Towards a Resolution of the Privacy Paradox

Kristof Madarasz (LSE) and Marek Pycia (U Zurich)*

Abstract

This paper provides an explanation of the so-called privacy paradox and describes a more general informational 'irrelevance' result. We show that in a large class of imperfect information dynamic games between the buyer, the seller, and privacy platforms, the buyer chooses not to bear any direct cost of protecting her privacy even if leakage of her information affects the prices she faces and hence her surplus from trade. More generally, we show that the informed party's choice of privacy (mode of communication) is driven solely by the direct cost of talk rather than by the information such talk conveys: choosing between different privacy options, the buyer always chooses a cheapest option irrespective of its and its alternatives' informational characteristics.

1 Introduction

George Stigler (1980) opens his piece on the economics of privacy by claiming that "the enormous increase in the interest in privacy in our society is evident in the public press and in the statute books." He continues that "In some respects this interest in privacy is paradoxical, for the average citizen has more privacy—more areas of his life in which his behavior is not known by his fellows—than ever before." While the former claim still holds, the latter is unlikely to be true. In fact, currently there is an extensive discussion of

*This draft February 2022. First presentation March 2016; first posted draft February 2020. For their comments, we would like to thank Dilip Abreu, Elisabetta Iossa, Stephen Morris, Rakesh Vohra, and the audiences at Spring 2016 Job Talk Lunch Seminar at LSE and at doctoral courses at UCLA and U Zurich. We first drafted this paper while visiting the William S. Dietrich II Economic Theory Center at Princeton University, and we would like to thank the Center for its hospitality. Keywords: Privacy Protection, Tracking, Defaults, Dynamic Pricing.

a different paradox. The so-called privacy paradox refers to the fact that research on online behavior has revealed *discrepancies* between user attitude and their actual behavior: While users claim to be very concerned about their privacy, they nevertheless undertake very little to protect their personal data (see e.g., Barth and deJong 2017 or Kokolakis 2017 for reviews).

Indeed, the presence of unverifiable private information remains central to the interaction between firms and consumers. With the rise of the internet and traceable online search, however, the ease of recording and processing data about people has increased substantially. Many business models in the economy are based on collecting, storing, and processing such data about observable individual behavior. While buyers preferences are the buyers' soft information, such activities, in principle, allow firms to uncover information about otherwise unobservable and unverifiable personal preferences from patterns of behavior. It is then no surprise that firms value such information and engage in costly effort to obtain it (e.g., booking.com purchases from google). Information about preferences decreases the information rent that a potential seller needs to offer to a buyer when selling, which, in turn, it increases the seller's expected profit and in many contexts such rise in profits is likely to also lead to a decrease consumer surplus.[1]

People are generally aware of such activities. Furthermore, they appear concerned about firms' investments in technologies that serve the purpose to obtain, store, and interpret information about them. For example, Turow et al. (2009) report, based on a representative survey, that "contrary to what many marketers claim, most adult Americans (66%) do not want marketers to tailor advertisements to their interests. Moreover, when Americans are informed of three common ways that marketers gather data about people in order to tailor ads, even higher percentages - between 73% and 86% - say they would not want such advertising." They also report that 63% believe advertisers should be required by law to immediately delete information about their internet activity.

Despite being concerned about the loss of their privacy, however, consumers appear to do little to protect the privacy of their behavior, e.g., conceal their behavioral patterns, pay extra for apps or products that do not track or verify them, request access to various aspects of their data. For example, Rose (2005) finds that although most survey respondents reported that

[1] For example, Mikians et al (2012) suggest that considerable price differences, 10-30%, exist for identical products based on a variety of online characteristics, such as location or browser configuration.

they were concerned about their privacy, only 47 percent of them expressed a willingness to pay any amount to ensure the privacy of their information. Similarly, although the evidence is at times harder to interpret, others, e.g., Tsai et al. (2011) and Beresford et al. (2012) also provide findings consistent with the general consensus that people appear to be willing to pay very little extra to use platforms that provide a greater level of privacy.

To explain this apparent paradox, various authors have proposed that lack of awareness, or some form of bounded rationality, psychological factors such as a preference for immediate gratification or miscalibration of probabilities are at play, e.g., Acquisti and Grossklags (2005), Acquisiti, John, and Loewenstein (2013). While such behavioral factors are likely important, attempts to theoretically explain and practically solve this puzzle are still scarce and the subject may deserves more research attention. Indeed, the above cited surveys also suggest that further research in this area is likely to be key.

In this paper we emphasize another channel that imply that a buyer's willingness to pay extra for platforms that limit the collection of their data, thus allowing the seller to learn about their private and unverifiable preferences, may be very small. In particular, we describe a simple, but robust reason that might help explain this phenomenon given the underlying strategic interaction.

