Protecting Consumers in Privacy and Data Security: A Perspective of Information Economics

Ginger Zhe Jin and Andrew Stivers[1]
Discussion Draft

May 22, 2017

Executive Summary

This note provides an economic approach to consumer privacy and data security based on the extensive economic literature on how information flows, and is used, in the marketplace. We apply that approach to consumer protection in privacy and data security, as a step toward the ultimate goal of facilitating well-grounded cost-benefit analysis of future policy and law enforcement action in this area.

Over the past two decades, the FTC has led the governmental effort to protect the integrity of consumer privacy choices in the market.[2] This note attempts to describe the economic basis for that work in one coherent piece. Given the scope of the topic, this note is only a first step in providing clarity on the economic perspective on consumer protection in this area. We hope the note will improve future actions in privacy and data security. As a matter of scope, we do not discuss the potential implications of privacy and data security for antitrust or competition. Nor do we discuss data access and data use beyond domestic commerce.[3]

We also do not claim to provide *the* economics of privacy and data security. Other authors have provided the basic research and surveys that this work builds on. Other perspectives might usefully focus on how property rights in information influence the creation and flow of information through the market, or how structuring privacy as a human right would change markets and influence social welfare. Such other perspectives might also facilitate cost-benefit analyses of privacy and data security policy and enforcement.

In comparison, we articulate privacy and data security issues primarily in information economic terms. In particular, we highlight the distinction between process and outcome: while an individual's privacy outcome is the realized restriction on the flow and use of information, the

---

[1] Director and Deputy Director, respectively, Bureau of Economics, Federal Trade Commission. The views expressed here are our own and do not necessarily reflect the views of the Federal Trade Commission or any individual Commissioner.

[2] FTC actions include over 100 cases on privacy and data security, a guidance of privacy to business and policy makers (FTC 2012), a report on data brokers (FTC 2014), a report on big data (FTC 2016a), staff comment on the privacy regulation proposed by the Federal Communication Commission (FTC 2016b), and numerous blogs, statements, and public workshops on privacy and data security.

[3] The access and use of personal data for national security and law enforcement purposes are beyond the scope of this note.

process that leads to that outcome depends on many parties. The decisions that each of those parties make about how that information flows, and the control that each party exercises over the flow of the individual's information, all contribute to the privacy outcome. The distinction between process and outcome is important. This is in part because while consumers may prefer more or less privacy – the outcome – given a particular situation, they all want themselves, and by extension the sellers they interact with, to have a certain amount of control over the flow. In line with other areas of market intervention, a focus on the process ensures that consumers and sellers have the tools to exercise appropriate control on the process. In turn, this should help bolster a healthy market to facilitate and honor their choice of privacy. This approach is in contrast to a more paternalistic approach that attempts to determine consumer preferences on privacy outcomes and directly impose that determination on the market.

Next, we articulate how consumer-to-seller information flows are more complicated in practice than the typical seller-to-consumer information flows that typically concern the FTC in other areas, such as advertising enforcement. In particular, information flows generated by a transaction can persist over time and create effects outside of that initial transaction. These persistent effects often complicate market incentives. Consequently, three market failures may arise:

a. **Information Asymmetry**: Data security and privacy policies are largely credence[4] characteristics even with direct partners in a transaction. In addition, data persistence means that consumer valuation of an information flow is a function of the network of entities that access and use that flow. The complexity of this network, combined with the difficulty in credibly conveying and committing to these policies, creates an information problem: it is difficult for consumers to be fully informed of the potential network of decisions and outcomes, process that information, and decide whether to allow their private information to flow to the network. The information asymmetry means that consumer decisions are potentially dependent on prior beliefs, or assumptions about data policies and the trust-worthiness of seller claims about those policies.

b. **Externality**: The opaqueness of the network often makes it difficult to establish causal linkages between seller policies and their effects, both positive and negative, on the consumer. In some cases, there may only be a probabilistic linkage between action and realized harm. This creates positive and negative externalities that no one actor may have full ability or incentive to internalize.

c. Commitment: The persistence of information means that transfer of private information is a sunk investment which could allow ex post (perhaps unilateral) renegotiation of how that information could be used or protected. Even an actor that knows all of the current protections and uses for her data cannot count on future use and protection. The

[4] A credence characteristic is an attribute of a product that is never directly evident to the consumer before or after the transaction.

commitment problem is especially relevant in privacy and data security, where perceptions about the value and use of consumer data is rapidly changing, and the technologies needed to control that data are also evolving.

Because the persistence of information can cause commitment problems, and tends to exacerbate information asymmetry and externality, it is more starkly important for policy makers to foster a healthy information environment about privacy outcomes and processes, to encourage industry to develop standards and mechanisms that support a healthy information market, and to police that market if necessary. In this document, we identify some market mechanisms that have arisen to address potential market failures, and when interventions in consumer protection are most likely justified. In light of these potential market failures, we then list potential policy tools and discuss their pros and cons.

Table of Contents

Section I: Privacy and data security in information economic terms

In commercial transactions, buyers and sellers often hold "private information" about attributes, preferences, costs and willingness to pay (WTP). In the information economics literature – and by extension this paper – "private information" is a value-neutral term referring to how the information is initially revealed. Private information is only revealed to a limited set of parties, in contrast to "public information" which is revealed to all parties. The term is not a normative judgement on whether that information "should" remain private. If information is private in this sense, it means there are choices that could be made about access to that information, i.e., about privacy and data security.

Privacy refers to some restriction in the flow and use of (initially) private information – in particular, what information, to whom, and how it is used. An entity has more privacy as the flow and use of information about it is more restricted.

Broadly speaking, both buyers and sellers may have privacy: consumer privacy restricts the access and use of that consumer's private identity, preference and WTP; whereas seller privacy restricts the access and use of that seller's private information in identity, attribute, price, quality and cost. The dominant public discussion focuses on consumer privacy, but a comparison of consumer privacy and seller privacy will highlight key economic issues surrounding consumer privacy. For this reason, we will discuss both before devoting the policy discussion to consumer privacy.

In the context of privacy, we note a distinction between process and outcome: while outcome is the realized result of who has what information in which use, the process that leads to that outcome depends on many parties and their decisions about and control over the flow of information. Examples of less restricted privacy outcomes include harms – a thief gaining access to credit card information– and benefits – a mortgage company identifies a good credit risk. Similarly, examples of more restricted privacy outcomes include benefits – thieves do not gain access to credit card information – and harms – a mortgage company cannot identify good credit risks. The process that would lead to these outcomes potentially includes a long chain of decisions by consumers, intermediaries, banks and others that are given access to data, and how tightly those entities control that access.

Consumers and sellers may prefer a more or less restricted privacy outcome given a particular situation, but they all want themselves, and the parties they interact with, to have an appropriate control on the flow so that they can have confidence in their understanding of the outcomes.[5] In that sense, the ability to control the information flow can be more universally thought of as a "good" across different actors and contexts, all else equal. In particular, we

[5] According to the literature review conducted by Acquisti et al. (2016), consumer and societal preferences for privacy outcomes are context and condition-specific, and this heterogeneity is acknowledged in both theory arguments and empirical analyses. However, an overwhelming majority of US adults, as shown in the 2015 survey by Pew Research Center, believe that it is important to be in control of who can get information about them (Madden and Rainie 2015).

emphasize that the universal preference for a tighter control is conditional on the same cost of control. It is likely that people will have different preference on how much *more* they are willing to pay in order to enhance the control on the process, but they agree that more control is better if it entails no extra cost in time or money.

As articulated later, the process involves many parties in and beyond the initial transaction, and these parties often have control of only part(s) of the process. For example, when a consumer initially decides to purchase and use a smartphone, she is exercising a data-process decision, which is the only independent opportunity for her to control completely her privacy outcome with respect to information collected or generated by her phone. Once the phone is in use, the consumer is typically given some choices by the seller of the smartphone about the amount and type of data she is willing to pass to (or through) the seller of the smartphone, and some choices about how secure she wants that flow to be. This flow typically includes private information of the consumer that exists independent of her phone use, as well as data that is generated by the consumer's interaction with the phone. This data could initially be considered jointly private information of the consumer, the phone company, and potentially other parties such as the producer of the phone's operating system or an app developer.[6] Beyond that point, the consumer's privacy outcomes are dependent on the seller's decisions about what to do with the data and how tightly to control it.

