TECHNOLOGY POLICY INSTITUTE
Studying the Global Information Economy

IN DEFENSE OF DATA:
INFORMATION AND THE COSTS OF PRIVACY

May 2009

Thomas M. Lenard and Paul H. Rubin

IN DEFENSE OF DATA: INFORMATION AND THE COSTS OF PRIVACY

By Thomas M. Lenard and Paul H. Rubin*

EXECUTIVE SUMMARY

The commercial use of information on the Internet has produced substantial benefits for consumers.[1] But, as the use of information online has increased, so have concerns about privacy. This paper discusses how the use of individuals' information for commercial purposes affects consumers, and the implications of restricting information availability in the interest of privacy. We make the following points:

Targeted advertising gives consumers useful information. The online advertising industry uses customer information to target advertising messages to consumers' specific interests. Such targeting reduces the cost to producers of communicating with consumers and the cost to consumers of obtaining useful information. Internet advertising often introduces consumers to products they were unaware of and therefore unable to seek out on their own. If information about consumers becomes less available and more expensive, sellers rely more on sending messages to poorly targeted sets of consumers. As this occurs, consumers receive more irrelevant messages and find it more difficult to obtain useful information.

Advertising revenues support new services on the Internet. New business models based on advertising revenue support new services, often provided to consumers free of charge. The most prominent example is the search engine, which would likely not be available (or would not work as well) were it not for the ability of Google, Microsoft, Yahoo! and others to develop new sources of revenue based on targeted advertising. These companies use individuals' data to target advertising; improve their algorithms; protect against a variety of threats, such as search spam, click-fraud, and malware and phishing; and develop innovative new services. For example, Google has unveiled a new flu-tracking service that shows flu activity around the country based on searches for flu-related words, and can be useful to public health officials and perhaps consumers.

Information can be "reused," increasing its value. A key property of information is that once produced, it can be used multiple times at low cost. This "public good" characteristic of

---

* Thomas M. Lenard is president and senior fellow at the Technology Policy Institute. Paul H. Rubin is senior fellow at TPI and Dobbs Professor of Economics and Law at Emory University. The authors thank Arlene Holen and Scott Wallsten for helpful comments, and James Riso for very able research assistance.

[1] This study does not address categories of sensitive information, such as health information, personal financial information, or information about children. These types of information present separate issues and are subject to specific regulatory programs tailored for them (e.g., the Health Insurance Portability and Accountability Act of 1996 and the American Recovery and Reinvestment Act of 2009 for health information, The Gramm-Leach-Bliley Act of 1999 for financial records, and the Children's Online Privacy Protection Act of 1998 for children's information). We also do not cover government collection and use of information, which involves a different set of issues.

information is a major reason for its productivity. Some argue that information should be used only for the purpose for which it was collected, as called for in the European Directive on the Protection of Personal Data. Such a restriction on information use would preclude many productive uses, and actually lead to reduced security for consumers.

Information is used anonymously. The major categories of online advertising that rely on user behavior—search advertising, display ads, and email advertising—use that information anonymously. The process of targeting messages based on an understanding of users' interests, derived from information collected about their activities on the Internet, is entirely automated. Advertisers are not interested in individuals, but rather in blocks of people who are good targets for a specific product. This focus on aggregates shows up in pricing—ad prices are usually quoted in CPM: cost per thousand ad views or click-throughs.

Online information may facilitate differential pricing. Online information may make it easier for sellers to charge different prices to different consumers based on their willingness to pay. While the welfare effects of such differential pricing are ambiguous, it can improve welfare by making possible the production of goods that otherwise would not be produced. Information goods are prominent examples, because of their high-fixed, low-marginal cost structure.

Responsiveness to privacy concerns. The competitive online market structure suggests that firms do have incentives to satisfy their customers' privacy preferences and that consumers' behavior in the market reflects their preferences. Numerous privacy tools on the market enable individualization of privacy settings. Recent episodes involving AOL and Facebook, who were punished for violating privacy expectations of their customers, illustrate the costs to firms of deviating from acceptable practices.

Restricting legitimate information use is not likely to reduce identity theft. While people may be comfortable with intended uses of their data (by search engines, for example), they are worried about unintended uses, such as identity theft. Identity theft is perhaps the major specific harm alleged to result from the use of online information. However, restricting the use of information by legitimate firms is not likely to address the identity theft problem. One reason is that the Internet is involved in only 11 percent of identity theft cases, according to the most recent data. Moving transactions online reduces the risk of identity fraud. Moreover, use of information can reduce identity theft by making it easier for legitimate sellers to verify the identity of consumers.

Privacy advocates suggest privacy is a "free lunch." Privacy advocates argue that online practices violate individuals' rights and therefore should be curtailed. Innovations, such as the development of search engines or, more recently, the possibility that Internet Service Providers might use deep packet inspection as an online-advertising tool, have led to increased apprehension. However, more privacy implies less information available for producing benefits for consumers. Privacy advocates have provided little detail on the benefits of more privacy and have typically ignored the costs or tradeoffs associated with increasing privacy (i.e., reducing information). Their analysis suggests they believe that privacy is a "free lunch" consumers can obtain more of without giving up anything else.

Reducing online information use would be costly to consumers. Policy proposals that reduce the availability of information, such as an opt-in requirement or a Do Not Track list would be costly to consumers because they would receive fewer of the benefits that online information provides. The purpose of obtaining information about consumers is to provide them with targeted advertising and services, such as free search and email, for which consumers indicate they would willingly trade their information. Under an opt-in system much of this consumer value would be lost because opt-in rates would likely be quite low. Also, by increasing unwanted ads, a Do Not Track List would have the opposite effect of the popular Do Not Call List. A Do Not Track List would increase the volume of unwanted marketing messages.

In sum, good public policy requires that the benefits of more information be balanced against the benefits of greater privacy. Regulation should be undertaken only if a market is not functioning properly and if the benefits of new measures outweigh their costs. Our analysis suggests that proposals to restrict the amount of information available would not yield net benefits for consumers.

I. INTRODUCTION

You are an avid Washington Redskins fan and a frequenter of sports web sites. While catching up on national news on the *New York Times* web site, you see ads for golf clubs. This is probably not an accident. Based on your interest in football and the *New York Times*, perhaps combined with other characteristics, advertisers have determined that you are likely to be a golfer.[2]

This is just one example of a common though relatively recent phenomenon. But, while the techniques are new, the basic practice is not. Indeed, firms and consumers have collected information about each other since long before the advent of the Internet. Without knowing something about their customers, firms would have no idea what goods and services to produce, or where and how to sell them. Similarly, customers need information about firms and what they sell to know what is available, as well as to compare prices and other product attributes.

The information economy, as its name implies, has made these types of information more readily available. Firms can better satisfy their customers' needs, and consumers are able to find what they want more easily.

As the Internet has developed, firms have introduced innovative new business models that make better use of user information. These new business models have supported an array of new goods and services, often provided to consumers free of charge. For example, search engines place a wealth of information at individuals' fingertips that was unimaginable a few years ago. These services would likely not be available (or would be of greatly inferior quality)

---

[2] This is a hypothetical example, but it reflects a common phenomenon. See, for example, Jennifer Slegg. 2006. "What's the Buzz Behind Behavioral Advertising," http://searchenginewatch.com/3605361, May 11.

were it not for the ability of Google, Microsoft, Yahoo! and others to develop new revenue sources based on targeted advertising.

As the use of information online has increased, so have concerns about privacy. Privacy can be defined in many ways, but in the current context it can be viewed as the withholding of information—that is, where there is more privacy there is less information available for use in the marketplace. Thus, there is a tradeoff between the benefits of increased privacy and the benefits of increased information in the marketplace.

Quantifying the benefits and costs of privacy—or conversely, the use of personal information for commercial purposes—is difficult. Nevertheless, more information about these benefits and costs is needed before we are able to make an informed decision about whether new regulations that limit the use of information are in the interest of consumers.

The rest of this paper is organized as follows:

In Section II, we discuss the role of information in making markets better at delivering the goods and services people want. Information technology has reduced the cost of obtaining information, which has improved market efficiency and benefited consumers. An efficient market should also take into account consumers' preferences concerning privacy. The competitive online market structure suggests that firms do have incentives to satisfy their customers' privacy preferences and that consumers' behavior in the market reflects their preferences. We also discuss a key feature of information—that once produced, it can be used multiple times at low cost. This "public good" characteristic of information is a major reason for its productivity. This characteristic applies to many information-based goods, such as software programs and media, and also to information about consumers.

In Section III, we discuss the online advertising industry, which has used customer information to develop new business models that deliver advertising messages better targeted to consumers' interests. These models reduce the cost to producers of communicating with consumers and the cost to consumers of obtaining useful information.

In Section IV, we discuss search engines, the most prominent example of a new service built on advertising revenues. Search engines use data about individuals to target advertising, improve their algorithms, and protect against a variety of threats. Many other online services are provided free of charge because they are supported by advertising revenues.

In Section V, we discuss the fact that information used by online advertisers is anonymous. Advertisers are not interested in any individual, but rather in identifying groups of individuals who may be interested in their product.

In Section VI, we discuss how online information can affect the ability of sellers to charge consumers different prices depending on their demand for a product. While the welfare effects of differential pricing are indeterminate, the practice can improve welfare unambiguously when it allows for the production of goods that otherwise would not be produced. This is more likely to happen with information goods because of their high-fixed, low-marginal-cost structure.

In Section VII, we discuss identity theft, perhaps the major specific harm alleged to result from the spread of online information. The loss of personal data, such as a credit card number, puts an individual at financial risk. We discuss why restricting the collection and use of information by legitimate firms is not likely to remedy the identity theft problem, and why other regulatory measures such as notification requirements are not cost-effective. One reason is that the Internet is involved in only 11 percent of identity fraud cases, according to the latest data.

In Section VIII, we examine the arguments of privacy proponents, who generally favor limitations on the collection and use of information.

In Section IX, we discuss policy proposals, such as an opt-in requirement—i.e., requiring consumers to give affirmative consent before their data are used—or a Do Not Track list, that have been proposed to address consumers' privacy concerns. Our analysis indicates such measures would be costly to consumers, because they would receive fewer of the benefits that online information provides. Moreover, such measures would not necessarily reduce identity theft. We also discuss self-regulation.