We first consider the classic problem where an uninformed seller faces a private-informed buyer over time. The seller can invest in increasing the precision of his/her information about the buyer over time. The buyer can also invest in decreasing the precision of the seller's information. Despite the fact that buyer has substantial information rents to collect thus suffers a great loss of such rents when her privacy is lost, we show that the buyer is never willing to invest in a greater protection of his privacy and, as long as such protection is costly, does not limit the seller's learning process.

More generally, we establish that a buyer in any perfect Bayesian equilibrium of a wide class of games ignores privacy considerations. Suppose that a consumer, who is privately informed about her preferences, can choose between different experiments, privacy platforms. Each privacy platform is associated with some signal structure that is transmitted to the seller. The information that is revealed by the buyer's type may depend on the buyer's type in an arbitrary way provided it never rules out a type with absolute certainty. Each experiment has some cost associated with it that the buyer has to incur. For example, protecting privacy may be more costly, then not protecting it at all. Such costs may also depend on the buyer's type, but there is a common default or least costly experiment. We show that all buyer types pool and choose the least costly experiment irrespective of its information content. In other words, it does not matter whether this platform offers full revelation of the buyer's preferences, or provides information about some types but not about a set of other types, all buyer types choose the same platform. The buyer always sticks to the default privacy platform irrespective of its privacy characteristics or that of the available alternatives.

In Section 3, we then endogenize the choice and the pricing of privacy platforms by a platform provider. Suppose first that the platform provider is a monopolist who can choose from an arbitrary set of technologically feasible platforms. The monopolist can decide which platforms to offer and at what price each. The platform provider also contracts with the seller and agrees on some profit-sharing, that is, on a contract where in exchange for the information provided to the seller, the seller's transfer back to the platform provider is increasing in the boost in the seller's profit given the information provided. Finally, a buyer then decides which platform to choose, may decide to shop 'offline directly from the seller, or exit the market. In this context we show that in equilibrium the platform provider offers platforms that maximizes the value of the information passed to the seller.

Finally, maintaining the same setting we also allow for competition between platform providers. In equilibrium, again, the platform providers offer platforms that maximize the value of the information passed to the seller.

In Section 4 we generalize our results by considering a large class of dynamic games between the buyer, the seller, and the privacy platforms. We assume that at each of her decision nodes, the buyer either takes a privacy or a purchase decision. At each purchase node, she decides whether or not to buy products offered by the sellers at prices determined by the seller prior to the buyer reaching this node. At each privacy node, the buyer makes an observable choice between different experiments where an experiment affects the information that is subsequently revealed to the other parties. Each experiment has some associated flow cost. This privacy choice of the buyer does not affect the subsequent game tree, e.g., the objects that sellers can offer or the prices he can set for their objects, only the information that the other parties may obtain in the continuation game. The seller has pricing nodes at which he can set prices, which remain in effect, until the seller reaches a potentially new pricing node, and contracting nodes with the platforms that again do not affect the game tree but affect the Nature's move, e.g., nature of the experiments, in the continuation game. Despite the generality of the setup, we show that in any perfect equilibrium of any game in this large class, at each of her privacy nodes, each buyer type always chooses an experiment with the lowest cost at that node irrespective of its information content. In other words, what affects the information that the seller gets, and then accordingly the profit he makes, is not the value (harm) of such information revelation to the buyer, but simply the cost of different privacy experiments.

Our paper then finds that in monopolistic markets buyers lack the proper equilibrium incentives of protecting their privacy as long as such protection is not the directly cheapest option they can pick. At the same time, parties that benefit from acquiring costly information have an incentives to do so provided competition does not fully restrict the ability of the latter parties to extract surplus from the former. This sharp asymmetry then has various implications. It may shed light on how the widespread ability to track consumers may have very significant impact on economic outcomes and the distribution of gains from trade even when, in principal, consumers seemingly have access to cheap methods of protecting such informational rents. In turn, policies aimed at consumer protection that do not take this asymmetry into account may be completely ineffective. Furthermore, policies that directly regulate information gathering, or information sale, may be needed to safeguard consumer welfare. We discuss such policy implications in the Conclusion.

Related Literature Our setup relates to various strands in the literature. It relates to classic mechanism design in that a privately informed party with non-verifiable information chooses between different alternatives and that choice may help the uninformed party learn her type. In our setting these alternatives are different information revelation technologies offered at potentially different – and potentially type-dependent — benefits / costs. Here, we show a general pooling result. It also relates to the literature on information design, e.g., Kamenica and Gentzkow (2011), Bergemann and Morris (2019), where the sender commits ex ante to an experiment to persuade a receiver. We relate to this literature in that the platform designer offers a set of platforms experiments. We also relate to the voluntary disclosure of verifiable information literature, e.g., Grossman and Hart (1980) and Milgrom (1981) who show an unraveling result that under certain assumptions (including: costless disclosure and verifiable information) all types disclose their private information. Among others, Jovanovich (1982) considers costly disclosure and shows that if disclosure is costly, then in some settings high quality sellers disclose and low quality sellers do not. Instead, in our setting, the informed party has no verifiable information and instead can choose from an arbitrary set of partial and dynamic revelation technologies.

