Volume Title: Economic Analysis of the Digital Economy

Volume Author/Editor: Avi Goldfarb, Shane M. Greenstein, and Catherine E. Tucker, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-20684-X; 978-0-226-20684-4

Volume URL: http://www.nber.org/books/gree13-1

Conference Date: June 6–7, 2013

Publication Date: April 2015

Chapter Title: Comment on "Information Lost: Will the 'Paradise' That Information Promises, to Both Consumer and Firm, Be 'Lost' on Account of Data Breaches? The Epic is Playing Out"

Chapter Author(s): Amalia R. Miller

Chapter URL: http://www.nber.org/chapters/c13025

Chapter pages in book: (p. 351 – 356)

Sullivan, Richard J. 2010. "The Changing Nature of the US Card Payment Fraud: Issues for Industry and Public Policy." Working Paper, Workshop on the Economics of Information Security.
Taleb, Nassim Nicholas. 2001. *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets*. New York: Random House.
———. 2010. *The Black Swan: The Impact of the Highly Improbable*, 2nd ed. New York: Random House.
Tang, Zhulei, Yu (Jeffrey) Hu, and Michel D. Smith. 2007. "Gaining Trust through Online Privacy Protection: Self-Regulation, Mandatory Standards, or Caveat Emptor." *Journal of Management Information Systems* 24 (4): 153–73.
Thomas, Russell Cameron, Marcin Antkiewicz, Patrick Florer, Suzanne Widup, and Matthew Woodyard. 2013. "How Bad Is It? A Branching Activity Model to Estimate the Impact of Information Security Breaches." Working Paper, Workshop on the Economics of Information Security.
Yan Chen, Grace YoungJoo Jeon, and Yong-Mi Kim. 2013. "A Day without a Search Engine: An Experimental Study of Online and Offline Searches." Working Paper, University of Michigan.
Verizon. 2014. Data Breach Investigations Report. http://www.verizonenterprise.com/DBIR/.

Comment

Amalia R. Miller

As more personal information about consumers is collected, stored, and transmitted by businesses in electronic form, the chances increase that records will be lost. Data breach incidents caused by malicious hacking or theft, or even by accidental equipment loss, can harm the consumers whose information is breached. In the wrong hands, personal information about consumers, such as their Social Security numbers, Internet search and browsing histories, insurance claims, financial transactions, and purchases can be used to harass, embarrass, impersonate, or steal from them.

This chapter argues that data loss is an important concern to be addressed in the digitization research agenda. Mann points in particular to the facts that data breaches at firms remain regular occurrences and that reported cases of breaches affect millions of individual records each year. Even if these breached records comprise only a small fraction of the total amount of data collected, consumer concerns about data breaches can have broader effects. For example, as their actual or perceived risk of data loss increases, consumers may engage in costly behaviors to protect themselves and become less willing to share their personal information with firms. Similarly, firms

Amalia R. Miller is associate professor of economics at the University of Virginia and a research associate of the National Bureau of Economic Research.

For acknowledgments, sources of research support, and disclosure of the author's material financial relationships, if any, please see http://www.nber.org/chapters/c13025.ack.

incur costs in protecting the digital information in their care about consumers. This is true for measures aimed at reducing the likelihood of successful breaches, such as data encryption, use of passwords, locks and physical security around computer equipment and storage devices, and security training and procedure for employees. It is also true for measures that reduce the harm from breach incidents (and the appeal of the data to potential thieves), such as collecting less, storing less, aggregating less, and transmitting less data.

This chapter represents an initial attempt to set out an agenda for economic policy research on the issue of data security that considers the tension between the costs of data breaches and of security efforts to prevent them. The chapter first proposes some possible frameworks for assessing the trade-offs and presents some arguments for government intervention. It then reports summary tabulations on disclosed data breach incidents in the United States or affecting US consumers in an empirical section. Finally, a section on policy approaches discusses data security regulations adopted in the United States and elsewhere. In this comment, I summarize the key contributions of each section and suggest some additional topics and issues for consideration in future research on the economics of data security.

The central theme of the framework section is that there are several potential market imperfections that would lead to underinvestment in data security by firms that possess private information on individuals. In particular, firms may not internalize the benefits of their investments in data protection because of incomplete markets for data safety. One reason may be that property rights are not clearly defined for data that is created, collected, and maintained by private companies, but that is *about* particular individuals (who can be harmed by its dissemination). Even with clear property rights, there are information asymmetries between consumers and firms when it comes to data protection, and without policy intervention it could be impossible for consumers to discover what steps are being taken to protect their information or when their information is lost or stolen. As Mann points out in the chapter, the inherent difficulties in assessing the risks of and harms from data disclosure may also prevent markets for data protection from developing because of high transactions costs in devising appropriate contracts. Hence, an agency problem can arise between consumers whose information is being collected and the firms that are entrusted with that information leading to insufficient investment in data protection.