Section X presents our conclusions.

II. SOME ECONOMICS OF INFORMATION

A. Consumer Preferences for Information and Privacy

The standard textbook economic model of perfect competition assumes perfect information. In reality, of course, information is never perfect because it is not costless. Individual firms and consumers make (often implicit) cost-benefit calculations, as they do with other economic goods, to determine how much information to obtain for decision-making. As the cost of information falls, market participants produce or gather more of it and are able to make better decisions. One of the major benefits of the Internet, and information technology in general, is that it has reduced the cost of obtaining all kinds of information and therefore has increased its availability.

Better information makes markets work more efficiently and do a better job of satisfying individuals' wants. For example, if merchants can better estimate demand, they are less likely to

purchase excess or insufficient inventories, thereby reducing costs to consumers. Geographic computer-based information can enable brick-and-mortar merchants to build new stores where they best serve consumers, and then stock those stores with the most locally relevant merchandise. Consumers searching for a particular product—say, a flat screen TV—can visit websites that provide detailed information about product characteristics and features. Moreover, Internet advertisers may use the fact that a consumer has searched for a TV to provide additional information about other features or brands through advertising associated with further searches. There are numerous such examples—all agents in the economy typically benefit from better information. It follows that policies that reduce the amount of available information in the interest of privacy impose a cost on consumers and the economy, and this cost should be compared with any gains from increased privacy.

Different industries use information differently. Credit agencies, insurance companies, and potential employers are interested in data about particular individuals. A person seeks to borrow money or buy insurance or get a job, and approaches a lender or an insurance company or employer. The lender or insurer or employer will want data about that individual before it decides to provide a loan or insurance, or to hire her.

By contrast, advertisers are not interested in data about individuals. The transaction unit in the advertising market is often a block of 1000 people who are good targets for a particular ad. This focus on aggregates shows up in pricing—ad prices are usually quoted in CPM: cost per thousand ad views. Part of the reason that consumers worry about online privacy may be confusion about these different uses. They are aware that some companies use information about individuals but not that others use only aggregated information.

B. Information as a Public Good

An important economic characteristic of information is its "publicness."[3] Once produced, information can be used multiple times by multiple parties without being "used up." Advertisers, credit institutions, and insurance companies may all use the same information.[4] Indeed, various information users cooperate in generating information because they all find it valuable. This public good characteristic of information is a major reason for its productivity.

Some argue that information should be used only for the purpose for which it was collected, as called for in the European Union Directive on the Protection of Personal Data. However, such a restriction on information use would preclude many productive uses, including uses associated with increased security.

The public good nature of information means that the externalities associated with the commercial use of information are more likely to be positive than negative. This in turn implies that it is more likely that not enough, rather than too much, information is available. This is why most advanced countries subsidize information production such as scientific research. Regulation that would reduce the availability and use of information would exacerbate the underproduction of information.

The public good nature of information has another implication: high fixed and low marginal costs. It is often expensive to gather or produce information but relatively inexpensive (or even free) to use it additional times once it is obtained. Such a cost structure creates difficulties for markets. Pricing at marginal cost means the good won't be produced because the

---

[3] Public goods are not diminishable and not excludable. Information is not diminishable, but is typically excludable—which is where information security and privacy come in.

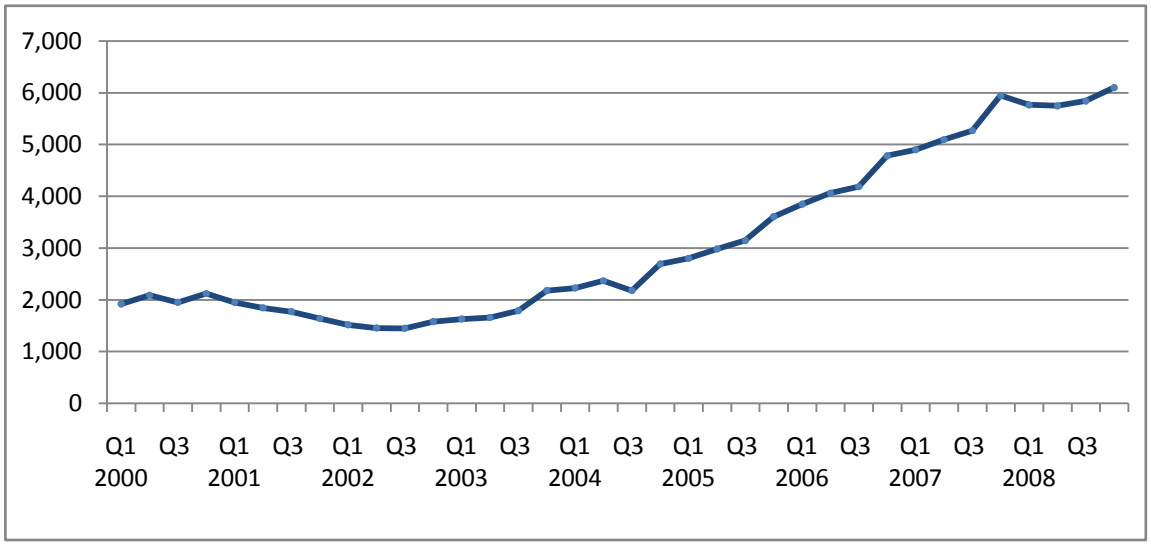
[4] The FTC has recently studied the use of credit data in automobile insurance markets and found that it is predictive of risk. See *Credit-Based Insurance Scores: Impacts on Consumers of Automobile Insurance: A Report to Congress* (July 2007).

price is not sufficient to cover its fixed costs. Pricing above marginal cost can produce deadweight losses due to reduced consumption by those who are priced out of the market. As we discuss below, differential pricing (price discrimination) is sometimes an answer to this problem.

III. ONLINE ADVERTISING

Online advertising spending is only about 9 percent of total media advertising spending, but it has been growing rapidly, as shown in Figure 1. Spending on online advertising was more than $23 billion in 2008.

Figure 1
Quarterly Internet Advertising Revenue 2000-2008, in millions USD



Source: PwC/IAB Internet Advertising Revenue Report, March 2009.

As shown in Table 1 below, search advertising is the largest component of online advertising spending, followed by display ads. To some extent all the types of online

advertising use individuals' information, such as browsing behavior, to improve marketing success.

Table 1
US Online Advertising Spending by Format 2008-2013, in millions USD

	2008	2009	2010	2011	2012	2013
Search	$10,691	$12,285	$13,880	$15,552	$17,686	$19,530
Display ads	$4,629	$4,933	$5,448	$6,182	$7,175	$7,958
Video	$587	$850	$1,250	$1,850	$3,000	$4,600
Rich media	$1,888	$2,030	$2,252	$2,560	$2,960	$3,660
Classifieds	$3,139	$2,956	$2,936	$2,944	$2,960	$2,982
Lead generation	$1,605	$1,645	$1,682	$1,792	$1,998	$2,268
Sponsorships	$590	$514	$542	$576	$629	$672
E-mail	$472	$488	$513	$544	$592	$630
Total	**$23,600**	**$25,700**	**$28,500**	**$32,000**	**$37,000**	**$42,000**

Source: eMarketer, November 2008.

A. The Value of Advertising

Advertising is a major source of information for consumers, as well as a major user of information to determine how to reach a target audience. Some privacy advocates view advertising as limited in value and even "manipulative."[5] However, the modern theory of advertising indicates that advertising is useful to consumers.[6]

Advertisers facilitate trade and, in general, there are gains from trade. Advertisers value information because it helps them match products with consumers who are good candidates for those products. This is why businesses are willing to pay for opt-in lists, which identify consumers interested in receiving certain types of messages.[7] Consumers value learning about

[5] See discussion in Section VIII.
[6] Summarized in Paul H. Rubin, "Regulation of Information and Advertising," *Competition Policy International*, Spring 2008, v. 4, No. 1.
[7] A Google search of "opt in list" provides many websites that advertise that they specialize in providing or building such lists.

products they are likely to buy. In order to receive useful messages consumers are willing to opt-in to these lists (or refrain from opting-out). When advertisers engage in targeted advertising they make it easier for consumers to get information on relevant products. Similarly, there is an active market for catalog mailing lists. Consumers who are interested in one sort of product may be interested in related products, so catalog vendors will often buy or trade lists in order to better target their catalogs.

If information can be used to target ads more inexpensively to blocks of individuals more likely to purchase the advertised product, then the information has benefited both the buyers and the sellers.[8] A well-known example of this practice is provided by Amazon. When a regular customer visits the website, she is greeted with a list of products based on past purchases and browsing. This feature has helped Amazon sell many useful books and other items consumers otherwise might not have found, to the benefit of both parties. And even if the information does not lead directly to a sale it may still be valuable to the consumer. For example, it may help consumers compare prices or determine what products are available.

HDTVs are complex, and searching various websites can enable one to learn about types of screens (LCD versus plasma), picture quality (1080p versus 720p) and other more esoteric characteristics so that a final purchase decision will be based on a great deal of information. Moreover, once a consumer has searched for information about a TV online, it becomes more likely that advertisers will provide her with additional information, perhaps about previously unknown characteristics or products such as cables that are complements to the TV itself. This is because many advertisers engage in "behavioral marketing," whereby a past search or other

[8] For an important discussion of the value of information from an economic perspective see Shapiro and Varian (1999).

expression of interest increases the probability of seeing advertising messages related to that search.

B. Targeted Advertising

As advertising becomes more targeted, recipients avoid the nuisance and inefficiency of being confronted with ads that are of no interest—in the marketing jargon, "consumer-borne marketing costs" (Petty 2000, pp. 42-53). Both advertisers and consumers gain from advertisers having access to information to reduce these nuisance costs. As a result, consumers receive fewer irrelevant messages, and advertisers deliver more messages that are actually read (Petty 2000).

For example, a mortgage lender trying to reach potential borrowers through email messages or by placing ads on websites must decide how widely to spread the message. If she chooses wide dissemination, many consumers will see the message, but most will be indifferent or even annoyed by it. If she can target her audience more narrowly—for example, by advertising to individuals who have been looking at online real estate ads—fewer consumers will see the message, but it will be relevant to more of them and a higher percentage will respond.