In a simple setting of disclosing all or no verifiable information, Acquisti et al. (2016) note the relevance of this classic full disclosure for consumer privacy. In independent and simultaneous work, Ali et al. (2021) consider disclosure choices of a buyer with verifiable information to a monopolist seller. In their setting, disclosure is still costless, but they allow for a richer space of signals. They show that the buyer may choose and benefit from partial disclosure of her information. We differ from their setups in that we allow for arbitrary information revelation technologies and, crucially, that in our setting different privacy platforms can differ in how costly they are for consumers.

In a manner that also builds on the linkage between the choices of the different types of the informed party, we then instead show generally that in the context of monopolistic markets the informational aspects of such privacy technologies are irrelevant for the informed party's choice as long as different privacy options may carry even minimally different direct costs or benefits. Such cost differences, no matter how arbitrarily small, then imply that all buyer types, irrespective of the signal technologies available to them, or the dynamic nature of the information revelation, pool on the cheapest privacy option. In other words, what matters is the direct cost (or benefit) of different privacy options, such as the price for or extra inconvenience in setting up limited tracking, or the different entertainment values of search sites, rather than the information consequence of different platforms and thus the consequences they entail for price discrimination.

2 The Value of Protecting Privacy: Examples

Consider the following classic monopoly problem. The seller owns an object that he produces at a normalized cost of zero. The buyer's privately knows her valuation which is denoted by θ and assumed to be positive; its cdf $F(\theta)$ is strictly increasing on some positive support $[\underline{\theta}, \overline{\theta}]$. In each period $t = 1....$ the seller makes a price offer that the buyer then accepts or rejects. If the buyer accepts, she obtains her valuation net the price and the seller receives the price and the game ends. If the buyer rejects, then the game continues. The seller and the buyer discount the future at same fixed rate. We employ the standard notion of perfect Bayesian Nash equilibrium (equilibrium or PBE henceforth) as our solution concept.

We study the privacy of the buyer's information in this classic problem. We assume that the type θ, e.g., describing the buyer's preferences, cannot be credibly revealed to the seller (it is not hard information), but there are some signal processes that over time provide information about θ; we model these processes below. We are interested in the question whether the buyer would choose to influence these signal processes in a way that would protect the buyer's privacy. We hence assume that at the interim stage, that is, before the seller makes any price offer, but once the buyer has private information, the buyer can invest into information (privacy) protection. As our focus is on the buyer's privacy protecting decisions, we initially simplify the seller's information collecting decisions and assume they will be costless to the seller or to the intermediaries (platform providers) we study in subsequent sections. Let us note however, that unlike the buyer, the seller (and the platform providers) would be willing to pay for information collection. Finally, we need no assumptions as to whether the seller can commit to a price sequence or if his pricing decisions need to satisfy sequential rationality as our results hold in both cases.[2]

2.1 Example

As an example, suppose that at the beginning of each period t there is an independent leakage probability α such that the buyer's valuation leaks to the seller, that is, the seller privately learns θ with this probability. For example, the seller, or the operator of an online platform on her behalf, may be able to figure out the buyer's preferences for the seller's product from observing the buyer's activity online. We can think of this leakage probability as one publicly set by the seller at some cost initially. Since the result does not depend on the exact values of α we do not model this stage explicitly.

Suppose that, having learned her type, but before the seller makes any offers, the buyer can invest $m \in 0 \cup [a, b]$ to protect his privacy, where $b > a > 0$. Protecting his privacy means that this investment decreases the arrival

[2]Note that in the absence of the leakage possibility, Myerson (1983) shows that, for the seller who can commit to a strategy, the optimal selling mechanism is for the seller to maintain the same posted price in all periods. Such an optimal mechanism then also characterizes the information that the seller can optimally elicit from the privately-informed buyer in an incentive-compatible manner. The seller's profit is constrained by the information rent that the seller needs to offer to the buyer in terms of consumer surplus to be able to sell. Gul, Sonnenschein, and Wilson (1986) show that, for the seller who can commit to a strategy and under the so called "gap" assumptions on the values, the price will quickly decrease to the seller's cost.

probability; i.e., $\alpha(m)$ is decreasing in m. Such protection is costly and the cost function $c(m)$ is such that $c' > 0$ on $[a, b]$ and there exists $\varepsilon > 0$ such that $c(a) > \varepsilon$. We interpret $m = 0$ as the no investment decision, leaving the leakage probability unaffected, and normalize $c(0) = 0$. In other words, we assume that there is an (arbitrarily small, but positive) cost to pay for protecting privacy. We also assume that the event $m > 0$ is observable by the seller.