The seller may choose not to use the data that it has access to. It may choose to use that data to influence its interaction with that consumer or other consumers – for example with pricing or advertising decisions. Alternatively, the seller may intentionally share data with business partners. The consumer's privacy outcome is then a function of what the firm and each of its partners does with the data – in essence, how privately each of the partners holds the consumer data. The consumer's privacy outcomes are also dependent on how tightly the seller and its partners control the flow of consumer information within and between themselves. Any of these businesses may attract unwanted "partners" who break into the data flow. In sum, from the consumer's point of view, her privacy outcome depends on the data process decisions adopted by herself, the seller, the seller's intended partners, and possibly unintended "partners" that break into the data flow.

Given this dependency, the consumer's initial data process decision relies on the information she receives about the privacy and security policies of all the partners that will have access to her data. For example, a consumer may prefer that information about her healthcare is not generally accessible, but she is comfortable with a wide sharing of that data between her caregivers and her insurer. A perfectly informed choice about whether to share her information with that network would require the consumer to know whether all partners have agreed not to share the data outside the network, and whether they have adopted reasonable security measures to restrict the flow of that data. Given the information asymmetries, the consumer is dependent on representations made by the partners, or third parties about these policies.

[6] See footnote 2 for the definition of private information. A phone number, for example, is at its creation jointly private between the phone company and the consumer.

These representations are then subject to all of the FTC's usual concerns about market information.

As with privacy, data security applies to both sides of a transaction in a market setting. We typically are not concerned with the security of data as it flows from seller to buyer because we often presume that the seller has incentives to exercise tight control on what information goes to consumers, and any information shared with at least one consumer need not be kept private by that consumer.[7] In the other direction, there are varying contexts and opinions about how carefully a company should control the flow of data it receives from consumers. For example, the presumption on an individual health record is that it must have high security (be tightly controlled) by both consumer and seller, but consumers might be comfortable with a lower level of security on, for example, information about their coffee buying habits. In general, privacy outcomes cannot be given a general preference ordering because people have differing preferences for who accesses their private data and how they use the data. This heterogeneity holds not only between different people but also within an individual across different contexts.

On the other hand, the process that leads to the privacy outcome, including but not limited to data security, is a measure of how tight and transparent the control is, and can be more plausibly given a general preference order. It is costly to provide more of it, but keeping cost and all else equal, entities affected by some flow of initially private information will weakly prefer a more controlled flow to a less controlled one.[8] In addition, a focus on the process helps ensure that consumers and sellers have appropriate tools to exercise control on the process and that a healthy market exists to facilitate and honor their choice of control. In contrast, a focus on privacy outcomes requires determining which privacy outcomes are preferred in what circumstances, and is therefore both more difficult and more invasive.

To better illustrate the concepts of privacy and data security in the market context, we lay out a stylized transaction. Consider a potential buyer for a product X (e.g. loan contract, physical good, communication service, etc) who has private information about, for example: her name, physical and email addresses, credit card number, income, expenses and product – and other – preferences. There is a seller with private information about the product bundle offered, including customer service quality, customer data collection, storage and use, and other physical or performance attributes associated with the product itself. When the two negotiate over a potential transaction, the buyer may learn about some seller attributes and the seller may learn about some buyer attributes. The details of the information exchange may affect the final contract terms for product X (price, service specifics, contract length, etc), both in terms of bargaining power and in expectations about the value of the transaction. The expectation of the value of the transaction is in turn influenced by how the buyer and the seller value privacy and data security as part of the product bundle.

[7] The obvious exception is data as product i.e., digital content. Exploring the parallels between securing digital content and securing consumer information is left for later work.

[8] Note that control can equally imply the ability to credibly share information and the ability to withhold it.

For example, knowledge of the home address of a consumer shopping for a gym membership could influence how aggressively the manager of a gym might price membership. An address very close by might signal that the consumer would be willing to pay a price premium, while an address far away might signal a need to offer a discount. That is, knowledge of the consumer's address might change both the expectation of value and the bargaining position of the manager. If given a choice of providing identifying information to the manager, the consumer might consider the benefit or cost to her in terms of how it would influence the price she would be offered for a membership.

However, that flow of information may also trigger other changes to the consumer's valuation of the product bundle. The consumer may have direct preferences about sharing personal information, and the consumer may have some expectations of how that information might be used in other transactions. Questions about the persistence and value of information beyond the initial transaction are reasons why privacy and data security deserve extra scrutiny in the world of consumer protection economics.

The consumer may also care about the extent to which her private data are accessible more broadly. Her potentially accessible data profile would depend on not only her information exchange with the gym manager, but also many other transactions she engages in with grocery stores, banks, personal laptop, Internet browser, phone service providers, and many apps on her cell phone. Non-commercial activities may contribute to this data profile as well, as she updates her personal websites, uses her library card, interacts with local police, and files federal and local taxes. The further the data flow away from the consumer, the more the consumer has to rely on other actors' data policies for her privacy outcome.

Altogether, there is a network of data relating to the consumer: a public profile may contain limited information about the consumer and be available to anyone; a number of "private" profiles may only be accessible to authorized employees; and some version of the "private" profiles may be found on the black market. Market or non-market actors may be able to combine the public and "private" profiles to increase the amount of information about the consumer. Exactly who ends up having what information defines the privacy outcome of the consumer (at some point in time), which in turn depends on data security and other control measures adopted by the consumer, the entities that she interacts with directly, and all third parties that access the data.

Generally speaking, privacy and data security are potentially relevant to consumer protection in three ways: (1) through the flow of data that may change the efficiency and surplus distribution of an initial transaction; (2) through the persistent effect that the data flow of the initial transaction may have on transactions or parties separate from the initial transaction (the magnitude of which may overshadow the narrowly defined value of the initial transaction); and (3) through consumer preference over the data flow in the initial transaction that may enter consumer's utility function directly as an attribute, independent of bargaining power or other interactions. We discuss each separately below.

(1) Privacy and data security governs information flow within the initial transaction (a classic asymmetric information problem)

A typical transaction has information flow in both directions: sellers may inform buyers of product attributes and price; buyers may reveal her preference for product attributes and WTP. The information flow can be controlled or uncontrolled.

a. Seller privacy in the initial transaction

Many economists have studied information flow between buyers and sellers, but most efforts focus on one direction: namely, sellers' abilities and incentives to hide, share or misrepresent their product information in front of consumers.[9] A large literature demonstrates that seller privacy may reduce market efficiency because it may encourage bad-quality products to flood the market (Akerlof 1970 and the follow-ups). This in turn increases consumer search cost for price and quality information, and makes it more difficult to match products with consumers.

The key question in that literature is whether there are enough market mechanisms to overcome seller privacy. Both high and low quality sellers have incentives to convince buyers that they are high quality, but that potential pooling of signals is hard to achieve if sellers lack data security, and thus have low privacy. In particular, buyers may have access to existing, credible seller identity, product attributes, transaction history and consumer experiences via consumer-review websites. Sellers typically cannot control such information sharing, although some sellers have attempted to use gag clauses in contracts. These are essentially data security measures to increase low quality sellers' privacy. On the other hand, high quality sellers – sellers with a negative value to privacy – have incentives to reduce privacy by credibly revealing their high quality via advertising, signaling, reputation, voluntary disclosure, and third party certification[10]. By purposefully reducing their privacy, high quality sellers exercise their control over the information flow about their product's quality. However, this also effectively reduces the privacy of the non-disclosing sellers[11], even if non-disclosing sellers adopt gag clause and other data security measures. This indirect effect helps illustrate that control of privacy is rarely absolute, and that there are potentially complex interaction effects.

[9] This note focuses on information exchange between sellers and buyers. Information exchange between competing sellers could raise a concern of collusion. In that context, maintaining seller privacy from other sellers may promote market efficiency. Similarly, greater privacy in a seller's intellectual property may encourage the seller to invest in the technology in the first place, which could have a positive impact on the overall market efficiency.

[10] Some classical theories include Spence (1973) on signaling, Akerlof (1970) on adverse selection, Grossman (1981) and Milgrom (1981) on voluntary disclosure, Nelson (1974) on advertising, and Shapiro (1983) on price premium for reputation.

[11] For example, Jovanovic (1982) models disclosure decision in a competitive market, and allows consumers to form a rational expectation on the true quality of non-disclosing sellers in equilibrium.

