Abstract

This chapter considers the growing body of theoretical and empirical research on the economics and behavioral economics of privacy, and how those streams of research can be applied to the investigation of the economic implications of business analytics and consumer data mining. The chapter begins by highlighting why privacy-sensitive scenarios and outcomes can be often interpreted through economic lenses. It then summarizes micro economic approaches to privacy. It continues by describing the current markets for privacy and personal data, consumers’ privacy valuations and behaviors, and the potential trade-offs associated with those behaviors. It concludes by discussing the role of privacy enhancing technologies and market forces in balancing the value of data and the value of privacy.
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 assume 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 or the cures for common ones. Now, consider the same world, but assume that employers become able to predict job candidates’ future health conditions from few data points extracted from the latter’s social network 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 the protection or disclosure of personal data - 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 scope of economic benefits and dangers alike.

Since the second half of the last century, progresses in information technology and the transformation of the advanced economies into service economies have made it possible for organizations to monitor, collect, store, and analyze increasing amounts of individual data. They 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

---

This chapter is partly based on previous work by the authors, including Acquisti (2010) and Brandimarte and Acquisti (2012).
assumption that all modern privacy issues have explicit monetary dimensions. Rather, this type of analysis stems from the realization that, whether or not consumers are aware of it, decisions that data subjects and data holders make about personal data often carry complex trade-off. The analysis of personal data can increase welfare, lower search costs, and reduce economic inefficiencies; at the same 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, 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. It starts with a brief summary of some relevant results from the micro economic theory of privacy (Section 2). It then describes the potential trade-offs associated 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 (see 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.

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, takes positions that may be considered alternative to Stigler and Posner. Hirshleifer 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

HIGHLY PRELIMINARY, INCOMPLETE
have a private incentive to invest more than socially optimal into collecting larger amount of consumer data).


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) study the economic impact of tracking technologies that make customer identification possible. If 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

HIGHLY PRELIMINARY, INCOMPLETE
offer personalized services that consumers will value).

This series of microeconomic results suggest 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. 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 three different types of markets.

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.

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 type 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 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 type of exchange involves 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.

A third type 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 they aim at giving consumers more “ownership” over (exchanges involving) their personal information, including - sometimes - the potential ability to monetize it.

The evolution and success of data mining and business analytics tools is and will keep affecting these various markets, especially in the form of both positive and negative externalities that arise when a consumers’ data is combined, or compared, or inferred, via the analysis of the data of many other consumers. As anticipated in Section 2, the emerging trade-offs 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 the following section we provide a few key examples that are especially relevant to the context of data mining and business analytics.
4 Privacy Trade-offs

Firms\(^2\) can significantly benefit from the ability to learn so much about their current, or potential, customers. Rich datasets of consumers can improve firms’ marketing capabilities, boosting their ability to address specific target markets or customers, and lowering their advertising costs (Blattberg and Deighton, 1991). Firms can therefore increase revenues through targeted offers, innovative coupon strategies such as Groupon.com (Pitta, 2010), and improved CRM (Richards and Jones, 2008), as well as increased consumer loyalty (consumers’ switching costs increase when a firm is able to use her information for personalized services; Ball et al. (2006)).

By analyzing large amounts of consumer data, firms are able to predict aggregate trends (such as variations in consumer demand) as well as individuals’ preferences (Linden et al., 2003), thus minimizing inventory risks and maximizing returns on marketing investment. They can improve their ability to offer useful recommendations to consumers (Bennett and Lanning, 2007), as well as their ability to enforce profit-enhancing price discrimination (Varian, 1985). Furthermore, by observing individual behavior, firms can learn how to improve their services, or re-design it in order to take advantage of the observed behavior.