Proposition 1 *The buyer never invests in protecting her privacy.*

This and all subsequent proofs are provided in the Appendix. The logic of the proof is as follows. Note that a pooling equilibrium with no investment, that is when no buyer type protects his privacy, exists. This is true because given such expectations if any buyer type were to invest, the seller can always attribute the investment coming from the type with the highest willingness to pay. In turn, the seller would not drop the price below that level but only once leakage occurs which, however, still leaves the deviating type with no surplus. In contrast, a separating equilibrium cannot exist because if it did, then one can always consider the lowest type, or the infimum of types in case the set of types who invest is not closed, who decides to invest some positive amount. Such a type, however, can never gain from investment since it is also the infimum of the seller's posterior conditional on the buyer's equilibrium investment choice. This type would then benefit by deviating and not investing and thereby not protecting his privacy such benefits include a lower cost of investment and a (weakly) lower sales price.

2.2 More general information revelation processes

The above proposition does not hinge on the details of the example and remain valid for general partial dynamic information revelation. Specifically, suppose that in each period the seller obtains a signal $s_t : [\underline{\theta}, \overline{\theta}, \{\widehat{s}_k\}_{k=1}^{t-1}] \to \Delta Z$ with probability α_t where Z is a finite realization space. We also allow the signal realization in period t to be a function of the past signal realizations, $\{\widehat{s}_k\}_{k=1}^{t-1}$, in case signals in the past have arrived. We assume now that the signal is no longer fully revealing and assume that each realization of s_t occurs with positive probability given each element of the domain, that is, observing a signal s_t per se, leaves the seller with a posterior that has full support. Of course in equilibrium, the seller's posterior in period t may rule out a positive measure of types, but such a feature of the seller's beliefs is a consequence of equilibrium behavior directly. Investment is costly just as before, but we can also let $\alpha_t(m)$ be an arbitrary function which is strictly positive on $[a, b]$.

Proposition 2 *The buyer never invests in protecting her privacy.*

The logic of the above result is the same as the one described before.

2.3 Privacy Platforms

Building on the logic of the above, suppose that the buyer can choose between N different privacy platforms (approaches, experiments) $\{s_1, ...s_N\}$. Each privacy platform, $s_j : \Theta \to \Delta(Z_1 \times Z_2 \times ....)$ is a function which assigns to each type a probability distribution over a set of, possibly correlated, signal realizations over time. We again assume that the each probability distribution has full support. Let $c(s, \theta)$ be the cost of *choosing* platform s for type θ. The privacy platform chosen is observed by the seller. We assume that there is a platform that is the cheapest for all types, that is, there exist $s^* \in S$ such that $c(s^*, \theta) < c(s', \theta)$ for all θ and s'. We can interpret this platform as the default as without loss of generality we can assume that if the buyer does nothing then he or she uses this platform (notice that this is just a normalization calling the choice of the cheapest platform as doing nothing). The motivation for this terminology is that defaults are typically less costly to choose than other options.

Theorem 1 *The buyer always chooses the cheapest privacy platform.*

Remark 1 *This result remains true—with no change in the proof—if we allow a richer set of privacy platforms in which having selected privacy platform s_j the agent can further choose a message (or an action) from a set M_j which is costless and not observed by the seller. Thus the privacy platform can be defined as a mapping $s_j : \Theta \times M \to \Delta(Z_1 \times Z_2 \times ....)$ which assigns to each type and message pair a probability distribution over a set of, possibly correlated, signal realizations over time. We can interpret these privacy platforms as contracts which tell the buyer that signals z_1, z_2 etc. will be communicated to the seller at respective times 1, 2, etc.*

3 Privacy-Paradox in Dynamic Markets

We conclude the analysis with showing that our insights are valid beyond the specific models discussed so far. We look at finite and bounded length extensive form games with perfect recall and imperfect information.[3] The set of players is $\mathcal{N}$ and we also allow for moves by Nature. We are particularly interested in the behavior of special player whom we call the buyer. The buyer is privately informed about their valuation for each subset of objects X sold by the seller: at the beginning of the game, nature draws the type of the buyer, $\theta \in \Theta \subseteq \mathbb{R}$.[4] The buyer's utility is determined as follows. By the end of the game, the buyer has bought a set of objects X at their respective prices and made a sequence of privacy choices y. The utility of the buyer is then:

$$u = c(y) + V(X, \theta) - \sum_{x \in X} (p(x))$$

where:

- $c(y)$ is the cost of the sequence of privacy choices y, which we assume is additively separable in "flow costs" incurred at each privacy decision node: at any privacy node h, let $C(h, a, \theta)$ be the flow privacy cost of decision a at node h for type θ, which is such that the action with the lowest cost at h is unique and is type-independent;

- $V(X, \theta)$ is the value of the purchased bundle X. We assume that $V(X, \theta)$ is continuous in θ and $V(X \cup X', \theta) - V(X', \theta)$ is strictly increasing in it for any disjoint X and X' such that X is nonempty. For instance, we can have $V(X, \theta) = \sum_{x \in X} v(x, \theta)$, where $v(x, \theta)$ is the value of object x for type θ of the buyer that is continuous and strictly increasing in θ. We assume that the buyer buys at most one unit of each object.[5]

- $p(x)$ is the price paid for object x.