Driven by efficiency arguments, many consumer protection policies aim to reduce the costs or other barriers to the flow of credible information from seller to buyer, i.e., *reduce* seller privacy in price, quality, and other transaction terms.

However, revealing sellers' private information to consumers does not necessarily benefit *every* consumer, because credible public information about a high quality product may allow a seller with market power to charge a higher price and, and thus make the product unaffordable to some consumers.[12] In other words, less seller privacy may generate a distributional effect where some consumers are better off and some consumers are worse off.[13] Competition may help to reduce the undesirable (from some consumers' points of view) distributional effect of information revelation, and to reinforce the market efficiency arising from more seller information made public.

b. Consumer privacy in the initial transaction

Changes in retail markets have elevated interest in the other direction of information flow, namely buyers' abilities and incentives to hide or share their private information about personal attributes, WTP and product preference. Similar to the seller-to-buyer information problem, greater buyer privacy may reduce market efficiency within a particular transaction because it may: increase sellers' search cost for consumers; increase the chance of inefficient pricing and product offerings; and discourage new sellers from market entry.[14]

Again, the key question in the economics literature has been what market mechanisms arise to address the information asymmetry due to high consumer privacy. Both high and low WTP buyers have incentives to convince sellers that they demand high quality but have a low WTP. That pooling is hard to achieve if buyers lack privacy. As with seller privacy, buyers may have low privacy for multiple reasons. For example, a consumer may choose to share data by boasting about luxurious possessions in her public Facebook postings. This consumer may then find it difficult to pretend to be low-income in front of a seller.

Alternatively, firms could exploit existing holes in consumer data practice by creating tools to scrape public tax records for the consumer's home ownership information, or by tracking unencrypted identifying information as the consumer navigates the internet. For the buyers that have high (initial) data security, sellers may offer "free" software applications to incentivize them to allow seller access to their private information. Even if sellers do not offer any direct

[12] At the same time, this could increase seller incentives to invest in new products, which may reinforce or counteract the distributional effects on consumers.

[13] In the large economic literature about price discrimination, the common wisdom is that price discrimination requires market power, and using it (as compared to uniform pricing) will increase firm profits, lower price for some consumers, and raise price for others (Tirole 1988). Whether the total welfare increases or decreases with price discrimination depends on change in total output (Varian 1985; Holmes 1989), potential new market opening (Hausman and Mackie-Mason, 1988), and the regulatory environment for a regulated monopolist (Armstrong and Vickers 1991).

[14] Papers by Stigler (1980) and Posner (1980) lay out some of the standard economic effects of greater privacy.

incentives for data access, some consumers may choose to use mechanisms that credibly share their private information with the seller – for example, certified health records, driving safety or credit worthiness – if doing so will improve price or other terms of the transaction. The fact that some consumers are willing to reduce their privacy may allow some sellers to infer the hidden information of non-disclosing consumers. In short, like seller privacy, buyer privacy is the result of data security and privacy decisions by a variety of actors, and markets have evolved to reduce consumer privacy as a way to potentially increase efficiency of the focal transaction.

As in the seller-to-buyer problem, more information from buyer to seller can have an ambiguous distributional effect on consumers. A seller with market power and buyer information can offer a low price to low-WTP buyers and a high price to high-WTP buyers. Conversely, in a market with heterogeneous WTP, high consumer privacy forces sellers to offer a single price to all consumers.[15] This typically allows consumers to capture a greater share of surplus, because privacy endows consumers with an informational asset. For those consumers with a true WTP greater than the single price, high privacy yields a positive return to private information. [16] In contrast, those with lower WTP may not be profitable for the seller to serve, and therefore may not be able to buy. This would create a negative return to private information, and these consumers would be better off with less privacy. Again, competition may alleviate the distributional problem if buyer classification is available to many sellers and there is enough competition among informed sellers for a particular type of consumers.

In summary, *within a transaction* privacy protection *increases* the bi-directional information asymmetry between buyers and sellers, which may reduce market efficiency and affect surplus distribution in ambiguous ways. Increased data security may or may not increase privacy, depending on the incentives of buyers and sellers, but will also have generally ambiguous effects. While neoclassical economists often advocate for more information flow from sellers to buyers because they believe market efficiency dominates distributional concerns, the same tradeoffs exist in the flow of information from buyer to seller in the context of the focal transaction. Limited to this narrow framework, whether the efficiency gains are dominated by distributional concerns depends on the objectives of the policy maker, the market institutions influencing the flow of information, and the specifics of the particular situation. Overall, these informational effects within a transaction are well defined: any harms are bounded by the potential surplus generated by that transaction alone, and are thus relatively straightforward to address under a traditional consumer protection framework.

(2) Privacy and data security can have persistent effects beyond the focal transaction

[15] In some markets, sellers may be able to offer a menu of prices to all consumers (e.g. volume discount) and let consumers sort themselves into different volume or package options. Still, knowing more about consumer heterogeneity will allow sellers to adjust the menu above and beyond what they have achieved without additional consumer information. See footnote 10 for a brief summary of the price discrimination literature.

[16] An economic rent is an unearned return, often because the asset in question is scarce and has been endowed to the owner.

As is well known, information flows generated by a transaction can persist over time and create effects outside of that initial transaction. Indeed, market participants – both buyers and sellers – may generate transactions specifically to create information flows that they can use to affect future transactions. These outside effects can be positive or negative, or both, as are those in the initial transaction. The welfare effects from this persistence – positive or negative – can potentially outweigh the initial effects, or even the value of the initial transaction. Here we first note several well known outside effects, to illustrate how the persistent effects complicate buyer and seller calculations of the benefits and costs of information flows, and therefore both increase the potential for needing consumer protection efforts, and complicate the determination of what that effort should be. We also discuss some market responses to these persistent effects.

a. Seller privacy with persistent effects

Much of the information that can be brought to bear on negotiating the focal transaction is generated by a previous one. One of the best-studied persistent effects is seller reputation. For example, consider the reputation of a home building contractor. As discussed above, information about the quality of the contractor's work that is revealed in previous and current projects may influence the final negotiated price between the homeowner and that contractor. This effect occurs within the focal transaction. In addition, that information flow associated with this project could influence the price paid by new clients. A sophisticated homeowner could, by offering (or threatening) to give referrals, take advantage of the contractor's low data security (inability to control information about herself), to motivate the contractor to internalize that persistent effect. Conversely, the persistence of that information could lead the contractor to increase her data security, by including a "gag-clause" in her contract that would give her more control over information about her work. Internalizing these effects into the initial transaction would have efficiency implications (a referral could result in a job otherwise not completed) as well as distributional effects for buyers and sellers.

Another well understood persistent effect related to seller privacy is technological spillover. New product characteristics shared with a buyer are not necessarily limited to enjoyment by that buyer. Other sellers can incorporate those new product characteristics into their own products. These spillovers can be positive for the seller – for example in creating standardization of features that broadens the market for particular products – or negative – for example in increasing competition for that product while being likely beneficial to the market as a whole. Either of these effects can increase, or decrease market efficiency, in part depending on how much of the persistent effects can be internalized. Patents and trade secrets, as controls on the flow of technology information, function as data security in this situation.

b. Consumer privacy with persistent effects

One prominent effect of persistent consumer information appears in targeted advertisements. This advertising can be valuable to consumers if it reduces the search costs for products and

services in which consumers have demonstrated an interest. By increasing the relevance – and thus the value – of a fixed amount of advertisements, low consumer privacy could also increase the value of the organic, advertising-supported content that consumers view. For example, news reports can be timelier and online video postings can be more entertaining, if targeted advertisements that support these contents are more effective, and therefore generate more income for the content provider. However, by increasing the marginal value of advertising to consumers, the platform that sells tracked advertising also has an incentive to increase the amount of advertising that a consumer receives.

The net effect is often ambiguous.[17] For those that perceive all tracked ads as having positive value, low privacy could be beneficial. However, consumers have heterogeneous valuations on receiving advertising, and ad targeting is often imperfect. This means that targeted ads have distributional effects on consumers, even if they may improve the overall matching efficiency of between products and consumers.