If the marginal cost of spreading a message to more recipients is zero, it might appear that the seller should spread her message as widely as possible, perhaps through the use of "spam," because there is some increased probability, even if small, of additional sales. However, there is a cost to spreading the message too widely in that consumers may in the future disregard this seller's messages, or even program their email clients' spam filters to delete such messages. If in the future the seller has some product that is of interest to these buyers, she will have more difficulty contacting them. In other words, there is a reputation cost to a seller from screening

the audience for a message insufficiently. The cost is higher for more "reputable" sellers—for sellers who invest more in reputation and plan to continue selling.

The value of a seller's reputation provides an incentive to spend resources to identify the best recipients for her message. It may be simple geographic information: if the seller is local, zip codes of potential customers are useful. It may be product-specific "contextual" information, as when an advertiser puts banner or text ads for automobiles on automobile-related websites. It may be sending mortgage ads to people who are in the market for a new house, as suggested above. Or it may be more complex information based on interrelations between various sites visited by individuals. For example, it is claimed that people interested in romantic movies are more likely to rent cars for the weekend.[9] Advertising based on past search and browsing behavior is called "behavioral" advertising.

Such information is valuable to a seller because it enables her to target consumers and avoid diminishing the value of her reputation by communicating with uninterested consumers. But like any other business investment, the amount of information that a seller uses depends on its price. As information becomes more costly, the seller uses less and accepts a greater diminution in reputation. As information is easier to obtain, a seller uses more and targets ads more carefully.

Additionally, much of the information that sellers use for targeting consumers is for statistical purposes. For example, information—such as the above hypothetical that viewers of romantic movies are more likely than average to rent cars for the weekend—is more useful if collected from a large number of consumers. Information about a single consumer is only one data point for performing the statistical calculation. If sellers have information about fewer

[9] Mentioned in Stephen Baker's *The Numerati* (New York: Houghton Mifflin, 2008).

consumers, then it will be more difficult to measure such relationships with precision, and therefore advertising will be less targeted, with the associated social costs.

Such information also has a public good component. Recipients of irrelevant messages do not only hold the particular seller accountable; there is a spillover to other sellers (Loeser 2000). Every time an irrelevant message is received the expected value of future messages from all senders is reduced, because consumers are more likely to simply ignore all messages. Thus, the lack of information for targeting consumers creates a public harm (or, alternatively, the ability of sellers to use such information creates a public good). One email service provider has made this externality argument explicit. In discussing the optimal rate for email, this firm recommends no more than one message per week; the firm believes that too many messages from any sender can be harmful and that "one of our biggest challenges" is that "we can't control what other people are doing out there."[10] The issue of excess messages leading consumers to ignore all messages is called "marketing clutter" (Petty 2000).

The same issue arises for search engines. If a search engine shows many irrelevant or annoying ads, users may well decide to ignore ads ("ad blindness"), stop clicking on ads, or even switch search engines. This gives the search engines a strong incentive to manage the relevance of ads that users are shown.

The "public" nature of the harm, if information about consumers becomes less available and more expensive, is this: sellers use less information and rely more on sending messages to less well-targeted sets of consumers. As this occurs, consumers receive more irrelevant messages and pay less attention to messages in general. The cost to producers of communicating

---

[10] Thomas E. Weber, "Why Companies Are So Eager To Get Your E-Mail Address." *The Wall Street Journal*, February 12, 2001.

with consumers is thereby increased as is the cost to consumers of obtaining valuable information, resulting in a loss of consumer welfare.

C. Search Costs and New Products

The economics literature characterizes search as the effort by consumers to find information about prices or other characteristics of products. Under conventional search theory, discussed in the literature going back to Stigler (1961), the consumer knows (at least approximately) what she wants, and searches for the best terms (e.g., price, location) for this product. Search may also occur for products with specific characteristics. The benefits from reduced search costs include the direct cost savings as well as improved matching of consumers and products.

Advertising through the Internet and email (as well as other media) goes well beyond this, providing what might be characterized as "push" information, where the seller takes an active role in providing information to consumers about products of which they may be unaware. New technologies and products—both Internet and non-Internet related—are constantly being introduced. Consumers are not in a position to seek out such products unless they know of their existence, and Internet advertising is an important source of such information.

As an example, consider Amazon's recently updated electronic book reader, Kindle. Kindle is almost entirely promoted through emails and promotions on the website to existing Amazon customers. A consumer would not seek out an electronic book service unless she already knows one exists. Various forms of push advertising by Amazon are an efficient way of informing consumers of new products or promotions. As a more general example, some websites have "wish lists" and inform consumers of changes in prices of selected products. Such

wish lists greatly reduce search costs for consumers. Consumers can also join various discussion groups that provide information about new products or accessories to existing products.

IV. SEARCH ENGINES AND OTHER FREE SERVICES

A. Search Engines

Search engines provided by Google, Yahoo!, Microsoft, and others—fueled by $10 billion in search advertising revenues—place a wealth of information at our fingertips that would have been hard to envision just a few years ago. Current web search technology relies on a range of information about individual users' interactions with the search engine. Although search engines *could* operate in today's Web without access to any user information, collection of information such as IP addresses, search queries, and result click-through history greatly improves their functioning, and is necessary for technological innovation. Search engines use data collected from their users to optimize search results, target advertisements, and protect against a variety of threats.

For example, a search firm can observe if searchers click on the first link in the results (a successful search) or if they must go further into the results (a less successful search). This information can then be used to refine search algorithms. Moreover, when a search algorithm is changed, it is important to determine if the change is beneficial. In order to make this determination, it is necessary to compare the success of current searches with past searches, based on an examination of search history.[11] Past history can also be used for seemingly simple but important tasks, such as the spelling query that often precedes search results ("Did you

[11] See Hal Varian, "Why Data Matters," March 24, 2008, http://googleblog.blogspot.com/2008/03/why-data-matters.html

mean…?"). These queries are based on the extent to which searchers responded favorably to similar questions in the past, and this history is used to refine future suggestions.

The practice of IP logging allows Google and other search engines to record search histories for unique IP addresses. Almost all websites keep some internal user log. A persistent tracking cookie placed on each user's hard drive contains both user preferences—things like SafeSearch settings, number of results to display per page, and language settings—along with a unique alphanumeric identifier. When combined with logged IP data, this pseudonymous identifier gives Google the ability to compile search histories for individual users across several different IP addresses—for example, if a user's home Internet connection assigns dynamic IP addresses, or if a user visits Google from several different wireless connections. Google stores log data for nine months (recently reduced from 18 months), during which software engineers extensively mine and analyze it. Yahoo! recently announced it would anonymize its log data after 90 days. MSN saves data for 18 months before redacting IP information. Data on user searches and clicks are used to improve the performance of the search and ad algorithms, just as data on consumer shopping behavior improves the layout and display of products in a physical store.

Search history is only one component in the search engines' complex proprietary algorithms. Search providers tightly control information on search engine algorithms, but firms offering "search engine optimization" services have deduced the basic principles of most search systems, both by examining patent information and through trial and error. Google, for example, still bases searches on the PageRank link analysis algorithm, which has been refined over time to account for approximately 200 other weighting factors to prevent enterprising website operators from manipulating page rankings. Search engines constantly update their algorithms to address

new manipulation techniques and increasingly rely on user information to provide customized search results.

Search logs are also used to protect against a variety of threats: search spam (the efforts to insert spurious results into the search stream),[12] click-fraud (fraudulently clicking on advertisements in order to drive up the costs to advertisers),[13] and other threats such as malware and phishing.[14] When search engines design methods of protecting against these threats, they can test their tools using past searches to determine if they are effective and if they eliminate too many useful results. Without sufficient policing these threats could virtually shut down the search process, or in the case of click-fraud, seriously undermine the business model on which search is based. These are also examples of the public good characteristic of information since search log information can be used for multiple purposes once collected.

Different applications run by search services—toolbars, email accounts, desktop search, personalized homepages—may collect different and more extensive information that can be combined with personal data collected by search engines themselves. Typically, separate privacy policies cover collection and use of information by these different services.

At the highest level of data collection, users can opt in for Google's "Web History" service for users with Google accounts (those who use Gmail, Google Documents, Google Reader, or most other Google services), which tracks, indexes, and archives search history. Data are available both to the user for future reference and to Google itself for analysis. Yahoo! offers

---

[12] See, for example, Matt Cutts, "Using data to fight webspam," June 27, 2008, http://googleblog.blogspot.com/2008/06/using-data-to-fight-webspam.html
[13] See Shuman Ghosemajumder, "Using data to help prevent fraud," March 18, 2008, http://googleblog.blogspot.com/2008/03/using-data-to-help-prevent-fraud.html
[14] See Niels Provos, "Using log data to help keep you safe," March 13, 2008, http://googleblog.blogspot.com/2008/03/using-log-data-to-help-keep-you-safe.html

a similar service, called “Personal Search,” and uses data collected by these applications for behavioral targeting in advertising.

Users concerned about ensuring personal privacy when using search engines have a variety of technological tools at their disposal to choose a level of activity-monitoring with which they are comfortable. Search engines provide some of these themselves; for example, users can opt out of Google’s Web History (which is opt-in to begin with), pause monitoring, or delete their collected search and browsing history altogether. Yahoo!, AOL, and MSN also allow users to opt out of similar behavioral targeting systems.

Other privacy protections rely on client-side techniques. Users of Internet Explorer and Firefox can easily delete Google's tracking cookie, which is essential for tying together separate personal data streams. In addition, several free browser extensions and utilities can clear the cookie or require Google to provide a new one at the start of each browsing session. Web proxies and anonymizing applications like Tor easily conceal user IP addresses, although because of their architecture they often reduce bandwidth speeds. At the most basic level, a dedicated user could potentially even “spoof” TCP source addresses to prevent Google from monitoring immediately previous search results. The costs in time and difficulty of these solutions tend to increase as the desired level of privacy increases, but minor actions can have huge marginal effects on privacy protection—for example, opting out of Web History takes only a few clicks but prevents collection of a significant amount of personal data, whereas browsing entirely anonymously requires more effort to set up.

Finally, all major search engines offer privacy policies in compliance with the requirements of both United States and European data security laws. These policies disclose the

companies' use of personal information and require user notification and consent before transferring personal information to others.