The set of players also contains an agent we call the seller, who is different from the buyer. We use the letter j to refer to the seller. Let $\bar{X}_j$ be the set of objects that can be produced by seller j. We assume that, for all types $\theta \in \Theta$

[3] An analogue of our results obtains for infinite games with discounting. Note that while we do not explicitly model discounting, as long as the agents are risk averse we can map games with discounting into our framework because our treatment of nature moves is very general; in particular the class of games we study contains games in which at the end of each period nature terminates the game with probability $1 - \delta$ for some $\delta \in (0, 1)$ and each player's decisions in such games are the same as in the analogous games with exogenous termination and with discount factor δ common to all players.

[4] Assuming that the type space Θ is finite would simplify the exposition; we do not rely on this assumption.

[5] This single-unit demand assumption is made for notational simplicity only. In particular, note that we allow different objects to be units of the same good (variety). Furthermore, the additive aspect of the utility across objects.

and objects $x \in \bar{X}_j$, the set of prices that seller j can set is $P = \mathbb{R}_+$. By the end of the game seller j has sold some objects X_j at price $p(x)$ for a given $x \in X_j$ and made a sequence of contracting decisions with platforms z_j. The utility of seller j is then:

$$u_j = \pi\left(z_j, X_j, p(x)_{x \in X_j}\right)$$

and we assume that for each set of decisions z_j and sales X_j, the utility of the seller is strictly and continuously increasing in each price $p(x)$, $x \in X_j$. For instance, this assumption is satisfied if

$$u_j = c(z_j) + \sum_{x \in X_j} (p(x) - c(x)).$$

where $c(z_j)$ is the cost of contracting decisions z_j and $c(x)$ is the cost of producing good x.

The set of players might also contain further players, e.g., in applications those additional players might be called platforms whose decisions affect the buyer's privacy nodes and seller's profits. The utility of these additional players is determined by their moves, buyer's privacy choices, and sellers contracting decisions.

We restrict attention to the following class of games. The decision nodes at which the buyer moves are partitioned into two types:

- Privacy-decision nodes. At each privacy node there is a finite set of actions and each action leads to the same continuation game tree. The privacy choices affect only
  - buyer's payoffs: the payoffs at corresponding terminal nodes following any two different choices at the node differ by a constant; the constant depends on the two choices and it is the flow cost of privacy decision mentioned earlier. We may interpret the payoff impact of the decision as reflecting both the cost of protecting privacy and the consumption value of privacy.
  - the probability distribution of actions that nature takes at the continuation nodes at which it moves: The probabilities are determined by both the buyer's choice and buyer's type and hence we interpret the nature's choices as signals that others might perceive about the buyer's type. We assume that at each privacy node,

the support of subsequent nature moves does not depend on the buyer's type; we refer to this as *the full support assumption.* Note also that our framework allows for multiple classic cheap talk messages as well, that is, messages that are both free and uncorrelated with the buyer's type, or potentially costly pooling messages that are uncorrelated with the buyer's type.

- Purchase nodes. At each purchase node there is a finite set of actions and each action at the purchase node leads to the same continuation game tree with the same distributions of nature's moves. Each decision at the purchase node is a purchase of an object or a set of objects.

We assume that the buyer's privacy choices are observable by other players, while her purchase is observable only by the seller.

The decision nodes at which the seller j moves are also partitioned into two types:

- Pricing decision nodes. At each pricing node, the seller is able to offer any object at any price. Thus, each action at the price-setting node sets prices $p(x)$ for objects $x \in \bar{X}_j$ sold by the seller and any mapping from $\bar{X}_j$ to P is allowed. The seller can change these prices at their subsequent moves (no commitment assumption). The seller's pricing decisions are observed by the buyer and when a buyer i is offered object x the associated price is the last price set by the seller for it. The decisions at a purchase node lead to the same continuation game tree not affecting directly the distributions of nature's moves and thus to the same payoffs of all parties except that directly attributable to the purchase (the revenue for the seller and the consumption utility for the buyer).