As with seller reputation, buyer reputation – especially in the form of WTP – can affect the efficiency and distribution of surplus in future transactions. Credit scoring is a widespread application of buyer reputation mechanisms that take advantage of low consumer privacy with respect to previous transactions.[18]

Similar to the seller side effect of technology spillovers, consumer information could motivate the development of new, better-customized products, either by the initial seller or by others that receive that data. Buyers could even benefit if they have negative returns to private information in the initial transaction. For example, the cost of finding low WTP consumers may be too high for a low-cost basic-function product to exist for those consumers, but the low cost of sharing information that has already been collected may encourage product entry. Conversely, access to better consumer information, and the targeting of high value buyers with new products could result in some buyers either not being served at all, or having to settle for less valued products. As above, sharing this information with other sellers could exacerbate the effect: some buyers would suffer if they have positive returns to private information beyond the initial one, as they potentially face higher prices in future transactions.

All of the above effects are *intended by sellers*. Just as sellers can take advantage of consumers' lax data control to use their information, other entities may take advantage of a seller's low data security to capture and use that same information. Most worrisome is that consumer data, aggregated by large sellers with low data security, may create a cost-effective target for

[17] See Acquisti et al. (2016) section 3.1 for related empirical studies. Consumers recognize the ambiguity as well. According to the 2016 Survey by the Pew Research Center, consumers acknowledge free service, grocery discount, and other data-dependent benefits, but they are also upset about unwanted contacts and lack of control after they give out their private data (Rainie and Duggan 2016).

[18] For the same reason, the collection, use and correction of consumer credit information are governed by a series of legislations, including the 1970 Fair Credit Reporting Act, the 1996 Consumer Credit Reporting Reform Act, the 1999 Gram-Leach-Bliley Act, and related local regulations.

malicious actors to exploit.[19] The resulting costs to consumers could include direct financial costs (in identity theft for example) as well as indirect financial costs arising from reputational shocks – particularly in terms of job loss.[20]

Finally, some classic externalities may arise to third parties that are not directly involved in the initial transaction. Non-market agents may also want to use buyers' data to improve access to credit, security, health, infrastructure or other social services that return benefit to the society as a whole. The controversy around Google flu trend – its ability to help the government better predict influenza and related privacy concerns[21] – highlights the outside effects that consumers' internet search data could generate beyond the search itself.

c. Market incentives to internalize persistent effects

Valuing these persistent effects is difficult, both because of the uncertainty over whether they will have a significant impact on the parties generating the information, and because of uncertainty about how significant that impact would be, if it happens. These persistent effects may be externalities – a classical source of market failure – when the market does not provide sufficient information or tools for participants to fully internalize those effects into the initial transaction. Thus, before deciding on whether and how to intervene in the market, we must understand the extent to which the market has already internalized these persistent effects.

In the absence of intervention, sellers in competitive markets have incentives to implement relatively strong data security to allow them to internalize the benefits to themselves from using consumer data. These benefits may arise from developing new products, carrying out better product customization, maintaining their own market advantage and otherwise maximizing profit. If consumers know those effects and themselves can control seller access to their data, the seller should have an incentive to incorporate those effects – positive or negative – into the product bundle that it offers to the buyer as part of the initial transaction that generates the information exchange. As evidence that sellers are internalizing these

[19] According to the Identity Theft Resource Center, hacking, skimming, phishing attacks were the leading cause of data breach recorded in 2016 (www.idtheftcenter.org/2016databreaches.htm).

[20] According to the Bureau of Justice Statistics, about 7% of age 16 or older (17.6 million) were victims of identity theft in 2014, similar to findings in 2012. The number of elderly victims of identity theft increased from 2.1 million in 2012 to 2.6 million in 2014. Moreover, 9 in 10 identity theft victims did not know anything about the offender, and about two-thirds of identity theft victims reported a direct financial loss. The combined direct and indirect loss is on average $1343 per victim in 2014, with a median around $300 (Harrell 2015). In addition to loss from identity theft, consumers subject to data breach are also motivated to monitor their credit scores and freeze access to their credit reports, though not necessarily change their lenders (Mikhed and Vogan 2017).

[21] Dugas et al. (2012) demonstrated the correlation between Google Flu trends and official data on influenza from the US Center of Disease and Prevention. However, both Butler (2013) and Lazer et al. (2014) pointed out problems in the prediction algorithm, in addition to privacy concerns raised by the Electronic Privacy Information Center (EPIC, https://epic.org/privacy/flutrends/EPIC_ltr_FluTrends_11-08.pdf.)

benefits, many sellers currently offer free service, loyalty programs, and price discounts to persuade consumers for more data provision.[22]

Nevertheless, it is still possible that consumers do not fully internalize all persistent effects even when they have a tight control on their data. In the consumer's mind, the offer of discounts and the like in exchange for consumer data may be tied more to overcoming the direct cost of providing that information, rather than to the persistent effects outside the transaction. Consumers may not recognize extent of the persistent effects, may have biased beliefs about them (in either direction), or may not know about them at all.

Sometimes, consumers choose to release their private data to sellers because the cost of tightly limiting the initial outflow exceeds the perceived benefits. When the persistent effects on consumers are positive and aligned with seller benefits, limiting information outflow will not be valuable to the consumer (especially if sellers still have high data security and are transparent about their actual use of the data). However, when persistent effects on consumers are negative (or perceived negative), sellers may take advantage of consumers' trust, ignorance (or other reason) not to control the outflow of their data to avoid having to internalize that harm. For example, in the case of criminal data breach, sellers have arguably little incentive to internalize the negative external effects because the buyer may have interacted with many sellers and it is often difficult to link a particular seller's lax data security practice to the realized harm. This is a classical common good problem. No seller would have full incentive to adopt socially optimal data security measures to reduce the risk of data breach. Sellers in this situation, facing consumers with lax data control, will prefer to simply take the data rather than incur data security costs to prevent data breach or disclose the potential risk of future data breach thus reducing the perceived value of the product bundle it is offering.[23]

Similarly, sellers have incentives to be opaque about how they use and share consumer data within their business network, and whether a data breach has occurred recently.[24] This is

---

[22] Technically, nothing prevents consumers from reselling access to their data to other sellers, assuming that they can find a technically feasible way to do so. This could prevent an internalization of the benefits.

[23] This is still true even if buyers know that data breach at a particular seller has occurred. In that case, publicity of data breach hurts the seller's reputation, which motivates the seller to internalize the reputation cost in her data practice. Research has shown that the drop of stock price, in response to publicized data breach, is often small and temporary (see review in Acquisti et al. 2016). Sellers may also suffer other loss from reputation and direct expenses. According to a survey by Ponemon (2016), the average per capita cost of data breach was $221 in the US. On average, each data breach episode cost the breached organization $7.01 million, of which 3.97 million was due to lost business and 1.72 million was due to organizational response to data breach. These are all direct costs to the breached firm. Sellers experiencing these direct costs may still have sufficient incentives to internalize the potential harm of identity theft and the like, if the link between data breach and consumer harm is probabilistic.

[24] According to a survey by Thales (2017), 68% of (corporate) respondents have experienced a breach at some point, and 26% have experienced a breach last year (2016). Based on media tracking instead of corporate survey, the Identity Theft Resource Center has recorded 7,356 breaches since 2005, involving 895 million records. Nevertheless, these alarming numbers likely understate the actual incidence of data breach, as not all states require private or governmental entities to notify individuals of security breaches of information involving personal identifiable information. As of 4/12/2017, 48 states, the District of Columbia, Guam, Puerto Rico and the Virgin Islands have enacted such legislation, according to the National Conference of State Legislators (accessed at

especially true if one cannot establish a causal link between how a consumer experiences the use of data (for example, pop-up ads on the consumer's computer) and exactly which seller is responsible for enabling that experience (for example, who shares the consumer's data with an ad network). In this example, sellers have strong incentives to recoup the positive returns they may earn from selling the data to the ad network, but they have weaker incentives (if any at all) to consider any annoyance consumers may experience from targeted ads versus non-targeted ads. This incentive asymmetry will exacerbate consumer harm, as the chain of data sharing goes further and further away from the initial transaction that originates the data.

Finally, even if sellers reveal how they use and share consumer data at the time that consumer information is collected, they may not have incentives to abide by the initial agreement. Knowing this, sophisticated consumers may refuse to participate in efficient exchanges, as is well documented in the "hold-up" literature.[25]

In summary, expanding our view to the possible persistent effects of consumer data shows the increased scope for information problems and misaligned incentives. These economic forces may yield suboptimal outcomes and potentially harm consumers when data security and other control of the privacy process are too low due to structural issues or information asymmetries. At the same time, these persistent effects yield an even wider mix of hard-to-measure ambiguous efficiency and distributional results tied to the spread of information itself – i.e. to privacy.