Users of search engines essentially face a tradeoff between protecting their personal privacy and the speed and relevance of their search results. As personalized search algorithms and behavioral targeting techniques grow in popularity and precision, this tradeoff will likely become more and more evident to the everyday user of search technology.

Considering the current availability and ease of use of tools for protecting personal information, the greatest threat to individual privacy is not search engines themselves, but the governments that may seek their records. Companies vary in the degree to which they have protected data from government requests. In 2006, Google resisted a Justice Department subpoena for millions of user search records, while Yahoo!, AOL, and MSN complied and handed over detailed server logs. It is unclear whether search engines have aided government agencies in other surveillance efforts. Google declared publicly in March 2008 that it "was not part of the NSA's Terrorist Surveillance program," but this statement does not rule out the possibility of collaboration in other monitoring schemes.

Could search engines exist and organize information without collecting any personal information? They could—consider Google's early years, when search rankings were based primarily on analysis of incoming links. However, Google attributes much of its success in developing better search algorithms to careful analysis of consumer behavior that is stored in its logs. Going forward, it is likely that user information will continue to be useful in providing searchers with relevant results and sustaining the business model that makes free search engines viable.

B. Other Free Services

The $24 billion Internet advertising industry supports many other services provided to consumers at no charge. They include customized pages from firms like Yahoo! with information of direct interest to the particular individual, and also free email services from many providers.[15] The major Internet advertising firms cooperate with operators of lower volume websites to provide customized advertising, and the revenues from this advertising enable many firms to remain in business and support their websites. Moreover, some bloggers earn sufficient revenues from advertising that they can devote more time to blogging. Table 2 lists some of the services available.

Table 2
Online Advertising-Supported Content

Video (e.g., Hulu)	Phone	Weather
TV	Internet portals	Donation sites (freerice.com)
Webmail	Maps	Translation services
Newspapers	Social networks	Online dictionaries
Games	Dating websites	Local event calendars
Educational services	Travel planning	Music (Pandora)
Blogs	Search	Job boards
Product rating and pricing services (e.g., CNET)	Information websites (About.com, etc.)	Classified ads (e.g., Craig's list)

[15] One of the authors uses a free customized page from Yahoo! as his homepage. This contains information in many categories that he has selected: headlines on selected topics from Reuters and AP, information about chosen stocks and stock indices; weather in selected cities, and movies in his neighborhood. Many other categories are also available.

C. Location Data: New Technologies

Firms are actively developing new technologies based on available data. For example, a new technology enables firms with access to cell phone data to track subscribers' movements.[16] Businesses are beginning to use these data to track customers and other patterns of behavior. Because of privacy concerns, the data are used anonymously but may still be useful for a range of purposes including emergency response, epidemic prevention, and urban planning. This type of information—location data from cell phone use—was not previously available, but may become extremely valuable in the future. The researchers and firms involved appear to be sensitive to privacy concerns.

Google has unveiled a new flu-tracking service that shows flu activity around the country based on searches for flu-related words. This service can be useful to public health officials and perhaps consumers.[17] Recently, Google released an experimental version of this service tailored to Mexico, in hopes of helping track swine flu trends in that country.[18]

---

[16] The technology is described in Marta C. Gonzalez, Cesar A. Hidalgo and Albert-Laszlo Barbasi, "Understanding individual human mobility patterns," Nature, V. 453, no. 5, June, 2008, pp. 779-782. Popular discussions are in Robert Lee Hotz, "Cellphone Data Track Our Migration Patterns," *Wall Street Journal*, June 10, 2008; Page A12, and Michael Fitzgerald, "Predicting Where You'll Go and What You'll Like*," New York Times*, June 22, 2008, p. Business 4.

[17] Google Flu Trends, which uses search engine queries, and is not based on cell phone data. Service available at http://www.google.org/flutrends. See discussion in the November 12, 2008 *Wall Street Journal*, http://online.wsj.com/article/SB122644309498518615.html.

[18] http://www.nytimes.com/2009/05/01/technology/internet/01google.html?_r=1&partner=rss&emc=rss&pagewanted=print

V. THE ANONYMOUS USE OF INFORMATION

A. Use of Information by Advertisers

The major categories of online advertising that rely on user behavior—search advertising, display ads, and email advertising—use that information anonymously. The process of targeting messages based on an understanding of users' interests, derived from information collected about their activities on the Internet, is entirely automated. No human is directly involved.

Advertisers are interested in locating consumers who have an interest in their product, but they have no interest in the identity or behavior of any individual. A seller does not ask "what can I sell to Paul Rubin?" Rather, a seller may ask an ad server such as DoubleClick or 24/7 to "put my ad on 1,000,000 pages viewed on computers of persons more likely than average to want a new car." Perhaps Paul Rubin's computer turns out to be one of those selected. But no person makes this determination; instead, it is made by various computers connecting with each other. Moreover, the unit of commerce in the online advertising market is typically 1000 impressions or click-throughs, not any individual.

Search engine advertising typically consists of text-based ads that users see alongside the responses to their search queries. The selection of ads that appears is determined by an auction process where advertisers indicate the price per click they are willing to pay, together with sophisticated models that predict click-through rates, since that is how search engines generate revenue. Search engines don't get paid unless they deliver ads in which consumers are interested. The models they use, as well as the selection of the ads that appear in response to any particular search query, may utilize user browsing behavior. However, all this is done by computers.

Similarly, email advertising may be based on scanning for key words in the email message. Again, this is automated and does not involve human interaction with individuals' data. Aside from advertising, email services routinely scan messages for a variety of reasons, such as virus detection.

Advertisers deliver targeted display advertisements across multiple websites using information they glean from a web surfer's activity: from the immediate query that a user makes (contextual information) and from storing and aggregating queries over time (behavioral information). Interpreting the immediate query involves using the IP address information attached to the query as a proxy for physical location in order to make some assumptions about the user's demographic profile. This also involves knowing in advance something about the type of pages the person is requesting and making assumptions about the individual's interests based on the content shown in the page.

Aggregation of queries over time allows the web surfer to be identified as a unique but anonymous individual. To do that advertisers make use of cookie technology that stores a unique identifier on the web surfer's computer. The aggregation of queries, each keyed to a unique identifier, provides the data needed for statistical models that categorize users into demographic or interest profiles. These profiles are used when matching a particular type of user with a relevant advertisement.

This cookie-based matching is imperfect. It is IP addresses that are matched, not individuals. Thus, if several people use the same computer, the "individual" that is recognized will be a composite. If one individual uses more than one computer (or even more than one

browser on the same computer) then her records will be fragmented. Thus, the advertising profiles used may be quite noisy.[19]

In thinking about privacy on the Internet, the metaphor sometimes used is of someone observing a consumer, learning about her personal characteristics, and trying to sell her something. A leading example of this metaphor is from a well-known (albeit somewhat dated) article by Jerry Kang from the 1998 *Stanford Law Review*:[20]

> By contrast, in cyberspace, the exception becomes the norm: Every interaction is like the credit card purchase. The best way to grasp this point is to take seriously, if only for a moment, the metaphor that cyberspace is an actual place, a computer-constructed world, a virtual reality. In this alternate universe, you are invisibly stamped with a bar code as soon as you venture outside your home. There are entities called "road providers," who supply the streets and ground you walk on, who track precisely where, when, and how fast you traverse the lands, in order to charge you for your wear on the infrastructure. As soon as you enter the cyber-mall's domain, the mall begins to track you through invisible scanners focused on your bar code. It automatically records which stores you visit, which windows you browse, in which order, and for how long. The specific stores collect even more detailed data when you enter their domain. For example, the cyber-bookstore notes which magazines you skimmed, recording which pages you have seen and for how long, and notes the pattern, if any, of your browsing. It notes that you picked up briefly a health magazine featuring an article on St. John's Wort, read for seven minutes a newsweekly detailing a politician's sex scandal, and flipped ever-so-quickly through a tabloid claiming that Elvis lives. Of course, whenever any item is actually purchased, the store, as well as the credit, debit, or virtual cash company that provides payment through cyberspace, takes careful notes of what you bought—in this case, a silk scarf, red, expensive.

This is not the way it works. Businesses and advertisers do not find a particular consumer and attempt to figure out what that consumer wants to buy. Rather, the process is the opposite: a

---

[19] Emily Steel, "Mistaken Identity," *Wall Street Journal*, September 20-21, 2008, p. W5.

[20] pp. 1198-99. As of December 26, 2000, Westlaw indicated that this article had been cited 100 times—a large number for a relatively new article. Moreover, it has been cited in many or most of the law review literature on privacy in cyberspace.

seller has goods for sale, and tries to find consumers who are more likely than average to buy those goods.

Some commentators—e.g., Daniel Solove (2001)—recognize this point:

> Since marketers are interested in aggregate data, they do not care about snooping into particular people's lives. Much personal information is amassed and processed by computers; we are not being watched by other humans, but by machines, which gather information, compute profiles, and generate lists for mailing, emailing, or calling. This impersonality makes surveillance less invasive.
>
> While having one's actions monitored by computers does not involve immediate perception by a human consciousness, it still exposes people to the possibility of future review and disclosure. In the context of databases, however, this possibility is remote. Even when such data is used for marketing, marketers merely want to make a profit, not uproot a life or soil a reputation.[21]

B. Deep Packet Inspection

A relatively new technology, "Deep Packet Inspection," permits ISPs to gather information by examining the contents of information—"packets"—that cross the network. This may enable sellers to obtain more specific data on consumer preferences and desires, and to better engage in behavioral marketing. A firm called NebuAd is a leader in this technology. Privacy proponents are opposed to using this technology to gather marketing-related information,[22] and Congress has held hearings on the issue. As a result, several ISPs have postponed or cancelled plans to utilize this method.[23] However, deep packet inspection has the

---

[21] This paper has been the subject of a *New York Times* story: Carl S. Kaplan, "Kafkaesque? Big Brother? Finding the Right Literary Metaphor for New Privacy," February 2, 2001. Solove is still in favor of regulation, however, and he perceives other dangers from online information, although the harm is "difficult to describe" and "difficult to quantify" (p. 40).