- Contracting nodes. At each contracting node there is a finite set of actions and each action leads to the same continuation game tree except that:
  - the contracting choice at a node may affect the payoffs of all parties, but only in a way that is additively separable from choices at other nodes
  - the contracting choice at a node may affect the probability distribution of actions that Nature takes at the continuation nodes at which it moves; we assume that at each contracting node, the

support of subsequent nature moves is the same after each choice (the second part of *the full support assumption*).

All decisions nodes of players who are neither buyer nor seller nor Nature are called contracting nodes.

In the sequel we restrict attention to single object $\overline{X}_j = x$. We also restrict attention to PBE satisfying the following *seller's payoff monotonicity condition*: if at pricing node h the seller asks p for object x, there is a positive probability the buyer can buy at a purchase node at which this price is current, at all subsequent purchase nodes at which the price remains current the buyer buys the object, and the same obtains if at pricing node h the seller instead asks for price $p' > p$ for x, then there is a strategy of the seller following asking p' at h such that the seller's expected payoff following p' at h is strictly higher than the seller's expected payoff following p at h. This monotonicity condition is satisfied for all PBEs under either of the assumptions on the environment in the following lemma:[6]

Lemma 2 *Suppose that either (i) the buyer and the seller are the only players in the game; or (ii) the price charged by the seller is only observed at the purchase node (and hence only by the buyer) and only the current price is observed. Then, all PBEs satisfy the seller's payoff monotonicity condition.*

Proof of Lemma. Consider a PBE in which if at pricing node h the seller asks for p for object x, there is a positive probability the buyer can buy at a purchase at which this price is current, and at all subsequent purchase nodes at which the price remains current the buyer buys the object, and the same obtains if at pricing node h the seller asks for price $p' > p$ for x. Consider the following strategy of the seller following asking p' at h: at all contracting nodes the seller chooses the same actions as at corresponding nodes following asking p at h and at pricing nodes following p' the seller chooses the same actions as at corresponding nodes following p. The buyer also chooses the same actions at the corresponding decision nodes following p and p'. Indeed, this is the case if the buyer does not observe whether the seller chose p and p'. If the buyer sees the price choice then the buyer has a purchase node at which these prices are current, and then by the assumption of the lemma, the buyer buys the object irrespective of whether the price is p or p'; following the purchase node the buyer only makes privacy choices and our assumptions

[6]This restriction on PBEs is also implied in the environment studied in the next section in which the seller and platforms share the profits from sale.

guarantee that the buyer always chooses the least expensive privacy choice. Because only the seller and the buyer might observe the price and they choose same actions (and with same probabilities), the other players also choose the same actions in their corresponding decision nodes following p and p'. Thus, following both pricing choices at h, all players choose exactly the same actions at the corresponding decision nodes and the only difference in the seller's utility between charging p and p' is the increased revenue when the good is sold at p'. QED

Theorem 3 *In any perfect Bayesian equilibrium, at any privacy node h on the equilibrium path, all buyer types choose the action which minimizes the flow cost of the privacy choice $C(h,a)$ across all actions a at h.*

Proof of Theorem. Consider the event that in the continuation game following node h there is a pricing node at which good x is priced and, with positive probability, following this pricing node a purchase node at which the buyer can purchase x. If this event has probability zero, then only the flow costs of privacy choices affect the buyer's utility and the theorem obtains. Hence, in the sequel we consider the case when the above event has positive probability.

By way of contradiction, suppose that there is a pure-strategy PBE that fails the properties from the theorem and thus there are on-path privacy nodes h that fail the property from the theorem. Because the game has finite length, there is a node h such that at no privacy node following it along the path of play the property from the theorem is violated. Consider such last violation and let type $\underline{\theta}(h)$ be the essential infimum of types for which h is on path.[7] As before, type $\underline{\theta}(h)$ never obtains a surplus from buying anything in the continuation of the equilibrium path, thus this type—if present at h—would make a lowest cost privacy choice at privacy node h. By the proof's assumption, some types make a non-lowest-cost privacy choice a', with positive probability. Let $\theta(h)$ be the infimum of these types; let $\Delta > 0$ be the cost difference between a' and the lowest cost privacy choice.[8]

[7] The essential infimum is the highest type such that the probabiliy that the type is lower is zero; the probability here is the conditional probability distribution over types present at node h on path of the PBE we study.

[8] There is such Δ because the lowest flow costs is type-independent and there is a finite set of privacy choices at each node.