(3) Privacy and data security preferences may enter consumer's utility function directly.

So far, we have discussed the role that information flows play in and out the focal transaction, assuming that consumers value the attributes of privacy and data security only to the extent that they affect the value of market interactions – including how that influences bargaining position, product offerings, and the risk of non-market harms. However, consumers may also have a direct preference for data collection or storage, independent of whether the use of that data ever directly affects them. For instance, they may feel that others' access to their private information is an intrusion to their private life, which generates a negative utility regardless of what, if anything, is done with the information. The presence of uncertain persistent effects could also enter utility negatively as a risk premium. Risk averse consumers may dislike the risk of their private information being uncontrolled regardless of the value of the actual outcomes (e.g. in terms of blackmail or identity theft), and would rather pay a premium to avoid the risk, over and above any expected harm.

Section II: What problem are we trying to solve?

---

http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx.)

[25] See, for example, Hart and Moore (1988).

In summary, privacy and data security often involve multiple parties and extend beyond the initial data-generating transaction. While the above discussion has touched on three classical market failures (commitment, information asymmetry and externality), we must distill them into concrete terms before we can articulate, and thus analyze, the benefits and costs associated with potential policy interventions. From a consumer protection standpoint, what problems might we be trying to solve beyond what the market has already done in this area? In this section, we make a few comments about privacy and data security, which that we hope with facilitate further economic input into practical policy development.

First, privacy and data security are part of the product bundle that a seller offers to a buyer. Viewing them as product attributes, consumers may value them – positively or negatively – because they affect the contract terms of the initial transaction, they generate benefit or harm to the consumer in future transactions, and they trigger direct feelings about privacy intrusion and risk aversion.

We noted a distinction between outcome and process: while privacy outcome is the realized restriction on the flow and use of information, the process leading to that outcome depends on many parties and their control (or lack of it) of the flow. As evidenced in many research studies, consumers have differing preferences for who accesses their private data and how they use the data.[26] Thus, it is difficult to give privacy outcome a general preference order that fits all consumers in all contexts. On the other hand, the process – which includes data security and information sharing decisions by all parties, as well as the flow of information about the actual use of consumer data – is about the ability to appropriately control the flow of data. Preference for the process can be more plausibly ranked in a general order. Keeping cost and all else equal, entities that faced with some flow of private information will weakly prefer a more controlled flow to a less controlled one. Admittedly, different consumers – or the same consumer in different contexts – may have different willingness to pay for a particular level of control by herself and the whole network behind the focal transaction[27], but this heterogeneity does not undermine the general order. As demonstrated in the 2015 survey by the Pew Research Center, 93% of US adults agree that being in control – on who can get information about them – is important.[28]

In short, we view privacy outcome and privacy process as two related concepts. Because it is easier to rank the general preference for the process than for the final outcome, it may be more straightforward for policy makers to target market failures in the process. In line with other areas of market intervention, a focus on the process ensures that consumers and sellers have the tools (including information) to exercise desired control on the process and that a healthy market exist to facilitate and honor their choice of control.

---

[26] See Acquisti et al. (2016) for a comprehensive review of this literature.

[27] Acquisti et al. (2016) section 3.6 reviews a literature of consumer willingness to pay for privacy, which is context-specific and ranges widely both across and within the cited studies.

[28] Madden and Rainie (2015), accessed at http://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/.

What market failures shall we target in the process? There is clearly an information problem: if we treat a seller's data practice as an attribute of the traded product, it is likely a credence attribute.[29] Consumers may not know what data policy they are exposed to when they purchase a product.[30] This could happen if the seller hides or misrepresents its data policy, if consumers fail to understand a truthful but complicated or misleading data policy, or if consumers and seller cannot associate a particular data policy with the real risk it exposes. Even if the seller articulates his data policy ex ante, consumers may not know what privacy consequence they should expect from that data policy. It can be equally, if not more, difficult to find out the actual data practice post transaction[31], especially when there is insufficient monitoring and enforcement from the government or other third parties.

This brings us back to the familiar, seller-to-buyer information problem. According to the economic literature, this problem can be addressed in many ways, ranging from the seller's incentive to advertise and disclose truthful information to regulations that mandate information disclosure or certain product attributes.[32] Received wisdom dictates that the market should adequately match the demand and supply of product attributes, if rational consumers have full and truthful information about the product attributes.[33] However, if consumers have wrong or insufficient information about a particular attribute, we may observe a race to the bottom for that attribute as predicted by Akerlof (1970). In our context, that could mean every seller provides the least protective data practice, even though most consumers prefer some protection.

Naturally, the next question is whether it is feasible to ensure that consumers are rational and have full and truthful information about sellers' data practice as a product attribute. There are two layers to this question: can consumers be fully informed about their choices, and can consumers make a sensible choice conditional on full and truthful information? The former depends on the extent to which the market – and policy makers if needed – can gather the right information, deliver it to consumers at the right time, and make sure that consumers fully understand the information as it is intended. This is a challenging task, as consumers need information on not only the seller's data practice in the initial transaction but also the consequence of that practice (or its breach) in future transactions.[34] As discussed above,

[29] Even when a consumer values privacy and data security because of the use that the data might be put to, data practices may be largely credence characteristics, as consumers cannot necessarily link out-facing seller practices to their data. However, in theory one might be able to link the two.

[30] Both Turow (2003) and Smith (2014) asked survey respondents to answer true or false on the following statement: "When a company posts a privacy policy, it ensures that the company keeps confidential all the information it collects on users." In 2003, 57% answered true; this number only dropped slightly to 52% in 2014.

[31] Lecuyer et al. (2015) show statistical evidence that the real use of Gmail ads (before Google abruptly shut them down on Nov. 10, 2014) contradicted two Google statements regarding the lack of targeting on sensitive and prohibited topics.

[32] See Fishman and Hagerty (1997) for an overview.

[33] O'Brien and Smith (2014) examine this wisdom to privacy in online markets, assuming sellers can commit to transparent privacy policies that are understood by consumers.

[34] The literature review by Acquisti et al. (2016) highlights consumers' inability to make informed decisions about their privacy (largely due to imperfection information), and because of that, heuristics can profoundly influence consumer's privacy decision making.

market-driven information exchange may suffer from misaligned incentives and externality, especially when the seller's data practice carries a risk of (unintended) negative consequence for consumers but it is difficult to trace it back to the seller when the harm is realized later.

The second layer of the question – can consumers make a sensible choice conditional on full information? – relates to consumers' cognitive and behavioral costs or biases. A large and longstanding economic literature has studied and documented the costs associated with processing and using information – and a more recent economic literature has highlighted consumers' limited ability to understand and act upon information – especially when information is complex, implicit, hard to understand, involves uncertainty, or requires prediction into the future.[35] Unfortunately, information about privacy and data security is likely to fall in these problematic areas. Thus, consumer ability to process information can be a real issue and could exacerbate other market failures.

One item worth highlighting is consumers' ability to control the flow of private information out of their own hands. A seller's data practice is an important product attribute, but how valuable (or harmful) that attribute is depends on what information the consumer is willing to give out and in what form. Conversely, consumer willingness to give out information depends on her perception of the seller's data practice. Adding to the complexity is that some or even most of the information at issue may be jointly, but not publicly, revealed at its generation. Given the interdependence, when there appears to be materially uncompensated flows of private information, we should examine the seller-to-buyer communication: are sellers imposing structural barriers to consumer's control, or are they manipulating the information environment to induce those flows? Are there transactional costs or behavioral choice considerations that impede consumers from implementing their preferences over information flow? We should also acknowledge potential problems in consumer-to-seller communication. That is, do consumers have credible channels to convey their private information to sellers, when desired? In short, the two halves of a consumer's data problem go hand in hand: consumers need to have appropriate tools to control the flow of information as they desire, while sellers and other stakeholders need tools to control the flow of consumer information in and out of their hands, and to convey the extent of that control clearly to consumers.