[22] See, for example, EPIC's deep packet inspection page, http://epic.org/privacy/dpi.

[23] Emily Steel and Vishesh Kumar, "Targeted Ads Raise Privacy Concerns," *Wall Street Journal*, July 8, 2008, p. B1.

potential to improve information flows in the economy. The information is anonymous, as we have noted is true of most online information. Indeed, NebuAd uses only a subset of information: "The NebuAd advertising service does not collect or use any information from password protected sites (e.g., HTTPS traffic), web mail, email, instant messages, or VOIP traffic. Using only non-PII [non-personally identifiable information], NebuAd constructs and continuously updates these unique and anonymous user profiles. In the course of these business operations, NebuAd's ad optimization and serving system does not collect PII or use information deemed to be sensitive (e.g., information involving a user's financial, sensitive health, or medical matters)."[24]

VI. DIFFERENTIAL PRICING AND THE AVAILABILITY OF GOODS

Online information can affect the ability of sellers to engage in differential pricing, or price discrimination, where different prices are charged to different consumers based on their willingness to pay (Shapiro and Varian 1999). This strategy enables firms to increase profits relative to charging a single price.

There are constraints to price discrimination. First, the firm must have some market power. In competitive industries, all units of a good sell for the same price, and no firm can charge more than the market price to any consumer. Second, firms must be able to segment the market between consumers with different demand characteristics. Third, the firm must be able to prevent low-price consumers from reselling the product to high-price consumers.

[24] Testimony of Robert R. Dykes, Chairman and CEO NebuAd, Inc.before the Subcommittee on Telecommunications and the Internet, Thursday, July 17, 2008, available at http://energycommerce.house.gov/cmte_mtgs/110-ti-hrg.071708.DeepPacket.shtml

By making pricing more transparent—for example, when a website provides a price that anyone can see—the Internet can make price discrimination more difficult. Also, consumers can use the Internet to more easily communicate among themselves and thus learn about ways to obtain lower prices.[25]

On the other hand, online information can make price discrimination easier, because sellers can gather information about consumers' preferences and use this information to charge discriminatory prices. Moreover, if purchases are made online, it may be more difficult for consumers to learn what other consumers are buying or see the prices they are paying. This facilitates price discrimination. We speculate that overall the availability of online data makes discrimination easier, not more difficult.[26]

Economists distinguish three basic forms or price discrimination. In Type I discrimination (perfect discrimination) each consumer pays his or her reservation price for each unit of the good purchased. This type of discrimination is efficient, in that there is no deadweight loss in the market, although what was consumer surplus is now entirely captured by sellers. In second-degree discrimination, different units sell for different prices, but the same price schedule is available to all; an example is a volume discount. Personal information is less useful for this type of discrimination since all consumers face the same schedule. In third-degree price discrimination, different categories of consumers pay different prices. Discounts to students or senior citizens are an example.

---

[25] For one example, the website TreoCentral is a chat site for consumers interested in the Treo line of mobile phones. One perennial topic is pricing of both phones and plans. Information about discriminatory low prices is available, and some consumers can learn about ways to get better deals.

[26] Most Treo owners do not monitor TreoCentral.

In general the welfare effects of price discrimination are indeterminate.[27] Some consumers lose and some gain, and there is no general way of determining the net effect. However, whenever price discrimination leads to increased sales of a product it is efficient (Varian 1996). Price discrimination is particularly useful if it allows additional markets to be served. A common example is "niche" markets—small markets that would not be served absent discrimination.

Price discrimination is clearly welfare-enhancing when the good could not otherwise be produced. This is especially likely for goods with high fixed costs and low marginal costs, because pricing at marginal cost (the competitive outcome) doesn't cover the fixed costs of producing the good. While many classes of goods have this cost pattern, informational goods are perhaps the most prominent category, as we have discussed. For example, it costs Microsoft billions of dollars to produce a new version of Windows, but only a few cents to duplicate the program once it is produced. In markets with high fixed costs, efficiency requires that marginal willingness to pay should equal marginal cost. This is also the profit maximizing condition for a firm able to engage in price discrimination.

Therefore, to the extent that online discrimination deals with informational goods such as software it can be efficient. With a high-fixed, low-marginal cost structure, price discrimination can make it profitable to develop new products that otherwise would not be produced. That is, price discrimination can lead high-demand buyers to pay a higher price, which helps cover the high fixed costs of producing the good. Then low-demand consumers

[27] For a discussion in the context of information see Kai-Lung Hui and I.P.L. Ping, "The Economics of Privacy," available at SSRN.com.

may pay something closer to marginal cost. Welfare will be higher relative to situations where the good is not produced at all.

An interesting form of price discrimination occurs when firms can obtain information about repeat customers. In this case, firms might want to discriminate but consumers may be able to foil this effort (by, for example, anonymizing their purchases). To combat this behavior, some firms offer services to repeat customers in order to make it worthwhile for consumers to identify themselves. For example, Amazon may not have the best price on a product at any given time, but it has invested heavily in its reputation for providing high quality service, including suggestions of additional purchases. Owing to these efforts it may still be the preferred choice (Acquisti and Varian 2005).

VII. COSTS OF INFORMATION: IDENTITY THEFT

People may be comfortable with intended uses of their data by search engines or advertisers, but they want their data to be secure. Identity theft—which involves the loss of personal data that poses a financial threat (such as a credit card number)—is perhaps the major privacy concern of individuals.[28] But the relationship between identity theft and online privacy is tenuous. Regulating the collection and use of information by legitimate firms does not appear to make it more difficult for criminals to access information such as credit card numbers, and therefore does little or nothing to deter identity theft. In fact, excessive control of information may increase the risk of identity theft by making it more difficult for sellers to

[28] Hal Varian, Google's chief economist, has observed that the most common privacy concern expressed by consumers in a focus group is that "someone might steal my credit card number."

determine if a potential buyer is fraudulent or not. Moreover, anything that encourages individuals to shift transactions offline is likely to be counterproductive.

A. The Incidence of Identity Theft

The FBI defines identity theft as "the illegal use of another person's identifying information (such as a name, birth date, social security and/or credit card number)," and calls it "one of the fastest growing crimes in the United States." However, the FBI's official data contradict this assertion. The Bureau reports that identity theft represents only 2.9 percent of all crimes reported to the Internet Crime Complaint Center and that one category of crime, non-delivery of merchandise, increased by more than the whole of identity theft (see Table 3).[29] Thus, even if there were no identity theft in 2006, it could not have increased by as much as non-delivery increased.[30] The FBI's characterization of identity theft as one of the fastest growing crimes is part of the general tendency to overstate the significance of identity theft.

Contrary to most popular perceptions, identity theft, while serious, is not a growing problem. A series of surveys by Javelin Strategy and Research between 2004 and 2007 found that both the total number of victims and overall monetary losses decreased every year during that period, presumably due to better risk management on the part of the industry (Javelin 2008). The most recent Javelin survey showed an increase in identity fraud victims in 2008 bringing it back approximately to the 2004 level, although the average cost per victim declined by 31 percent (Javelin 2009). These trends are shown in Figure 2.

---

[29] National White Collar Crime Center, Bureau of Justice Assistance, Federal Bureau of Investigation, *2007 Internet Crime Report*.

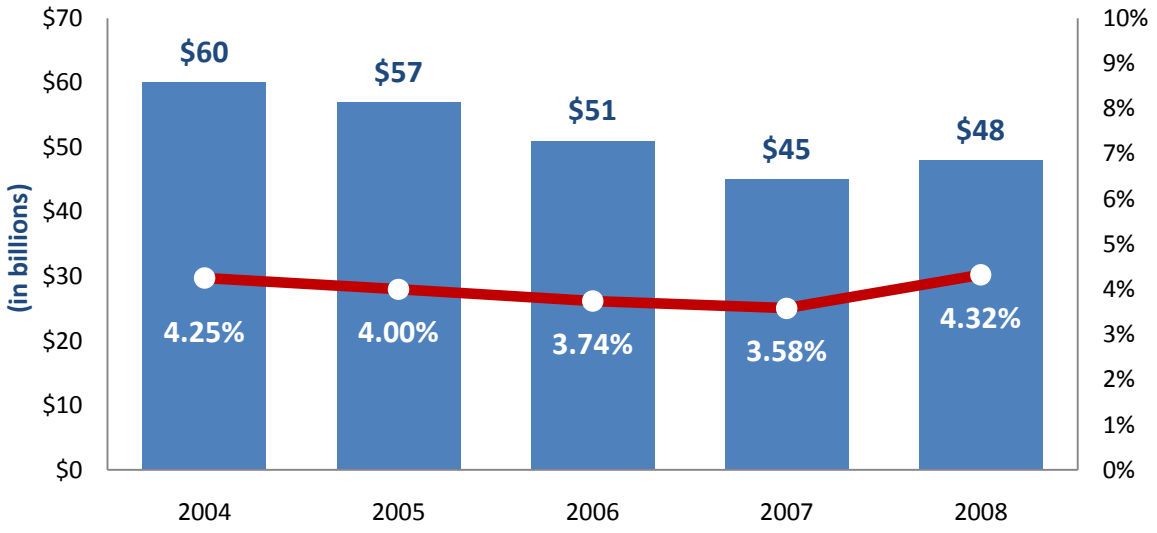
[30] The Report may mean that the percentage increase in identity theft was fastest growing, but the base is so low that even a small absolute increase would be a large percentage increase.

Table 3
2007 Top Ten Complaint Categories

Complaint Category	*% of Total Complaints Received*
Auction Fraud	35.7
Non-delivery	24.9
Confidence Fraud	6.7
Credit/Debit Card Fraud	6.3
Check Fraud	6
Computer Fraud	5.3
Identity Theft	2.9
Financial Institutions Fraud	2.7
Threat	1.6
Nigerian Letter Fraud	1.1

Source: Internet Crime Complaint Center. *2007 Internet Crime Report.*

Figure 2
Overall Fraud Amounts and Incidence Rates 2004-2008



Source: Javelin Strategy press release, February 2009.