The framework section also discusses the possibility that consumers are either uninformed about risks or not completely rational (or capable of understanding information about risks and rare events) as further reasons for underinvestment in security. The idea is not that consumers want more protection than they are able to obtain from firms but, rather, that they want less than they should want. While the arguments are reasonable, and it is possible that consumers should care more about security than they do (or that

they do care more than they show in their behavior), it is worth noting that this general type of argument based on irrationality or limited rationality could also lead to the opposite prediction. Consumers could easily overreact to small risks and demand too much costly data protection.

Although not discussed in the chapter, it may also be interesting to consider how the availability of insurance coverage that protects firms from the financial costs of data breach incidents (either as part of their casualty and property policy or as a separate plan) affects their decisions to invest in data security. Such coverage will typically dampen incentives for firms to invest in data protection, though the effects will be limited if coverage is incomplete (for example, because business loss and reputation effects are excluded) or if premiums are based on past claims experience. The effect could even be reversed in part if large insurers use their data on breach claims to provide incentives and useful guidance to firms about effective investments in data security.

Another issue that merits some attention is the possibility of externalities between firms from investments in data security. Negative spillovers will occur if the risk of theft increases after other firms make their data more secure. This spillover can lead to overinvestment in data security if companies feel compelled to match or escalate beyond the security levels of their competitors. However, if security protections reduce the value of data theft and are not visible to potential thieves, there may instead be positive externalities, comparable to those found in Ayres and Levitt's (1998) study of the LoJack device on auto theft.

In addition to outlining possible market imperfections that apply to data security, this section presents two separate discussions of the types of public policy responses used to address data security. In the first, presented in relation to Hirsch's (2006) pollution metaphor for the risk of data spills as an externality from greater data aggregation, command-and-control process regulation is contrasted with policies that target the outcomes of interest. Because security technologies and threats can evolve quickly, and firms may have better information about the costs and effects of different investments than regulators do, Hirsch argues the first approach, mandating specific data protection polices, is unlikely to be effective. There is some confirmation of this in the empirical finding in Miller and Tucker (2011) that state laws promoting the adoption of data encryption technologies led to an overall *increase* in incidents of data loss (driven by cases of internal fraud and loss of computer equipment). The second discussion of policy in the framework section takes a more legal approach. The options presented are the ones discussed in the policy section of the chapter: (1) mandated data security requirements, (2) government fines or penalties for data loss, and (3) mandated disclosure of data loss.

The empirical section of the chapter provides an overview of some of the recent trends in data breach incidents by economic sector, number of records

breached, content of breached data (including SSN or not), and source of the breach (such as computer hackers, paper document loss, computer loss, and insider fraud). While it is hard to infer much from this limited information, a few points are worth noting. Most interesting is that a large share of data loss incidents are not coming from hackers, but are instead a result of insider fraud, accidental loss, or unintentional disclosure. Second, there are differences across sectors in the numbers and content of breaches and trends over time. The trends generally show relatively stable numbers of breaches between 2005 and 2012 (the number of breaches is highest in 2006) and declines in the share of reported breaches with SSN data. However, as Mann points out, it is hard to know if this reflects a decline in actual incidents or worse reporting. Although the framework in the previous section and discussion in most of the chapter is focused on the private sector, the data section also includes summary information on data breaches in the government sector. In theory, comparing public and private sector breaches could provide information about the role of incentives or market imperfections, but that is not possible without a better understanding of the incentives in the public sector and differences in the types of information collected in each sector.

The evidence in this section is necessarily constrained by the limited information available about data breach incidents. As Mann acknowledges in the chapter, in order for the evidence to be useful to inform public policy, more information is needed. First, a major limitation is that researchers only know about publically disclosed (or discovered) incidents. The requirements for disclosure, as discussed in the chapter, are not comprehensive, which means that many incidents do not need to be disclosed to the public. Furthermore, even when disclosure is required, it is unknown how well firms comply with the requirements. It is also impossible to compute risks or rates of information loss from data on breaches alone. We need to know how to scale for the amount of data collected, which is surely increasing over the time period.

The next section discusses existing public policies addressing data security. Building on the second discussion of policy approaches in the framework section, the chapter expands the discussion of the relative benefits of disclosure requirements and subsequent "market discipline" compared to rules that mandate security protections or impose penalties for data breaches.