This last point tees up the commitment problem that may arise once consumers have given private information to sellers. Even allowing for perfect information flow about what sellers intend to do and have done with consumer data, and perfect control by sellers about what has happened to that data, the evolving eco-system for data makes commitment to those practices difficult and costly, even for sellers that want to commit ex ante. Here questions arise about the availability of commitment mechanisms in the market and the role and ability of government for enforcing those mechanisms. The fundamental trade-off for market efficiency is between mechanisms that are too weak – providing little confidence to consumers – and mechanisms

[35] Kahneman (2011) summarizes how two systems of human brain affect the individual decision making process, often with cognitive limits and behavioral bias. Conlisk (1996) summarizes the literature on bounded rationality, including cognitive costs.

that are too strong – inducing unacceptably high risks and costs on sellers. Either imbalance could prevent welfare enhancing trades.

Section III: Policy tools and their economic consideration

Policymakers in general can use a range of policy tools to address the above issues, each with its own pros and cons. We examine a broad spectrum of tools. However, not all of these tools, nor the specific approach articulated in our discussion of them, are necessarily appropriate or feasible for the FTC or any other current government agency. Therefore, our discussion should not be taken as endorsing their use in any particular case by any particular entity. Furthermore, the use of these tools implies that market solutions – including private tort, competition policy and the like – are insufficient, an assumption that itself deserves careful examination. As noted in the introduction to this paper, such an examination is a prerequisite to the cost benefit analysis needed for any particular application.

Before discussing potential policy tools, we should note that it is difficult to measure the harm these issues may bring to consumers, for all the reasons discussed above. In addition, applying generalizations about harms to any given case is complicated by the heterogeneous and interdependent value of the data at issue. However, before using a particular policy intervention, some estimation of the significance of the effects on consumers should be made to determine whether that intervention is likely to result in net benefit to consumers. Ideally, determinations about whether consumers are being harmed should be based on how consumers would act and gain in the market in the absence of market failures. That is, a policymaker should be as agnostic as possible as to what is important to consumers, and rely on consumer data to inform us about the magnitude of harms.

As a practical matter, for example, policymakers have used the "sensitivity" of consumer information as a proxy for the potential seriousness or magnitude of harm. While this may be relatively clear for some information – SSN, individual health records, and similar – other information is more contextually sensitive and requires a judgement, often in the absence of data, about whether the particular use or collection would result in net harm. For grey areas of sensitivity, a careful and contextually sensitive approach may be more appropriate than drawing a universal bright line.

Once a policymaker has decided that a market failure exists and it is likely to cause net harm to consumers, it has a variety of tools to apply. To focus the discussion on policy tools, this note shies away from the practical details of how to address a particular case or how to characterize consumer harm in that case. Rather, we discuss the general pros and cons of each policy tool, and leave practical details to follow-up work. The costs that various policy options impose on the market are difficult to measure, and can be hard to predict. They are equally important to address in practice, however, and we outline the cost implications of each tool below.

Tool 1: Educate stakeholders to reduce information asymmetries and encourage them to internalize externalities.

Arguably, a centralized policymaker has an advantage in gathering certain kinds of information and disseminating it to the whole market. Although individual sellers may have first-hand

experience and could collect some market information more easily than a centralized agency, they may choose that effort depending on how it benefits themselves rather than consumers or other sellers. As a result, individual sellers tend to underinvest in information gathering and market education. In comparison, and in theory, a government agency does not have the incentive to favor any particular seller and can easily reach out to all sorts of stakeholders.[36] For example, the agency could educate consumers on industry data practices and the likely effects of those practices on consumers. The agency could also educate sellers on what practices are material to consumers, and how those practices can be voluntarily disclosed. The agency can educate third parties on the value of tracking, clarifying, ranking and publishing sellers' privacy and data security practices.

No matter whether an agency education targets buyers, sellers or third parties, this is an *indirect* way to address concerns about consumer information flows. It uses the market force of consumer demand. More specifically, when consumers know the persistent effects associated with data practices, and those effects are material to their choices, they may implement better data control for themselves, prefer firms with better data practices, and avail themselves of third party certification tools. Such preferences can be manifested in higher WTP, which in turn motivates a greater supply of tools and practices for a better process on privacy.

For this market mechanism to work, a few conditions must be met: First, consumers must know firms' data practices and the associated persistent effects. This may be difficult to achieve: when consumers do not have a technological background; when firms lack an incentive to disclose their data practices in a simple and user-friendly way; when it is difficult to quantity the risk of data breach and subsequent harms; and when the link between data practices and risk of consumer harm is weak, noisy or hard to describe. Second, even given the availability of all that information, acting on all that information requires a significant amount of resources by consumers, which may prevent them from engaging if the cost outweighs what they perceive to be the benefit. This may be especially true if a data policy is not obviously linked to the primary material attributes of a product.

Significant questions remain in how best to ensure the most efficient outcomes using educational tools. Some market forces may help to meet the heavy requirements on consumers. Markets may develop technical tools to measure and convey the risk of data practices. However, before that happens, tool developers must have knowledge of firms' true data practices and of outcomes. This may not be easy to come by because firms do not have full incentives to reveal their true data practices or outcomes. Therefore, policy makers may need to provide those incentives for firms to reveal their true data practices and outcomes. In addition, it will be a technical challenge to cover all firms, but partial coverage can paint a biased picture if consumers do not understand why some firms do not disclose their data practices.

[36] This is not always true. See, for example, Laffont and Tirole (1991) and the literature on regulatory capture.

In summary, Tool 1 may address the likely market failures due to information asymmetry (about data policy) and negative externality (beyond the current transaction), but it assumes away the potential market failure due to consumers' information processing costs and potential behavioral biases. Note that this tool does not necessarily imply law enforcement actions. It could be a policy-oriented program that: defines elements in data practice; educates consumers and sellers about data practices and their persistent effects; disseminates the disclosed data practices; and encourages industrial self-regulatory body to monitor whether the disclosed practices are appropriate.

Tool 2: Enforce seller truth-telling by raising the cost of providing false information

Like Tool 1, Tool 2 attempts to address market failures due to asymmetric information and negative externality, but it has the following caveats:

First, unless there is a bright line between "true" and "false" information, raising the cost of providing false, but voluntary, information will likely raise the cost of providing true information as well. Economic theory and empirical work predicts that firms with more secure and more privacy-friendly data policies are more likely to disclose them when truthful (but voluntary) disclosure is enforced.[37] Hence the firms that voluntarily disclose their policies may be simultaneously mitigating the asymmetric information problem and opening themselves up to closer scrutiny as a target of investigation (as in Tool 1). Tool 2 may inadvertently reduce the amount of voluntary disclosure by incentivizing silence. Mandating disclosure (Tool 4 below) may mitigate this problem.

Second, in theory, the effectiveness of Tool 2 (in solving market failures) relies on the assumption that consumers read, understand and act upon the content of privacy notice. This assumption is not necessarily correct.[38] However, it may be that "expert consumers" help to drive the demand of less engaged consumers toward firms with better practice.[39] In this framework, the expert consumers would provide the fuel for voluntary disclosure, even if the disclosure is costly for the seller. In this case, its effectiveness depends on (1) the incentives of

[37] This is assuming that data policies are mostly concerned with laying out a firm's data security policies - which are likely to be considered a quality attribute by most consumers. When consumers are willing to pay more for higher quality and disclosure cost is independent of quality, high quality firms have more incentives to disclose than low quality firms. Mathios (2000) and Jin & Leslie (2003) examine the effect of mandatory disclosure for more traditional products.

[38] Cranor and McDonald (2008) estimated that it takes an average American 76 work days to read privacy notices encountered through a year, which explains why few consumers read privacy notices and few readers understand them.

[39] Starting from Stigler (1961), economists often believe that lower search cost (for price information) will lead to lower average price and lower price dispersion. However, both theoretical and empirical literatures suggest that the effect of lower search cost on price dispersion is ambiguous (see review by Baye, Morgan and Scholten 2005). Part of the price search literature highlights the positive externality from searchers to non-searchers: because searchers are more sensitive to price, their presence motivates some sellers to offer low price. Absent of perfect price discrimination, such low price is available to (some) non-searchers as well if the store they patronize randomly happens to be the low price store.

expert consumers (why they will spend significant time and efforts to digesting a privacy notice?); (2) the fraction of expert consumers in the whole population (it must be significant enough for the sellers to care); and (3) the information exchange from expert consumers to other consumers (it must be material enough to drive the demand of less engaged consumers). Under right conditions, this framework can motivate voluntary disclosure and is similar to encouraging third party activities as described in Tool 1.