An example of how consumer information can be leveraged for higher profit is online advertising. Such targetability implies that firms reduce the cost of ads wasted on consumers unlikely to be receptive to them. Furthermore, since online ad exposure, click-through behavior, and sometimes even post-exposure online behavior are often measurable, advertisers can monitor and improve the effectiveness of online advertising more than in other marketing channels. Primarily, this allows higher revenues for marketers and merchants (the price of behaviorally targeted advertising is almost 3 times as much the price of untargeted

---

\(^2\)IMPORTANT footnote for revision: this section is long, and - right now - is taken almost directly from a white paper I prepared for the OECD (Acquisti (2010)) Based on feedback, I will of course adapt (cut, change, add to) this section accordingly.

HIGHLY PRELIMINARY, INCOMPLETE
advertising: see Beales (2010)). Secondarily, this can also benefit the consumer: Targeted advertising may give consumers useful information, since the ads are tailored to consumers’ interests. Hence, such targeting may reduce the producers’ cost of communicating with consumers, and the consumers’ cost of obtaining useful information (Lenard and Rubin, 2009; Goldfarb and Tucker, 2010). In turn, revenues from targeted and untargeted advertising may support new services and business models, free content, or low-cost products - benefitting both consumers and firms.

Conversely, firms without access to consumer data may face significant barriers to entry and competitive disadvantage against firms with larger customer bases, thus limiting competition. Or, mandatory opt-in privacy policies for certain types of data may be costly for firms, when they result in the loss of valuable data (Staten and Cate, 2003). Furthermore, lack of consumer data may make it harder for firms to innovate and offer new services. For the same reason, uncertainty about (or fear of) possible legal reprisals following the collection or processing of consumers data may hinder product innovation.

Similarly, costs of undisclosed data may be suffered by society at large. During the summer of 2010, for instance, the Canadian Ministry of Industry announced that the long-form Census questionnaire would no longer be mandatory. The initiative was motivated by the Government’s stance that Canadians “should [not] be forced, under threat of fines, jail, or both, to divulge extensive private and personal information” (even though the Census data is actually never released to parties outside Statistics Canada in identifiable form). The transition from compulsory to voluntary, however, could likely result in a drastic decline of total respondents to the long-form questionnaire. The subsequent increase in the risk of non-response bias could, in turn, negatively affect the work of policy makers, researchers, or healthcare providers.³

Better marketing information in the hands of companies may also benefit customers and

society in general in an indirect way, by way of positive externalities. For example, better consumer data can allow firms to bring to the market niche products that, without focused data about potentially interested consumers, might have been too risky to develop (Blattberg and Deighton, 1991).

Prices might, sometimes, be reduced as an effect of more targeted (and less wasteful) advertising and marketing.

Online advertising - and in particular targeted ads - may both inform consumers (providing them better information at a lower search cost), as well as allow other services (for instance, news) to be provided for free to the consumers. Such ads may also be visually less intrusive than non-targeted ads (Goldfarb and Tucker, 2010).

The existence of a secondary market for customer data may also be a potential source of positive externalities for consumers. Such externalities may arise when, for instance, data provided to a website makes the service more convenient or efficient on another site, precisely because of sharing of data between different services (for instance, Facebook Connect enables seamless authentication on third-party Web sites, reducing the user’s cost of signing up across different platforms).\footnote{See http://developers.facebook.com/docs/guides/web.}

Furthermore, macro-economic benefits may materialize: the analysis and aggregation of the online behavior, sensor data, and individual decisions of a multitude of separate economic agents may allow the early identification of trends and patterns that would be otherwise hard or impossible to notice, or at least not within a limited period of time. This can benefit society as a whole: the monitoring and aggregation of web searches can allow the rapid identification of an infectious disease outbreak (Wilson and Brownstein, 2009); the combination of inputs from portable devices may be used for traffic and congestion control; data from remote and distributed sensors on consumers’ machines may be used for environmental monitoring (Dereszynski and Dietterich, 2007).

\footnote{See http://developers.facebook.com/docs/guides/web.}
As we discuss further below in this chapter, however, one can argue that such benefits may be enjoyed by consumers without their having to disclose personally identified data: the adoption of privacy enhancing technologies can make it possible to satisfy both the need for privacy and the need for sharing data, by selectively protecting and disclosing pieces of personal information.

