economic tools. Economic theory, as we have discussed above, has brought forward arguments both supporting the view that privacy protection *increases* economic efficiency, and that it *decreases* it. Empirically, the costs and benefits associated with the protection and revelation of consumers' data have not proven easily amenable to aggregation: First, as soon as one attempts an aggregate evaluation of the impact of privacy regulation, one faces the challenge of delimiting the problem: data breaches, identity theft, spam, profiling, or price discrimination are all examples of privacy problems, yet they comprise very different expected benefits and costs for the parties involved. Second, even within each scenario, it may be hard to statically measure at a point in time the aggregate costs and benefits of data protection and data sharing, since the benefits and costs of privacy happen over time (for instance, data revealed today may only damage the individual years from now). And third, in addition to measurable outcomes (such as the financial losses due to identity theft, or the opportunity costs of spam), other privacy invasions require an estimation of

consumers' valuations of privacy. Furthermore, as we have noted elsewhere in this chapter, numerous of the benefits associated with data disclosure may, in fact, still be gained when data is protected. Evaluations and conclusions regarding the economic value of privacy and the optimal balance between disclosure and protection are, therefore, far from simple.

Acknowledgement The author would like to thank the editors, the anonymous reviewers, Veronica Marotta, and Laura Brandimarte for particularly insightful comments and suggestions. This chapter is partly based on previous work by the author, including Acquisti (2010) and Brandimarte and Acquisti (2012).

Notes

[1] This section is based on material previously discussed in Acquisti (2010, sec. 3).

[2] Some consumer data firms advertise databases with contacts to individuals suffering from various types of addiction, such as gambling (see http://www.dmnews.com/media-one-gamblers-database/article/164172/).

[3] Knight (1921) distinguished between *risk* (the random outcomes of an event can be described with a known probability distribution) and *ambiguity* (those probabilities are unknown).

[4] See also Chapter 2 of this volume, by Barocas and Nissenbaum.

References

Acquisti, A. 2004. Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the ACM Conference on Electronic Commerce (EC '04)*, pp. 21–29.

Acquisti, A. 2009. Nudging privacy: The behavioral economics of personal information. *IEEE Security & Privacy* 7(6): 82–85.

Acquisti, A. 2010. The economics of personal data and the economics of privacy. Background Paper for OECD Joint WPISP-WPIE Roundtable.

Acquisti, A., I. Adjerid, and L. Brandimarte. 2013. Gone in 15 seconds: The limits of privacy transparency and control. *IEEE Security & Privacy* 11(4): 72–74.

Acquisti, A., and C. Fong. 2012. An experiment in hiring discrimination via online social networks. In *Privacy Law Scholars Conference*.

Acquisti, A., and R. Gross. 2006. Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *Workshop on Privacy Enhancing Technologies (PET '06)*.

Acquisti, A., and J. Grossklags. 2005. Privacy and rationality in individual decision making. *IEEE Security and Privacy* 3(1): 24–30.

Acquisti, A., and J. Grossklags. 2007. What can behavioral economics teach us about

privacy? In *Digital Privacy: Theory, Technologies and Practices*, ed. S. G. C. L. Alessandro Acquisti and Sabrina De Capitani di Vimercati, 363–377. Boca Raton, FL: Auerbach Publications.

Acquisti, A., L. K. John, and G. Loewenstein. 2013. What is privacy worth? *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. Sleights of privacy: Framing, disclosures, and the limits of transparency.

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

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

Beales, H. 2010. The value of behavioral targeting. Network Advertising Initiative.

Bennett, J., and S. Lanning. 2007. The Netflix prize. In *Proceedings of KDD Cup and Workshop*.

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

Brandimarte, L., and A. Acquisti. 2012. The economics of privacy. In *Handbook of the*

Digital Economy, ed. M. Peitz and J. Waldfogel. New York: Oxford University Press.

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

Brunk, B. D. 2002. Understanding the privacy space. *First Monday* 7(10).

Calo, R. 2011. The boundaries of privacy harm. *Indiana Law Journal* 86.

Camenisch, J., and A. Lysyanskaya. 2001. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In *Advances in Cryptology - EUROCRYPT '01*, LNCS 2045, 93–118. Heidelberg: Springer.

Camp, L. J. 2007. *Economics of Identity Theft: Avoidance, Causes and Possible Cures.* New York: Springer.

Canny, J. F. 2002. Collaborative filtering with privacy. In *IEEE Symposium on Security and Privacy*, 45–57.

Cate, F. H. 2002. Principles for protecting privacy. *Cato Journal* 22(1): 33–57.