Tool 3: Lower the cost of communication by directly monitoring firms' real data practice

As with Tool 2, Tool 3 is focused on the traditional market failures associated with information asymmetry between consumers and sellers. In Tool 3, an agency may itself act as a third party that directly monitors firms' real data practice.[40] For example, an agency might provide financial backing to researchers that audit the data practice of a random sample of big firms; it might investigate a firm's data practice in response to a media-reported data breach, a security researcher's discovery of security holes, or market complaints; and it might publicize the good and bad of data practices. Unlike Tool 1 and Tool 2, Tool 3 does not rely on firms' voluntary disclosure of their data policies. However, for the same reason, Tool 3 may be most costly to implement and require more strategy in targeting.

Tool 3 aims to address the likely market failure due to information asymmetry (on data policy), persistent negative effects (of using data beyond the current transaction), and lack of commitment (after the initial data collecting transaction). As with Tool 2, since it may not rely on the "average" consumers' ability to understand and act upon information about firms' data practices, it could also address the market failure due to consumers' cognitive and behavioral bias.

Tool 4: Mandate sellers to disclose privacy and data security practice

Tool 4 may be a more direct way to address informational distortions. As noted above, enforcing truth-telling in voluntary information disclosures (Tool 2) may have unintended consequences. In addition, other market failures could prevent voluntary disclosure and education campaigns from bridging the information asymmetries. Where these market failures can be identified, mandatory disclosure might be an appropriate solution.

However, many questions arise in how to implement the mandate: for example, if the mandate does not specify what to disclose and in what format, sellers may use complicated language or provide overwhelmingly long privacy notices, consequently consumers may not be able to obtain meaningful information from the disclosure.[41] This concern may motivate standardized mandates on the content and format of data policy disclosure. However, standardization on the

[40] This tool is similar to regulations that require or encourage a government entity to monitor the production process of certain goods. Examples include USDA meat inspection and FDA inspection of drug manufacturing plants.
[41] A growing literature demonstrates firm incentive to obfuscate consumers. See Gabaix and Laibson (2006) for an theoretical model, and Ellison and Ellison (2009) and Brown et al. (2010) for empirical evidence in online markets.

most effective elements would be difficult in this context, where the technology is evolving, the industry practice is diverse, and consumer preferences are heterogeneous.

Setting aside the difficulties in determining required content, enforcement would still require validation of seller practices. Otherwise, every seller would have an incentive to declare that she adopts the most desirable privacy and data policy (from the consumer's point of view), and consequently, that uniform declaration would not convey useful information.

Tool 5: Set minimum quality standard or other requirements on how firms should collect, store, use and share data.

Establishing and enforcing data security and privacy standards is the most direct way to address market failures in information asymmetry, negative persistent effects, and potential behavioral bias that could stand in the way of sellers making efficient investment in data policies. The persistency and probabilistic nature of outcomes for data security and privacy suggest that there may be useful lessons to be drawn from the approach other inherently risky, but valuable, products like foods and drugs. Requirements could be implemented by prescriptive legislation and regulation or through the market implementation of more general guidance backed by case-by-case government enforcement. The prescriptive regulatory approach is typically more transparent and precise, because government enforcement actions occur over time and in specific factual contexts that may not generalize and may cause difficulty in maintaining consistency. However, case-by-case enforcement is more flexible and capable of maintaining relevance over time. Especially because of the dynamic nature of the technology of collecting, storing and using of consumer information, more general requirements may be less costly to implement than explicit standards. As discussed above, because consumer preference is more diverse and heterogeneous for privacy outcomes, it may be more difficult to set up minimum quality standard for those outcomes than for process measures such as data security and transparency.

No matter how Tool 5 is implemented, it involves a number of key questions: First, on what basis are the standards set? And who sets the minimum quality standard? While there are a variety of "best practices" and guidelines, no global cost benefit analysis has been conducted, and it is not clear whether there is clear consensus on the right valuation of consumer privacy or data security. Second, is there a way to develop a meaningful standard that is flexible enough to account for rapidly changing technology in a timely manner? Third, how should standards be conveyed to firms so that they know the standard *before* an agency alleges they fail to meet the standard? This might be a chicken-egg problem. If an agency relies on casework to establish a minimum quality standard, the standard will need to allow for new technological developments. More important, they may be tempted to measure a firm's past behavior against a desirable standard, which may generate regulatory uncertainty and have perverse consequences in technological adoption and improvement.

In addition to the above economic tradeoff, one should not underestimate the many other technical economic considerations of Tool 5 that could be explicitly dealt with in rulemaking.

For example, how much leeway should be given to firms so that they can adjust to comply if a change in the standard occurs? At what frequency should the standard be reevaluated and the industry educated about it? When consumers have heterogeneous preferences, what consumer – or expert – preference should be used to gauge the standard? And for any standard, big and small firms may face different costs of compliance. Even if the compliance cost is the same, a fixed compliance cost can be trivial for a big firm but burdensome for small firms. Will this give an advantage to big firms and stifle competition?

According to the economic literature on minimum quality standards, the key shortcoming of this approach is that it may preclude consumers from accessing a segment of the market in which they prefer to trade. More specifically, assuming perfect knowledge and perfect enforcement, any firm producing a product with a quality below the standard is shut down. If some consumers truly prefer these products (with their associated relative prices) to the above-standard market, they are made worse off. This result hinges on whether consumers prefer the below-standard market because of their inherent preferences or because of their lack of knowledge or behavior bias.

The economic literature also points to the use of minimum quality standards when there are external effects on third parties.[42] The mandate of seatbelt use is a classic example. While individual privacy and data security choices may have implications for others' privacy and data security, this is a difficult issue for traditional consumer protection, as third party spillovers have not been a primary motivator. If there is no way to internalize these third party effects, and they are large enough to counterbalance the costs, it may be efficient to impose standard practices.

In summary, the five tools listed above target the main data security and privacy problems that face consumers, which in turn arise from market failures due to information asymmetry, negative persistent effects, and commitment problem. Some of them may also address problems arising from consumers' cognitive costs or behavioral biases. These tools are not exclusive to each other: in many cases, educating consumers and other stakeholders (Tool 1) can increase the effectiveness of voluntary or mandatory disclosure (Tool 2 and Tool 4); and direct monitoring (Tool 3) can produce valuable input for market-wide education (Tool 1).

The use of these tools is also likely to vary by context: for consumer-facing products, firms may spell out their data policy and obtain consumer consent before contracting. Thus, voluntary (Tool 2) or mandatory disclosure (Tool 4) can be an effective way to ensure transparent and informed choice from consumers. However, for products that do not face consumers directly – for example, a firm's contract with an ad network for consumer-customized ads – it is difficult for consumers to know and act upon the firm's real data practice. In these cases, direct monitoring (Tool 3) or minimum quality standard (Tool 5) may be more practical than disclosure.

[42] The classical economic theory on minimum quality standard starts with Leland (1979). A large literature follows, with a number of empirical studies estimate the actual effect of minimum quality standard in many markets.

Similar heterogeneity applies to different types of information. Sensitive information such as social security number, medical history, geolocation, and financial accounts can be more easily exploited for identity theft, blackmail, stalking and other crimes, which calls for a higher level of protection against unauthorized use. However, consumers are likely more wary about the potential danger when they consider transactions involving these sensitive information. Also, some seemingly non-sensitive information -- say zip code, demographics, and shopping habits -- can be put together to yield one's personal identity and other sensitive information. Taking all these into account, the definition of sensitive information is likely evolving as the technology develops, and the best tools for protecting sensitive information are likely different from the best tools for less sensitive information.[43]

In summary, we discuss the five tools as potential instruments to address information asymmetry, negative externality, and commitment problem. None of them directly addresses the potential market failure due to positive externality, though a reduction of the overall data risk may encourage consumers to share data (in a secure way) and alleviate their incentive to free ride on each other. Nor do they address product and pricing customization due to firms' better ability to classify consumers. We argue that this last one is not in itself an issue of market failure, and entails a policy tradeoff between market efficiency and the distributional effect of price discrimination.