The rest of the argument builds on the following:

Lemma 4 *Consider the above perfect Bayesian equilibrium and the continuation game following h. At each on-path node h' at which the seller sets price $p(x)$ for object x and there is a positive probability set of on-path game paths along which the buyer is able to buy x at price $p(x)$ on the game path after h' and before the next pricing node of the seller, then $p(x) \geq v(x, \underline{\theta}(h), h)$ where $\underline{\theta}(h)$ is the essential infimum of buyer types present at node h on path of the game.*

Proof of Lemma. The argument is by backward induction. If there are pricing nodes that violate the claim, then, because of the finite length of the game, there is a node h' such that all pricing nodes in the continuation game (if any) satisfy the claim. At h' the price $p(x) < v(x, \underline{\theta}(h))$. Consider a deviation to price $p'(x) \in (p(x), v(x, \underline{\theta}(h))$. With positive probability the buyer can buy x at price $p'(x)$ and given the assumptions on utility all still present buyer types will buy x at this price because, by the inductive assumption, the price of x will be higher than the essential infimum of the marginal value, and the essential infimum can only increase along on-path game paths. Thus, the assumptions of Lemma 2 are satisfied and by the lemma the deviation would be profitable. QED

Thus, in the PBE in the continuation game following h, the sellers will not set bundle prices strictly below the marginal values of type $\theta(h)$ when it matters. In effect, by choosing a' type $\theta(h)$ would not obtain a surplus from trading in equilibrium and would prefer to deviate to the lowest cost experiment. Furthermore, as $\theta(h)$ is the infimum of types choosing a' for any $\epsilon \in (0, \Delta)$ there must be a type $\theta_\epsilon \in [\theta(h), \theta(h) + \epsilon]$ that chooses a'. Because of continuity of utilities, type θ_ϵ profit from trading can be made arbitrarily small by setting ϵ sufficiently small; this type would benefit by deviating to the lowest cost experiment. The presence of this deviation is a contradiction that proves that no type chose non-lowest-cost experiment and completes the proof of the theorem. QED

4 Design of privacy platforms

We now turn to an application of our previous results in the context of the design of privacy platforms. So far we left the emergence of the buyer's privacy options to be an exogenous aspect of the setup. What happens when they are designed by a privacy platform provider, such as a search website or other online intermediary? Our main insight carries over to such a richer environment. In particular, we show that the observability of buyer's choice of privacy platform emerges endogenously in equilibrium.

We start by considering a monopolistic designer of the privacy platforms and we address competing platform providers in the next section. The platform provider can offer privacy platforms from an arbitrary set P of technologically feasible privacy platforms. Each platform consists of both the signal generating process and a contract with the buyer that specifies what type of information might be passed on by the provider to the seller of the good the buyer may want to buy, cf. Section 2.3 above. We assume that the set of feasible platforms is finite; but this assumption can be relaxed as long as the topology of the problem ensures that the provider has well-defined optimal choice. We also assume that all platforms are equally costly to the provider; this assumption can be fully relaxed.[9]

After selecting any subset, the designer decides at what price to offer each to the buyer. In other words, we now allow the cost of each platform to be determined endogenously in equilibrium by a profit-maximizing platform provider. Finally, the platform providers communicate to the seller the buyer's choice and all other information about the buyer that is allowed by the contractual agreement with the consumer. After offering a subset of the technologically available platforms and corresponding prices, the buyer picks whichever she prefers.

We allow the buyer to choose to shop 'offline' directly from the profit-maximizing seller, i.e., not use the platform. This option then allows for a, in principle, type-dependent outside option. We may also assume that the buyer can also decide to choose neither a platform from the provider, nor purchase offline, i.e., to exit the market. We furthermore assume that being on the platform offers some, potentially arbitrary small, pure benefit $b > 0$ which we can interpret as e.g., search or entertainment value of the platform.

The model so far did not specify the bargaining between the platform provider and the seller. To close the model let us assume that the platform provider makes a take-it-or-leave-it offer to the seller and then the seller responds, both events happening before the provider makes the offer to the buyer. The timing details turn out not to matter as long as the bargaining

[9] After we relax this assumption, we need to adjust the proposition below so that the platform provider chooses the platform that maximizes the provider's share of the seller profit net of provider's cost of operating the platform.

between the platform provider and the seller occurs while both parties are still uninformed. In particular, we could also assume that the provider makes a take-it-or-leave offer to the seller simultaneously with the offer to the buyer. The above bargaining procedure between the platform provider and the seller is also not crucial: the main insight goes through for all bargaining procedures in which the expected transfer from the seller to the platform is strictly monotonic in the seller's expected revenue, e.g. if the platform captures a fixed positive share of the seller's revenue. Below we assume without loss of generality that if a platform is never chosen by the buyer in equilibrium it is also not offered by the provider.

Proposition 3 *In equilibrium all platforms offered have a price of b, the set of platforms provided maximizes the seller's revenue among all feasible platform sets, and with probability* 1 *the buyer participates in one of the platforms provided.*

The above implies that when the full information extraction (no privacy) platform is feasible for the platform provider, then the following is true:

Corollary 1 *Suppose that the platform provider has a full revelation platform available. Then that is the platform that will be provided at the price of b. In turn, the seller and the platform provider will extract all the surplus from the market and the equilibrium is essentially unique.*

4.1 Competition amongst Platform Providers

While so far we derived our insight—that the cheapest privacy platform is chosen—assuming a monopolistic seller and an exogenous set of privacy platform or a monopolistic designer of the platforms, the insight remains valid when two or more privacy-platform designers (tracking institutions) offer their services to buyers.