In describing the policy environment, Mann contrasts the approaches taken in the United States and the European Union (EU). European policy has tended to favor global requirements for privacy and security protection measures on data holders, often specifying technology, staffing, and procedures for data collection, use, and transfers. These rules include strict limits on what information can be collected or shared and how long it can be stored, as well as requirements to obtain consent for different actions. US policy on data security is more heterogeneous, varying at the state level and according to the type of information. The main policy lever applied in the United States is the requirement that data breach incidents be disclosed pub-

lically and to affected consumers. There are federal disclosure requirements for certain types of information (breaches affecting children or including financial or health information) and broader disclosure rules in many states.

After distinguishing between the typical US and European approaches to data security policy (process regulation versus disclosure rules), the chapter notes recent convergence in the area disclosure rules. The EU and several other counties are now implementing or considering imposing these requirements. Interestingly, the EU disclosure rules will be more demanding than those in the United States (greater coverage and shorter time frame) and will be applied along with regulatory fines for data breaches. It is also worth noting that some US state and federal rules do mandate specific security procedures, such as data encryption or consent requirements for exchange, and other rules require that standards for "reasonable protection" be met without explicitly listing them or include incentives for the adoption of certain security technologies in the form of relief from disclosure requirements for data breaches (see, for example, Miller and Tucker 2009, 2011).

Assessing the appropriateness of different policy responses to data security requires empirical information about the relative costs to firms and consumers of data breach incidents and their prevention. This chapter reports some preliminary estimates of the costs to firms and consumers from disclosed data loss incidents, suggesting that costs are at least twice as large for consumers. Specifically, the chapter cites an estimated range out-of-pocket cost to consumers from each data breach of $400 to $700 from a report by Javelin Strategy and Research, and an estimated range of average costs to businesses per record lost of about $100 to $200 from a report by the Ponemon Institute. These initial estimates suggest that data disclosure requirements are not sufficient to cause firms to fully internalize even the costs of disclosed breaches, which suggests a role for greater intervention beyond disclosure rules to increase investment in data security.

However, uncertainty about these values, and the fact that they only reflect average costs, means that more reliable and extensive information could either support or overturn this initial conclusion. Furthermore, the actual costs of data loss incidents are likely to change over time as a result of evolving policies outside of data security. As discussed in the chapter, the costs to businesses from breaches should be expected to increase in the future if the legal environment evolves in such a way that consumers gain access to more recourse options to sue companies for breaches even without showing direct harm. The basis for these claims would be negligence for failing to meet industry standards for data protection or breach of contract for violations of privacy or security provisions. These legal options would make disclosure a more powerful tool for encouraging security investments, though enforcement would impose costs on consumers and lead to uncertain outcomes, which might be avoided with a more direct regulatory approach, such as penalties for breaches.

It is important for policymakers to recognize that the level of business investment in data security is not the only factor that affects the frequency or cost of data breaches. Consumer behavior, such as deciding what information to share and with what companies, as well as their reactions to disclosed breaches (involving their own data or not), can affect data loss and may itself respond to changes in public and corporate policies. For example, the cost of breaches may be low if consumers limit or distort the information they provide to firms. If improved data security at the firm level makes consumers more willing to share personal information, or less careful about protecting themselves from theft (by actively monitoring their financial accounts for fraud and checking credit reports for possible identity theft), the incremental cost from successful breaches could increase. The cost of breaches could also increase with better data security if the breaches that are least costly to prevent involve information that is least valuable to thieves and least harmful to consumers. Finally, the costs of data breaches are also affected by public and private efforts to prevent, detect, and penalize attempts to *use* lost or stolen data. For example, careful monitoring of credit card charges, financial transfers, and insurance claims by companies that process these transactions can prevent thieves from making use of the data. On the public policy side, data theft is already illegal, as are most fraudulent and malicious uses of lost or stolen data. However, enforcing these rules requires that resources be devoted to law enforcement for investigating crimes and developing new tools to address new threats. The ideal combination of data hoarding, data protection, consumer and firm efforts to detect fraud, and government efforts to investigate and punish fraud and theft, will depend on the costs of these efforts and their effects on the frequencies and costs of data loss.

In summary, this chapter about data loss introduces an important topic to the research agenda on the economics of digitization. It raises many questions for researchers and policymakers to consider and summarizes some of the initial empirical information on the topic. This is a research area with many open questions and opportunities for contributions by economists.

References

Ayres, Ian, and Steven D. Levitt. 1998. "Measuring Positive Externalities from Unobservable Victim Precaution: An Empirical Analysis of Lojack." *Quarterly Journal of Economics* 113 (1): 43–77.

Hirsch, Dennis D. 2006. "Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law." *Georgia Law Review* 41 (1): 1–63.

Miller, Amalia R., and Catherine Tucker. 2009. "Privacy Protection and Technology Diffusion: The Case of Electronic Medical Records." *Management Science* 57 (7): 1077–93.

———. 2011. "Encryption and the Loss of Patient Data." *Journal of Policy Analysis and Management* 30 (3): 534–56.