Perhaps most importantly, the Javelin survey found that "In 2008, online access, such as using virus-afflicted computers at home or at work, accounted for only 11 percent of the total fraud. Combined with the increased speed of misuse, this trend points to more attacks of opportunity, when a fraudster takes advantage of personal information to which they suddenly have access, such as a lost wallet or watching someone enter their ATM PIN" (Javelin 2009). Earlier estimates by The Nilson Report (2005) showed the total costs of credit card fraud to issuers decreased from $882 million in 2003 to $788 million in 2004—a 10-percent decline. Moreover, over a longer period—1992 to 2004—The Nilson Report found that the costs of these frauds decreased from $0.157 to $0.047 per $100 in credit card sales.[31] This is not surprising, because credit card firms are continually updating and improving security (Bank and Clark 2005, Pacelle 2005). The constant or declining incidence of identity theft is also reported by economists at the FTC, as shown in Table 4.

Table 4
Level of Identity Theft

Year	*Incidence*	*Source*
2003	4.60%	FTC
	4.25%	Javelin
2004	4.00%	Javelin
2005	3.70%	FTC
	3.74%	Javelin

Source: Anderson et al. 2008.

[31] This represents costs to card issuers, and so is not comparable to the FTC numbers, which represent total costs to all businesses and consumers.

Note that most of the actual costs of identity theft are borne by businesses, not by individuals. No one is liable for more than $50 for misuse of a credit card and in many cases firms bear the entire cost. There are sometimes additional costs of being victimized, but in general only a small portion of consumers bear such costs (Anderson et al. 2008, p. 179).

B. Identity Theft and the Availability of Credit Card Numbers

Credit card numbers are widely available to criminals. Data collected by Symantec (2008) show that credit card numbers can be purchased illegally online for about 40 cents each. This means that the cost of a credit card number is about equal to the transactions cost of the exchange. A group that had stolen over 40 million credit card numbers and offered them for sale was recently arrested.[32] There is no shortage of credit card numbers for those who want to purchase them.

Another group used a "Trojan horse" program that took over computers and stole at least 500,000 credit card numbers and other information (Markoff 2008). This type of theft is unrelated to use or storage of data on search history, browsing behavior and similar activities.

As a society we have made significant efforts to require firms to provide increased levels of security for electronic data, such as enacting laws mandating data encryption and notice when data may have been compromised. In spite of these efforts criminals seem able to access credit card numbers with little difficulty. On the other hand, even though criminals can access these numbers, the actual amount of identity theft has been decreasing, as we discussed above. This means that factors other than efforts by legitimate firms to protect

[32]Brad Stone, "Global Trail of an Online Crime Ring," *New York Times*, August 12, 2008.

credit card information are responsible for the reduction in identity theft. We cannot exactly identify these factors, because firms in the industry are understandably reluctant to reveal the methods that they use. However, efforts by the credit card companies to quickly identify and limit the use of stolen cards are an important component. Rather than concentrating on information usage by legitimate firms, law enforcement authorities would do better to devote increased efforts to catching and punishing the actual criminals who engage in illegal activities. This would be more effective than trying to force all information users into adopting costly and ineffective security procedures.

The use of personal information is an important method of reducing credit card fraud. Firms use patterns of past activity to identify behavior that is inconsistent, and which may indicate that a credit card is being used without authorization. In addition, if there is some doubt about whether a number is being used fraudulently, a consumer can be questioned based on stored information about this consumer (e.g., past addresses). In these cases, the use of information serves to reduce risk. Rules that limit the use of information for purposes other than the purpose for which it was originally collected thus serve to increase risk by making it more difficult to identify fraudsters.

C. Notification Requirements

In addition to recommending limitations on data collection and use, the response to the identity theft problem has been to require that consumers be notified if their data are at risk. As of 2007, 39 states had enacted laws that would impose a variety of obligations on both businesses and public-sector entities in the event of a security breach, and provide remedies for

individuals whose personal information was acquired by an unauthorized party (Symantec 2008). The only study (Romanoski et al. 2008) of the effectiveness of these laws shows that they do not serve to deter identity theft.

In addition, notification requirements are dubious on benefit-cost grounds (discussed in Lenard and Rubin 2006). The expected benefits to consumers of such a requirement are extremely small—probably under $10 per individual whose data have been compromised. There are several reasons for this. First, most cases of identity theft involve offline security breaches, which are not affected by notification requirements. Second, the probability of an individual compromised by a security breach becoming an identity-theft victim is extremely small. Third, most of these are victims of fraudulent charges on their existing credit accounts, for which they have very limited liability, rather than victims of true identity theft. Finally even a well-designed notification program is likely to eliminate only a small fraction of the expected costs.

There are significant costs to regulating legitimate firms that have to do with over-deterrence of various sorts. As firms increasingly concentrate on security it can become more difficult to use data legitimately.

© Scott Adams, Inc./Dist. by UFS, Inc.

A significant potential cost of notification requirements is that they make consumers afraid of doing business online (Fountain 2005). As consumers receive more notices and read

more about dangers of identity theft from online business, they may be more likely to avoid it. This would be a costly reaction; in fact, Javelin suggests that "consumers should move financial transactions online to eliminate many of the most common avenues fraudsters use to obtain personal information and gain more control compared to traditional channels. Moving online includes turning off paper invoices, statements and checks, including paychecks, and replacing them with electronic versions" (Javelin 2009).

Javelin data also indicate that the mean time for fraud detection for paper statement review is 114 days, with a mean cost of $4,543; the comparable numbers for electronic accounts are 18 days and $551 (Javelin 2005). It is quite plausible that a continual stream of warnings could lead consumers to decide that online commerce is riskier than traditional paper commerce and, consequently, shift away from the online mode. Such a shift would have the unintended effect of increasing the identity-theft risks to which consumers are exposed.

D. Technological Fixes

Most identity theft is due to human error. In a *Scientific American* panel discussion on data security, Art Gilliland, vice president of Symantec (a major Internet security firm), estimated that "98 percent of the data loss is through mistakes of human error and process breakdown."[33] While efforts can be made to reduce this loss, technological fixes are of limited use against human error.

There are also tradeoffs. If technological fixes designed to increase security make data use more difficult, users may compensate by going around the technologies, thereby diminishing security. The above-cited discussion addresses such work-arounds as emailing data to oneself to

[33]"Improving Online Security," Industry Roundtable, *Scientific American*, September 2008.

work on at home and copying files to a USB drive in order to avoid corporate security measures. Less stringent measures might actually be more effective if they were less susceptible to bypass by users.

More recently, some states have begun to require that data be encrypted so that if it is stolen or lost it will not be usable. These requirements may cost about $100 per laptop and impose additional costs as well (Worthen 2008). Since state requirements may differ, firms will be forced to comply with conflicting state laws. Since the cost of identity theft has been overestimated, these rules may not pass a cost-benefit test. We are not, however, aware of any formal analysis of the encryption rules.

VIII. ARGUMENTS OF PRIVACY ADVOCATES

The arguments of privacy advocates are prominent in the policy debates concerning the use of information online. Advocates such as the Electronic Privacy Information Center (EPIC 2008) believe that "the detailed profiling of Internet users violates the fundamental rights of individuals, diminishes the accountability of large corporations, and threatens the operation of democratic governments." This view is shared by the Center for Digital Democracy (CDD 2007), which asserts that "the online advertising industry continues to ride roughshod over basic privacy rights…."

These groups worry that consumers do not know the extent of data collected about them or how it is used. For example, according to EPIC (2008), "opaque industry practices result in consumers remaining largely unaware of the monitoring of their online behavior, the security of this information and the extent to which this information is kept confidential. Industry

practices, in the absence of strong privacy principles, also prevent users from exercising any meaningful control over their personal data that is obtained."

Innovations, such as the development of search engines or, more recently, the possibility that Internet Service Providers might use deep packet inspection as an online-advertising tool, have increased apprehension. The Center for Democracy and Technology (CDT 2008a) claims "existing privacy protections have been far outpaced by technological innovation," and the collection of data by ISPs in particular "appear[s] to defy reasonable consumer expectations, could interfere with Internet functionality, and may violate communications privacy laws." CDT (2008b) expresses concern with the growing use of deep packet inspection to collect data, saying that it "raises serious questions about the future of trust, openness, and innovation online."

Some privacy advocates question whether online data collection benefits consumers by giving them more relevant information. The CDD (2007) claims that personalized advertising psychologically manipulates people to buy things they would not otherwise purchase, noting that "the growing use of neuropsychological research suggests that increasingly digital marketing will be designed to foster emotional and unconscious choices, rather than reasoned, thoughtful decision making." Jeff Chester (2007), CDD's executive director, warns, "I fear that such a powerful psychosocial stealth-marketing machine, backed by the yearly expenditure of many billions of marketing dollars, will drive personal consumption to greater excess," and that "They [the ad and marketing agencies] must ensure that consumers fully understand and consent to digital techniques; make certain that approaches to target our emotions and other brain behaviors are truly safe…."

There is an extensive legal academic literature on privacy. Ohio State law professor Peter Swire (2003) suggests that comparing the advantages and disadvantages of more privacy is problematic because the harm caused by disclosure of personal information is difficult to measure. "A variable such as the taste for privacy is 'soft' in the sense that it is difficult to quantify. In any quantitative estimates of costs and benefits, the soft variables can readily be excluded from the main analysis. Even important variables can thus be treated as an afterthought when they do not fit neatly into the analytic structure."

Swire (2003, p. 2) has observed that "…economists have systematically given less weight to privacy protection than academics trained in other disciplines" because of their view that "a competitive market is characterized by *perfect information*. The closer a market comes to perfect information, the better can willing buyers and sellers find each other." He notes that "the key insight for the economist is that privacy rules systematically reduce information flows." Essentially, for the economist, privacy rules "reduce the free flow of information, make it more difficult for buyers and sellers to find each other, and prevent efficient transactions from taking place."

Swire argues that the economist's "efficiency analysis leaves out much of what people actually fear in the area of privacy protection." Indeed, a complete efficiency analysis should try to accurately incorporate consumers' preferences for privacy, which would include their "privacy fears." Normally, economists try to evaluate the strength of these preferences by looking at actual marketplace behavior. Although survey results often indicate that consumers place a high value on privacy, their behavior in the market indicates they frequently trade information for a variety of benefits such as services available without charge, more useful advertising messages, and a more secure computing environment.