Finally, it is important to emphasize that this note adopts an economic approach to identify potential market failures and policy tools that may address these failures. In practice, policy makers must evaluate whether a market failure *does* exist, to what extent market mechanisms cannot address it adequately, how long the failure has persisted and will continue in the future, and why the policy tools chosen to address the failure is more effective than existing market mechanisms. These benefit-cost analyses are likely context specific and evidence demanding.

[43] For similar reasons, the US Congress has enacted the Children's Online Privacy Protection Act (COPPA) for the collection and use of child information and the Gramm-Leach-Bliley Act (GLBA) for financial information. The FTC has advocated for different treatments of sensitive and non-sensitive information. (FTC 2012, FTC 2016b).

References

Acquisti, Alessandro; Curtis Taylor and Liad Wagman (2016) "The Economics of Privacy", Journal of Economic Literature, 54(2), 442-492.

Akerlof, George (1970) "The Market for 'Lemons': Qualitative Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, Vol. 84 (1970), pp. 488-500.

Armstrong, Mark and John Vickers (1991) "Welfare Effects of Price Discrimination by a Regulated Monopolist" *The RAND Journal of Economics*, Vol. 22, No. 4 (Winter, 1991), pp. 571-580.

Baye, Michael; John Morgan; Patrick Scholten (2005) "Information, Search, and Price Dispersion" *Handbook on Economics and Information Systems* (Elsevier, T. Hendershott, ed.), accessed at http://faculty.haas.berkeley.edu/rjmorgan/Information%20Search%20and%20Price%20Dispersion.pdf.

Brown, Jennifer; Tanjim Hossain and John Morgan (2010) "Shrouded Attributes and Information Suppression: Evidence from the Field", *Quarterly Journal of Economics*, Vol. 125 (2): 859-876.

Butler, Declan (2013) "When Google got flu wrong", *Nature*, Volume 494, February 14, 2013, pages 155-156.

Conlisk, John (1996) "Why bounded rationality?" *Journal of Economic Literature, Volume* 34 Issue 2, pages 669-700.

Cranor and McDonald (2008): "The Cost of Reading Privacy Policies" *I/S: A Journal of Law and Policy for the Information Society* 2008 Privacy Year in Review issue, http://www.is-journal.org/.

Dugas, Andrewa Freyer, et al. (2012): "Google Flu Trends: Correlation with Emergency Department of Influenza Rates and Crowding Metrics" *Clinical Infectious Diseases* 54(4): 463-69.

Ellison, Glenn and Sara F. Ellison (2009), "Search, Obfuscation, and Price Elasticities on the Internet," *Econometrica* 77(2), 427-452.

Fishman, M. & Hagerty, K. (1997), "Mandatory disclosure," in P. Newman, ed., *The New Palgrave Dictionary of Economics and the Law*, Macmillan Press.

FTC (2012): Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policy Makers, March 2012.

FTC (2014): Data Brokers: A Call for Transparency and Accountability, May 2014.

FTC (2016a): Big Data: A Tool for Inclusion or Exclusion?, Jan 2016.

FTC (2016b): FTC Staff Comment to the Federal Communications Commission: In the Matter of Protecting the Privacy of Customers of Broadband and Other Telecommunications Services, May 2016.

Gabaix, Xavier and David Laibson (2006), "Shrouded Attributes, Consumer Myopia and Information Suppression in Competitive Markets," *Quarterly Journal of Economics* 121(2) 505-540.

Grossman, S. "The informational Role of Warranties and Private Disclosure about Product Quality", *Journal of Law and Economics*, Vol. 24 (1981), pp. 461-489.

Harrell, Erika (2015): "Victims of Identity Theft: 2014", Bureau of Justice Statistics, NCJ 248991, accessed at https://www.bjs.gov/content/pub/pdf/vit14.pdf.

Hart, O.D. And Moore, J. "Incomplete Contracts and Renegotiation." *Econometrica*, Vol. 56 (1988), pp. 755-785.

Hausman, J. and Mackie-Mason, J. "Price Discrimination and Patent Policy." *Rand Journal of Economics*, Vol. 19 (1988), pp. 253–265.

Holmes, T. (1989), "The Effect of Third-Degree Price Discrimination in Oligopoly." *American Economic Review*, Vol. 79 (1989), pp. 244–250.

Jin, G. & Leslie, P. (2003), "The effect of information on product quality: Evidence from restuarant hygiene grade", *Quarterly Journal of Economics* Volume 118 Issue 2, pages 409–451.

Jovanovic, B. "Truthful Disclosure of Information", *Bell Journal of Economics*, Vol. 13 (1982), pp. 36-44.

Kahneman, Daniel (2011), *Thinking, Fast and Slow*, published by Farrar, Straus and Giroux.

Laffont, Jean-Jacques, and Jean Tirole. "The politics of government decision-making: A theory of regulatory capture." The Quarterly Journal of Economics 106, no. 4 (1991): 1089-1127.

Lazer, David; Ryan Kennedy; Gary King; and Alessandro Vespignani (2014) "The Parable of Google Flu: Traps in Big Data Analysis." *Science*, 14 March, 343: 1203-1205. Copy at http://j.mp/1ii4ETo.

Lecuyer, Mathias; Riley Spahn; Yannis Spiliopoulos; Augustin Chaintreau; Roxana Geambasu; and Daniel Hsu (2015): "Sunlight: Fine-grained Targeting Detection at Scale with Statistical Confidence", *Proceedings of the* ACM *Conference on Computer and Communications Security (CCS)*, Denver Colorado, October *2015*.

Leland, Hayne E. (1979) "Quacks, Lemons and Licensing: A Theory of Minimum Quality Standards" *Journal of Political Economy*, Volume 97, pages 1328-1346.

Mathios, Alan (2000) "The impact of mandatory disclosure laws on product choices", *Journal of Law and Economics* Volume 43 Issue 2, pages 651–677.

Mikhed, Vyacheslav and Michael Vogan (2017) "How Data Breaches Affect Consumer Credit", Federal Reserve Bank of Philadelphia Working Paper No. 17-06.

Milgrom, P.R. "Good News and Bad News: Representation Theorems and Applications", *The Bell Journal of Economics*, Vol. 12 (1981), pp. 380-391.

Nelson, P. (1974), "Advertising as Information," *Journal of Political Economy*, 82, 729-54.

O'Brien, Daniel P. and Doug Smith (2014) "Privacy in Online Markets:
A Welfare Analysis of Demand Rotations", *FTC Working Paper* #323, July 2014.

Ponemon Institute (2016): "2016 Cost of Data Breach Study: Global Analysis", accessed at https://public.dhe.ibm.com/common/ssi/ecm/se/en/sel03094wwen/SEL03094WWEN.PDF.

Posner, Richard "The Economics of Privacy", *The American Economic Review* Volume 71, Issue 2, Papers and Proceedings of the Ninety-Third Annual Meeting of the American Economic Association (May 1981), pages 405-409.

Rainie, Lee and Maeve Duggan (2016) "Privacy and Information Sharing", accessed at http://www.pewinternet.org/files/2016/01/PI_2016.01.14_Privacy-and-Info-Sharing_FINAL.pdf.

Shapiro, Carl (1983) "Premiums for High Quality Products as Returns to Reputations" *The Quarterly Journal of Economics*, Vol. 98, No. 4 (Nov., 1983), pp. 659-680.

Smith, Aaron (2014) "Half of online Americans don't know what a privacy policy is", accessed at http://www.pewresearch.org/fact-tank/2014/12/04/half-of-americans-dont-know-what-a-privacy-policy-is/.

Stigler, George (1961) "The Economics of Information", *The Journal of Political Economy*, Volume 69, Issue 3 (June 1961), pages 213-225.

Stigler, George (1980) "An Introduction to Privacy in Economics and Politics", *The Journal of Legal Studies* Volume 9, Issue 4 (Dec 1980), pages 623-644.

Thales (2017) “2017 Thales Data Threat Report: Trends in Encryption and Data Security”, accessed at https://www.thales-esecurity.com/company/press/news/2017/january/2017-thales-data-threat-report-security-spending-decisions-leave-sensitive-data-vulnerable.

Tirole, Jean (1988), *The Theory of Industrial Organization*, MIT Press, ISBN: 9780262200714.

Turow, Joseph (2003): “Americans and Online Privacy: The System is Broken.” Report of the Annenberg Public Policy Center, June 2003.

Varian, H.R. “Price Discrimination and Social Welfare.” *American Economic Review*, Vol. 75 (1985), pp. 870–875.