On the other hand, both data holders and data subjects may suffer tangible and intangible costs from disclosed data. For the data holder, some of these costs may be associated with the mere collection of that data (for instance, when consumers deem a certain strategy of data gathering too intrusive). Other data holder costs are associated with the actual use (and misuse) of collected data (such as data breaches of improperly secure consumer data, which may lead to fines, settlement costs, or stock market losses for the firms involved (Romanosky and Acquisti, 2009). The costs the data subjects can incur because of disclosed and abused data are complex to categorize, since they comprise tangible and intangible damages that may occur (if at all) long after the data was initially disclosed.

As an example of the nuances of privacy costs, consider Calo (2011)’s distinction between subjective and objective privacy harms. Subjective harms derive from the unwanted perception of observation. They include “unwelcome mental states – anxiety, embarrassment, fear – that stem from the belief that one is being watched or monitored.” Objective harms consist of the unanticipated or coerced use of information concerning a person against that person, and include outcome as diverse as identity theft, the leaking of classified information that reveals an undercover agent, or “the use of a drunk-driving suspects blood as evidence against him.” As Calo notes, the categories represent, respectively, “the anticipation and consequence of a loss of control over personal information.” While no less important, subjective harms are harder to describe in economic terms than objective ones. The latter can often be described in terms of tort (Prosser, 1960). Instead, the former are not usually recognized by US courts as actual damage (Romanosky and Acquisti, 2009); furthermore,
they often amount to expected (that is, future and probabilistic, as opposed to presently incurred) costs.

An intuitive way of describing the state of uncertainty associated with privacy costs is the “blank check” metaphor. As an individual reveals private information to other parties, she is signing a blank cheque. The cheque may never come back to her, or may come back for an indeterminably small or large price to pay. That price could be a mild embarrassment, an annoying spam, or a devastating case of identity theft. In short, the probability, form, and actual damage from disclosed data are, in Knight (1921)’s terms, ambiguous and, up to a point, unknown.\(^5\)

Some of those costs are immediate but intangible: the psychological discomfort with feeling observed or violated; the embarrassment or social stigma when personal data has been disclosed; the chilling effect of the fear that one’s personal sphere will be, in fact, intruded.

Some costs are immediate and tangible: time and efforts spent deleting junk mail; annoyances from telemarketing; higher prices paid due to (adverse) price discrimination.

Some costs are more indirect: for instance, segmentation and profiling (especially in the form of behavioral targeting and advertising) may manipulate the consumer towards services she does not need or cannot afford.\(^6\)

Other costs are only probabilistic (that is, expected, rather than occurred, damages): for instance, errors in consumer databases due to poor data handling procedures by firms may later cause a consumer’s request to be wrongfully denied; or, breached databases may later results in identity theft Camp (2007).

\(^5\)Knight (1921) distinguished situations characterized by risk (in which the random outcomes of an event can be described with a known probability distribution) from situations characterized by ambiguity (in which those probabilities are unknown).

\(^6\)For instance, some marketing databases explicitly list personal information of individuals suffering from various types of addiction. For an example of a “gamblers database”, see http://www.dmnews.com/media-one-gamblers-database/article/164172/.

HIGHLY PRELIMINARY, INCOMPLETE
Because of their uncertain nature, privacy costs are therefore often hard to assess, and act upon, for the individual, but by no means less real: they often take the form of high-probability events with negligible individual impact (for instance, spam); or, they materialize as high significance events with very low expected probability of occurrence (for instance, being wrongfully denied a mortgage after suffering from identity theft). In either case, because of their low likelihood of occurrence or their limited impact, they may be dismissed as unimportant at the individual level - even if, in the aggregate, they may amount to significant societal damage.

The existence of a secondary market for customer data can also become a source of negative externalities for consumers. Such externalities may arise when the data holding company extracts the full benefit of using the information in its own marketing efforts, or the full price it receives when it sells the information to third parties, but does not internalize the losses that the consumer may derive from the disclosure of private information. Because customers often do not know the sources of data disclosure or the ways in which their data is combined and used, they may not be able to discipline effectively the companies that exploit that data (Swire and Litan, 1998). In economic terms, the company internalizes the gains from using the information (without necessarily sharing a portion of those gains to the consumer), but externalizes some of the losses.