George Washington University law professor Daniel Solove (2008, pp. 82-83, 118-119, 123, 173-174) acknowledges that information is socially valuable and that there are tradeoffs between privacy and information. Nevertheless, his policy recommendations generally come down in favor of increased privacy. But he does not provide the data about the tradeoffs that are necessary for making an informed judgment about the desirability of policy proposals. For example, while he lists harms associated with information use, he does not quantify how frequent or serious they are.

Solove (2008, p. 73) believes that consumers are not behaving in their best interest due to asymmetric information; if they knew what information was being collected and how it was being used, they would be less willing to share their information. If this were the case, however, one would think that some firms would find it in their interest to educate consumers and highlight privacy protections to gain a competitive advantage. In fact, firms are quite sensitive to the privacy concerns of their customers. Recent episodes involving AOL and Facebook, who were punished for violating privacy expectations of their customers, illustrate the costs to firms of deviating from acceptable practices (Nakashim 2006; Havenstein 2008).

In addition, the market now provides numerous privacy tools that consumers can use. For example, the search engine Ask.com has an AskEraser option, which deletes an individual's search activity from Ask.com's servers.[34] New browsers increasingly allow personalization of privacy settings. IE8 has a number of features, including "inPrivate Browsing," which upon request omits browsing history, cookies, temporary files, and other data. It also improves the "delete browser history" function of IE7. Google's Chrome browser has an "incognito" setting which removes all traces of browsing, such as history or

[34] For details see "About AskEraser" at http://sp.ask.com/en/docs/about/askeraser.shtml

cookies. Firefox has various privacy settings, including a "Clear Private Data" button. These tools allow users to choose circumstances in which privacy is more valuable, and those in which information sharing is more useful.[35]

We would expect technologies that protect privacy to develop in a competitive market. The fact that they have developed suggests that firms are responding to consumers' preferences, even though consumers may not know the details about how their information is being used.

It is also important to distinguish between consumer preferences for "security" and for "privacy," which are not the same thing. People may be comfortable with the intended uses of data collected by search engines, for example. Search engines want to show relevant results and ads. Having more information helps them do this, as we discuss in Section IV. People are worried, however, about unintended uses of information such as identity theft, blackmail, extortion, embarrassment, etc. Their biggest privacy concern appears to be theft of a credit card number. Credit card theft is more of a security issue than a privacy issue. Credit card theft can also be exacerbated by misinformation about identity theft, as we discuss in Section VII.

IX. COSTS OF DECREASING THE AVAILABILITY OF INFORMATION

If one focuses on the qualitative benefits of privacy and ignores the costs of reducing the information flow, it is easy to conclude that more privacy is always better. But privacy is

[35] See also our discussion of tools available to consumers in Section IV on search engines.

not a "free lunch." Reducing the flow of information would require consumers to give up things they value.

There are a number of policies that would appear to give consumers more choices to enhance their privacy but would likely impose costs on consumers. Here we discuss two well-publicized examples.

A. The Default: Opt-In or Opt-Out

An important general issue in gathering information is the "default;" that is, if no decision is made, do consumers have the right to control the use of information, or do information gatherers have control? This issue is generally discussed as "opt-in" or "opt-out." Privacy advocates and legal privacy scholars are overwhelmingly in favor of an opt-in requirement.[36]

Under opt-in, the consumer by default controls information use and the website must obtain permission from the consumer to use information. Under opt-out, the website has the default right to use information about the consumer, and the consumer must make an effort to change this by opting out.

Richard Thaler and Cass Sunstein discuss the importance of the choice of default in their popular book *Nudge* (2008). They argue that "humans will often consider required choice to be a nuisance or worse, and would much prefer to have a good default… When choice is complicated and difficult, people might greatly appreciate a sensible default," (pp. 86-87). They go on to cite a study of opt-in versus opt-out enrollment in 401(k) plans, for which initial participation for opt-in plans was only 20 percent, eventually rising to 65 percent over three

[36] See, for example, CDD, CDT, and EPIC positions.

years, compared to 90 percent and 98 percent for an opt-out plan. They conclude that "automatic enrollment thus has two effects: participants join sooner, and more participants join eventually." (p. 109). Therefore, the choice of default has an appreciable impact on the eventual number of participants in a program.

The key insight here, as in many other situations, is that transactions costs matter. If transaction costs are zero, then the choice of the default—for our purposes, the original allocation of the right to control the use of information—does not matter. If, however, transactions costs are positive, it is efficient to give the right to the party who values it the most, or the party who would buy it if transaction costs were zero (see e.g., Posner 2007). It might appear that the transaction costs associated with making a decision about data collection are low, and therefore the default would not matter. However, as illustrated by Thaler and Sunstein, consumers have a tendency not to change the default, whatever it might be.

With respect to information usage, in testimony before the FTC a consultant indicated that with opt-out, one firm would lose about 5 percent of participants, while with opt-in they would lose about 85 percent.[37] This tendency of consumers to not change the default may be because the issue is not very important to them. Default inertia may also occur because transaction costs include learning about the nature of a choice, and may be higher than they first appear.

Businesses place a higher value on the right to use information than do consumers. Surveys show that most consumers would be willing to trade information for something specifically useful to them. As discussed above, the purpose of obtaining information about

[37]Testimony by Larry Ponemon, PriceWaterhouseCoopers, at the FTC hearing, *Wireless Web, Data Services and Beyond: Emerging Technologies and Consumer Issues*, December 12, 2000, Vol. 2, p. 232.

consumers is to provide them with targeted advertising—advertising of products likely to be of use to them—as well as with services, such as free search and email. These are the types of transactions consumers indicate they would like to engage in. This means that if transactions costs were zero, websites would end up with the information. Therefore, efficiency would argue for giving the initial right to businesses—that is, for opt-out. Noam (1997) and Varian (1997) also conclude that opt-out is the most efficient pattern. If the default were opt-in, then information would be lost—it would not flow to its highest-valued uses. This loss of information would be quite costly and would lead to price increases as firms attempt to compensate for the loss of information.

Sovern (1999), who is in favor of a mandatory opt-in system, provides an example that indicates the sort of transactions costs associated with opt-in:

> Evidence on how companies behave in an opt-in environment suggests that such a system may be more efficient for consumers than the current system. After the FCC ruled that phone companies seeking to use phone-calling patterns for marketing purposes must first obtain the consumer's permission, the telephone company in my area attempted to secure that permission. Its representatives called and sent mailings to subscribers. The company also set up a toll-free number for consumers with questions. The mailing I received was brief, printed in different colors, and written in plain English. It also promised, in words which were underlined, that "we'll never share this information with any outside company." A postage-paid envelope and a printed form were included for consumers to respond. Consumers who accept the offer need only check a box, sign and date the form, and print their name. The company also offered consumers incentives to sign up—such as a five-dollar check, two free movie tickets, or a ten-dollar certificate from certain retailers—thus increasing the likelihood that consumers will pay attention to the information. In sum, the company has done everything it can to eliminate consumer transaction costs.

Although this procedure may have reduced consumer transactions costs, it increased total transactions costs substantially. The increase transactions costs incurred by businesses

trying to induce consumers to opt-in are also a nuisance for consumers.[38] US West (using telemarketing) obtained an opt-in rate of 29 percent among residential subscribers at a cost of $20.66 per positive response (cited in Turner 2001). These higher transactions costs will ultimately be paid by consumers, either through higher prices or reduced services and benefits.

B. Do Not Track

Building on the popular Do Not Call List for consumers wanting to avoid telemarketing calls, a group of privacy and consumer groups proposed that the Federal Trade Commission implement a Do Not Track List that would allow consumers to block servers from tracking their online activities.[39] The proponents of a Do Not Track List assert that behavioral tracking "places consumers' privacy at risk, and is not covered by federal law," and that "it is time to move forward with something structured like the Do Not Call List to address problems we are seeing, and have now seen for seven years."

As we have discussed, the great benefit of the Internet as an advertising medium is the ability to target ads to consumers much more precisely than can be done through other media. This targeted advertising is based on developing an understanding of consumers' interests, then matching and delivering relevant advertisements. It utilizes personal information, sometimes including the past history of Internet browsing. Consumers get ads that are more useful to them and fewer unwanted ads. By increasing unwanted ads, the Do Not Track List would have

---

[38]Discussed in Fred H. Cate and Michael E. Staten, "Protecting Privacy in the New Millennium: The Fallacy of 'Opt-In'," Information Services Executive Council, available at http://www.the-dma.org/isec/optin/shtml.

[39]Discussed, for example, in Ryan Singel, "Privacy Groups Ask for Online 'Do Not Track' List," *Wired*, October 31, 2007, available at http://www.wired.com/politics/onlinerights/news/2007/10/do_not_track.

exactly the opposite effect of the Do Not Call List, which does reduce unwanted marketing messages.

Under the proposal, consumers on the Do Not Track List would still receive ads. The ads would just be less useful to them, because they would be less well targeted.

C. Self-Regulation

The Federal Trade Commission, which is the agency with primary jurisdiction over privacy matters, has proposed self-regulatory privacy principles, which represent a middle ground between a pure market solution and regulation (FTC 2009). Companies that agree to the self-regulatory principles would then be subject to enforcement by the FTC under the Federal Trade Commission Act. While self-regulation is more flexible than imposed regulation, it can still be quite rigorous.

Notwithstanding any specific provisions, a major adverse effect of self-regulation (or mandatory privacy legislation) would be to take privacy out of the competitive marketplace. Firms would agree to a common privacy regime that might be consistent with the preferences of a subset of consumers, but would likely not satisfy the preferences of the majority of consumers. Consumers' preferences for privacy are not homogeneous and there is no reason why firms shouldn't provide varying levels of privacy, just as they provide a variety of product and service characteristics.

X. CONCLUSIONS

This paper has discussed the substantial value for consumers produced by the use of consumer data for commercial purposes. The use of such data permits firms to target their

marketing messages to consumers' interests, pays for a wealth of content on the Internet, and helps protect consumers from a variety of online threats. It forms the basis for many of the business models that are fueling the growth of the Internet.

Since the online marketplace is quite competitive, with many participants, we would expect that any particular consumer's preferences for privacy would be satisfied by one or another of the sellers. If the preferences of a significant group of consumers were not being satisfied and could be met at a cost they were willing pay, it would be in the interest of some firms to satisfy those preferences. At several points in this paper, we have discussed how firms respond to their customers' privacy concerns, including introducing new technologies for consumers who want to shield their information.