Cate, F. H., R. E. Litan, M. Staten, and P. Wallison (2003). Financial privacy, consumer prosperity, and the public good: Maintaining the balance. Federal Trade Commission Workshop on Information Flows: The costs and benefits to consumers and businesses of the collection and use of consumer information.

Chaum, D. 1983. Blind signatures for untraceable payments. In *Advances in Cryptology*, 199–203. New York: Plenum Press.

Chaum, D. 1985. Security without identification: Transaction systems to make big brother

obsolete. *Communications of the ACM* 28(10): 1030–1044.

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.

Cvrcek, D., M. Kumpost, V. Matyas, and G. Danezis. 2006. The value of location information. In *ACM Workshop on Privacy in the Electronic Society (WPES)*.

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

Dingledine, R., N. Mathewson, and P. Syverson. 2004. Tor: The second-generation onion router. In *Proc. 13th Conference on USENIX Security Symposium*, 13:21.

Duncan, G. T., S. A. Keller-McNulty, and S. L. Stokes. 2001. Disclosure risk vs. data utility: The ru confidentiality map. In *Chance*.

Farrell, J. 2012. Can privacy be just another good? *Journal on Telecommunications and High Technology Law* 10:251–445.

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

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. March.

Gentry, C. 2009. Fully homomorphic encryption using ideal lattices. In *Proc. 41st Annual*

ACM symposium on Theory of Computing, 169–178.

Goldfarb, A., and C. Tucker. 2010. Privacy regulation and online advertising. Available at SSRN: http://ssrn.com/abstract=1600259.

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

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

Hirshleifer, J. 1980. Privacy: Its origins, function and future. *Journal of Legal Studies* 9, no. 4 (December): 649–664.

Huberman, B. A., E. Adar, and L. R. Fine. 2005. Valuating privacy. *IEEE Security & Privacy* 3:22–25.

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

Knight, F. 1921. *Risk, Uncertainty and Profit*. Boston: Hart, Schaffner & Marx; Houghton Mifflin.

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.

Lenard, T. M., and P. H. Rubin. 2009. In defense of data: Information and the costs of

privacy. Technology Policy Institute.

Loewenstein, G., and E. Haisley. 2008. The economist as therapist: Methodological issues raised by light paternalism. In *Perspectives on the Future of Economics: Positive and Normative Foundations*, ed. A. Caplin and A. Schotter. New York: Oxford University Press.

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.

Noam, E. M. 1996. Privacy and self-regulation: Markets for electronic privacy. In *Privacy and Self-Regulation in the Information Age*. National Telecommunications and Information Administration.

Png, I. 2007. On the value of privacy from telemarketing: Evidence from the 'Do Not Call' registry. Working Paper, National University of Singapore.

Posner, R. A. 1981. The economics of privacy. *American Economic Review* 71, no. 2 (May): 405–409.

Rabin, M., and T. O'Donoghue. 2000. The economics of immediate gratification. *Journal of Behavioral Decision Making* 13(2): 233–250.

Romanosky, S., and A. Acquisti. 2009. Privacy costs and personal data protection: Economic and legal perspectives. *Berkeley Technology Law Journal* 24(3).

Rubin, P. H., and T. M. Lenard. 2001. *Privacy and the Commercial Use of Personal*

Information. Boston: Kluwer Academic Publishers.

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

Spiekermann, S., J. Grossklags, and B. Berendt. 2001. E-privacy in 2nd generation e-commerce: Privacy preferences versus actual behavior. In *3rd ACM Conference on Electronic Commerce.*

Stigler, G. J. 1980. An introduction to privacy in economics and politics. *Journal of Legal Studies* 9, no. 4 (December): 623–44.

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.

Taylor, C. R. 2004a. Consumer privacy and the market for customer information. *RAND Journal of Economics* 35(4): 631–651.

Taylor, C. R. 2004b. Privacy and information acquisition in competitive markets. Technical report, Duke University, Economics Department, 03-10.

Tedeschi, B. 2002. E-commerce report; everybody talks about online privacy, but few do anything about it. *New York Times*, June 3.

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.

Varian, H. 1985. Price discrimination and social welfare. *American Economic Review* 75(4): 870–875.

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

Varian, H. R. 1996. Economic aspects of personal privacy. In *Privacy and Self-Regulation in the Information Age*. National Telecommunications and Information Administration.

Wathieu, L., and A. Friedman. 2005. An empirical approach to understanding privacy valuation. In *4th Workshop on the Economics of Information Security*.

Whitten, A., and J. D. Tygar. 1999. Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In *8th USENIX Security Symposium*.

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