Finally, a more intangible (but not less important) form of indirect consumers’ costs arises from the observation that, the more an individual’s data is shared with other parties, the more those parties gain a bargaining advantage in future transactions with that individual. Consider, for instance, behavioral targeting. While the consumer receives offers for products she is actually interested in, data holders accumulate data about her over time and across platforms and transaction. This data permits the creation of a detailed dossier of the consumers’ preferences and tastes, and the prediction of her future behavior. As the microeconomic models surveyed in Section 2 would predict, it is not hard to imagine that, in HIGHLY PRELIMINARY, INCOMPLETE
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. In other words, the disclosure of personal data ultimately affects the balance of power between the data subject and the data holder. The long-run effects of such altering of the balance of power are, of course, very hard to predict.

Conversely, some of the highlighted costs of disclosed data turn into benefits when consumer data is protected. For instance, when firms keep consumer data encrypted, they reduce the likelihood that, even if the data is breached, the consumer will suffer from identity theft. Similarly, consumers can benefit when certain personal information is not known by the merchant (such as information that the merchant may correlate with that individual’s willingness to bargain for a certain good: see Zettelmeyer et al. (2001)).

More notably, numerous of the benefits associated with data disclosure and highlighted above in this very section may, in fact, still be gained when data is protected.

For instance, Gellman (2002) points out that the privacy-enhancing restrictions in credit reporting brought about by the Fair Credit Reporting Act did not impose the impenetrable barrier to beneficial and profitable uses of consumer data that its critics feared before its passage. The welfare-diminishing effects of privacy regulation may have been similarly overestimated in other sectors, as markets find ways of adapting to new restrictions.

Similarly, in a recent paper, Goldfarb and Tucker (2010) find that while the enactment of privacy regulation limiting behavioral targeting did reduce the effectiveness of (and therefore welfare gains from) ads on websites with general content, it had no such impact on ads on sites with specific content, larger ads, or ads with interactive, video, or audio features.

Furthermore, while behavioral targeting reduces consumers’ costs of discovery of products that can match their preferences, so can less intrusive technologies: electronic marketplace in general, as well as search engines and sponsored searches in particular, reduce buyer’s search costs (Bakos, 1997) without linking consumer data across platform and transactions.
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.

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 who thought their information was less desirable and more deviant from the 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

HIGHLY PRELIMINARY, INCOMPLETE
a conjoint analysis, found that protection against errors, improper access, and secondary use of personal information was worth US$30.49-44.62 among US 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 US 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 provided evidence of how highly malleable
consumers privacy valuations can be: in 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.

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 (Taylor, 2004a).

There is, indeed, evidence that consumers face known decision making hurdles when facing privacy trade-offs: 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.

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 him through a website.

Furthermore, the “value” itself of the individual’s information might be highly uncertain

HIGHLY PRELIMINARY, INCOMPLETE
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, over-confidence, 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).

HIGHLY PRELIMINARY, INCOMPLETE
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.

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.

HIGHLY PRELIMINARY, INCOMPLETE
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?

HIGHLY PRELIMINARY, INCOMPLETE
Attempting to answer the above question may result in an interesting research agenda. A related 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 give market for data is not functioning properly, and when the benefits of new measures outweigh their costs.

HIGHLY PRELIMINARY, INCOMPLETE
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. Evaluations in this area, as we have discussed in this chapter, are far from stable.
References


HIGHLY PRELIMINARY, INCOMPLETE


HIGHLY PRELIMINARY, INCOMPLETE


Gellman, R. (2002, March). Privacy, consumers, and costs - how the lack of privacy costs consumers and why business studies of privacy costs are biased and incomplete.


HIGHLY PRELIMINARY, INCOMPLETE


HIGHLY PRELIMINARY, INCOMPLETE