Some privacy advocates and scholars argue that the market isn't working well because of asymmetric information—that consumers aren't well-informed about how their information is being used. Again, in a competitive market, firms willing to provide increased privacy have incentives to provide that information to consumers if they want it or would value it.

Moreover, given the amount of misinformation—e.g., anecdotes about the incidence of identity theft that have little or no basis in fact—about privacy-related issues, it is likely that the asymmetric information argument should go in the opposite direction. That is, the asymmetry would lead consumers to demand more privacy protection than is optimal.

Regulation should be undertaken only if a market is not functioning properly. Market failure here would mean that consumers' preferences concerning privacy are not being accurately transmitted and responded to in the market place. Good public policy requires that proposals for additional regulation be based on a showing that consumers are being harmed and that new

regulation would alleviate those harms in a way that the benefits are greater than the costs. There is no analysis that shows this.

REFERENCES:

Acquisti, Alessandro and Hal R. Varian. 2005. "Conditioning Prices on Purchase History" *Marketing Science*, V. 24, No. 3, Summer 2005, pp. 367-381.

Akerlof, George A. 1970. "The Market for Lemons: Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*.

Anderson, Keith B., Erik Durbin, and Michael A. Salinger. 2008. "Identity Theft." *Journal of Economic Perspectives*, V. 22, No. 2, Spring 2008, pp. 171-192.

Bank, David and Don Clark. 2005. "Visa sets antifraud system upgrade." *Wall Street Journal*, p. B4. June 13.

Bryan-Low, Cassell. 2005. "Identity thieves organize." *Wall Street Journal*, p. B1. April 7.

Bureau of Labor Statistics. 2005. "Real earnings in May 2005." Press release. June 15.

Campbell, Katherine, Lawrence A. Gordon, Martin P. Loeb and Lei Zhou. 2003. "The economic cost of publicly announced information security breaches: empirical evidence from the stock market." *Journal of Computer Security*, 11, pp. 431-448.

Center for Digital Democracy (CDD). 2007. "Supplemental Statement In Support of Complaint and Request for Inquiry and Injunctive Relief Concerning Unfair and Deceptive Online Marketing Practices." November 1.

Center for Democracy & Technology (CDT). 2008a. "Privacy Implications of Online Advertising." Statement of Leslie Harris Before the Senate Commerce, Science & Transportation Committee. July 9. http://www.cdt.org/testimony/20080709harris.pdf

———. 2008b. "What Your Broadband Provider Knows About Your Web Use: Deep Packet Inspection and Communications Laws and Policies." Statement of Alissa Cooper Before the House Committee on Energy and Commerce, Subcommittee of Telecommunications and the Internet. July 17. http://cdt.org/testimony/20080717cooper.pdf

Chester, Jeff. 2007. "Commentary: The Dark Side of Interactive Marketing," as reproduced on the Center for Digital Democracy website. http://www.democraticmedia.org/news_room/articles/2007/darkside_interactive_marketing

Dash, Eric and Tom Zeller. 2005. "MasterCard says 40 million files are put at risk." *New York Times*. June 18.

Electronic Privacy Information Center (EPIC). 2008. "Data Protection and Search Engines on the Internet: Google –DoubleClick and other case Studies." European Parliament. Brussels, Belgium. January 21. http://epic.org/privacy/ftc/google/EPIC_LIBE_Submission.pdf

Federal Trade Commission (FTC). "Take charge: Fighting back against identity theft." Available on the FTC website.

———. 2005. "Identity theft survey report." Synovate, September. available on the FTC website.

———. 2009. FTC Staff Report: Self-Regulatory Principles for Online Behavioral Advertising. February.

Fountain, Henry. 2005. "Worry, but don't stress out." *New York Times*, June 26, Section 4, p. 1.

Garg, Ashish, Jeffrey Curtis, and Hilary Halper. 2003. "Quantifying the financial impact of IT security breaches." *Information Management & Computer Security*, 11/2, pp. 74-83.

Havenstein, Heather. 2008. "After Beacon fiasco, new Facebook privacy controls score good reviews" *Computerworld*, March 19.

Internet Advertising Bureau and PriceWaterhouseCoopers. 2009. "IAB Internet Advertising Revenue Report: 2008 Full Year Results". March.

Jarrell, G. and S. Peltzman. 1985. "The impact of product recalls on the wealth of sellers." *Journal of Political Economy*, 93, pp. 512-536.

Javelin Strategy & Research. 2008. "2008 Identity Fraud Survey Report," Consumer Version available online, February.

———. 2009. "Latest Javelin Research Shows Identity Fraud Increased 22 Percent, Affecting Nearly Ten Million Americans: But Consumer Costs Fell Sharply by 31 Percent" press release. February 9.

Kang, Jerry. 1998. "Information Privacy In Cyberspace Transactions." *Stanford Law Review*. Vol. 50, April, p. 1193.

Lenard, Thomas M. and Paul H. Rubin. 2006. "Much Ado about Notification." *Regulation*, Spring, pp. 44-50.

Loeser, Julius L. 2000."Some Practical and Theoretical Thoughts about Privacy and Banking." Competitive Enterprise Institute, *The Future of Financial Privacy*, Washington.

Markoff, John 2008. "A Huge Cache of Stolen Financial Information," *New York Times*, Bits (Blog), October 31.

Morriss, Andrew P. and Jason Korosec. 2005. "Private dispute resolution in the card context: structure, reputation, and incentives." Case Research Paper Series in Legal Studies, Working Paper 05-12. June. Available at http://ssrn.com/abstract=735283.

Nakashima, Ellen. 2006. “AOL Takes Down Site With Users’ Search Data: Personal Details Posted in ‘Screw-UP’”. *Washington Post,* August, 8, p. D01.

New York Times. 2005. “Company news: LexisNexis restricts access to personal data.” March 19, nytimes.com.

Nilson Report. 2005. “Credit card fraud in the U.S.” March.

Noam, Eli. 1997. “Privacy and Self-Regulation: Markets for Electronic Privacy.” *Privacy and Self-Regulation in the Information Age*, U. S. Department of Commerce. Washington. http://www.ntia.doc.gov/reports/privacy/selfreg1.htm.

Pacelle, Mitchell. 2005. “How MasterCard fights against identity thieves.” *Wall Street Journal*, p. B1. May 9.

Pacelle, Mitchell and Christopher Conkey. 2005. “Card Industry Fights Breach Bills.” *Wall Street Journal*, p. C1. June 23.

Pegoro, Rob. 2005. “Voluntary disclosure is the threat to password security.” *Washington Post*, p. F7. June 12.

Peltzman, S. 1981. “The effects of FTC advertising regulation.” *Journal of Law and Economics*, 24, pp. 403-448.

Petty, Ross D. 2000. “Marketing Without Consent: Consumer Choice and Costs, Privacy and Public Policy,” *Journal of Public Policy and Marketing*, Spring.

Posner, Richard A. 2007. *Economic Analysis of Law*, Wolters Kluwer Law and Business, 7th Edition.

Romanosky, Sasha, Rahul Telang, and Alessandro Acquisti. 2008. “Do Data Breach Disclosure Laws Reduce Identity Theft?” September 16. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1268926

Rothschild, Michael and Joseph Stiglitz. 1976. “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information.” *Quarterly Journal of Economics*, pp. 629–649.

Rubin, Paul H. and Thomas M. Lenard. 2002. “Privacy and the commercial use of private information.” Boston, Kluwer Academic Publishers and the Progress and Freedom Foundation.

Rubin, P.H., R.D. Murphy and G. Jarrell. 1988. “Risky products, risky stocks.” *Regulation*, pp. 35-39.

Shapiro, Carl and Hal R. Varian. 1999. *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business Press.

Sidel, Robin and Mitchell Pacelle. 2005. "Credit-Card breach tests banking industry's defenses." *Wall Street Journal*, p. C1. June 21.

Singel, Ryan. "Privacy Groups Ask for Online 'Do Not Track' List," *Wired*, 10.31.07,

Solove, Daniel J. 2001. "Privacy and Power: Computer Databases and Metaphors for Information Privacy." *Stanford Law Review*, Vol. 53, p. 1393. July.

———. 2008. "Understanding Privacy." Harvard University Press. Cambridge, Massachusetts.

Solove, Daniel J. and Chris Jay Hoofnagle. 2005. "A model regime of privacy protection: Version 2.0." available from SSRN.com.

Sovern, Jeff. 1999. "Opting In, Opting Out, Or No Options At All: The Fight for Control of Personal Information." *Washington Law Review* 1033. October.

Stigler, George J. 1961. "The Economics of Information." *The Journal of Political Economy*, Vol. 69, No. 3, pp. 213-225. June.

Swire, Peter. 2003. "Efficient Confidentiality for Privacy, Security, and Confidential Business Information," available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=383180.

Symantec. 2008. *Symantec Global Internet Security Report: Trends for July-December 07*, available at http://eval.symantec.com/mktginfo/enterprise/white_papers/b-whitepaper_internet_security_threat_report_xiii_04-2008.en-us.pdf. April.

Synovate. 2007. "Federal Trade Commission—2006 Identity Theft Survey Report." Prepared for Federal Trade Commission. November.

Thaler, Richard H. and Cass R. Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven: Yale University Press.

Turner, Michael A. 2001. "The Impact of Data Restrictions on Consumer Distance Shopping." The Direct Marketing Association.

Varian, Hal R. 1996. "Differential Pricing and Efficiency". *First Monday*, V. 1,No. 2-5,August.

———. 1997. "Economic Aspects of Personal Privacy." *Privacy and Self-Regulation in the Information Age*, U. S. Department of Commerce, Washington, DC. http://www.ntia.doc.gov/reports/privacy/selfreg1.htm.

Wall Street Journal Online. 2005. “Without a trace: A *Wall Street Journal* online news roundup.” April 20.

Worthen, Ben. 2008. “New Data Privacy Laws Set for Firms.” *Wall Street Journal,* p. B1. October 16.

Zeller, Tom. 2005. “Black market in stolen credit card data thrives on Internet.” *New York Times*. June 21.