Consider the setting of Section 3 except that there are multiple platform providers competing for the buyers. As with the monopoly platform provider, the details of the bargaining between the platforms and the seller do not matter as long as the expected transfer from the seller to the platform is strictly monotonic in the seller's expected revenue, e.g. if the platforms capture a fixed positive share of the seller revenue from the buyers matched with the seller by the platform. We assume that each platform provider has access to the same set of feasible platforms, but otherwise allow asymmetric platform providers; thus different platforms might have different ability to negotiate with the seller.[10] The providers simultaneously offer menus of platforms to the buyers.[11] As before we allow transfers between the buyer and the platform providers; we refer to negative prices as subsidies.[12] The buyer sees the menus of each provider before choosing one of the platforms or choosing to shop offline or exit the market. Finally, the platform providers communicate to the seller the buyer's choice and all other information about the buyer that is allowed by the contractual agreement with the consumer.

Proposition 4 *In equilibrium, the platforms that maximize the seller's revenue are offered and, with probability 1, the buyer chooses to join one of these platforms.*

For instance, if the competing platform providers have access to the same set of feasible platforms and platform i captures the share λ_i of the seller's revenue, then each provider with highest λ_i chooses the platforms that maximize the seller's revenue and provides the subsidy equal to the product of the second highest λ_i times the expected revenue of the seller. Furthermore, with probability 1, the buyer chooses to join one of the platforms offered by providers with highest λ_i.

We thus see that the competition shifts some of the surplus towards the buyers while not affecting the insight that the privacy is not protected.

4.1.1 The role of subsidies

So far we allowed the platforms to subsidize the buyers. Suppose now that such subsidies (beyond the benefit of using platforms b) are not possible. We may want to impose this assumption e.g. if we are concerned that in practice subsidies may lead the buyers to collect the transfers—possibly multiple times, by posing as multiple buyers (shills)—without buying the product. No subsidy is then a simplifying assumption made in lieu of specifying a richer model.

In the presence of the no-subsidy assumption, multiple essentially different equilibria of the privacy game become possible. In each equilibrium, all platform providers offer their platforms for free, and, with probability 1, the buyer chooses to join one of the platforms. There is an equilibrium in which

[10] Allowing different platform providers access to different sets of feasible platforms does not change things substantially but its full analysis would require taking a less agnostic stance on the bargaining between the seller and platform providers.

[11] This assumption is to focus attention only. In particular, our results hold if the providers move sequentially.

[12] Below we separately consider the case when subsidies are not allowed.

the privacy is not protected and each platform provider chooses the platforms that maximize the seller's revenue but there are now also other equilibria, including an equilibrium in which the most privacy-protecting platforms are offered by providers and chosen by the buyers. The no-privacy equilibrium maximizes the joint surplus of the platform providers and as such is a focal one.[13] We formulate this insight for the case of symmetric platform providers.

Proposition 5 *Suppose that subsidies are not allowed. In the equilibrium that maximizes the joint surplus of the platform providers, each provider chooses the platforms that maximize the seller's revenue and offers it at price* 0. *Furthermore, with probability 1, the buyer chooses to join one of the platforms offered.*

4.2 Competition among sellers

In the presence of competition among sellers, all buyer's type choosing the lowest cost experiment remains a Perfect Bayesian Equilibrium as long as the resulting prices set by competing sellers are weakly increasing as the sellers' beliefs about buyer's value change in the first order stochastically dominant way. There might be however other equilibria.[14]

5 Regulation and Conclusion

Regulators are concerned that the loss of consumer privacy in the digital age allows firms to engage in pricing practices that both reduce consumer surplus from trade and lower the competitive pressure between firms, e.g., see the UK's Competition and Market Authority's 2021 report "Algorithms: How they can reduce competition and harm consumers."[15] A general policy recommended, both in Europe and in the United States, is forcing online platforms to offer consumers the option of better safeguarding their private information and the privacy of their online behavior; an example of which is GDPR passed by the European Union in 2016 (similar regulations have been proposed in, e.g., California). While GDPR has generated fine revenues from companies like Google or Facebook, survey evidence (from the UK, Ireland, Germany,

[13]As we prove in Section 4, the multiplicity is a by-product of the benefit b being the same on all platforms. The insight that buyers choose the cheapest (highest benefit) platform obtains in all equilibria irrespective of whether b is homeogenous or heterogenous.

[14]The uniqueness does obtain if the competition among sellers implies that their belief about buyer's value does not impact prices they charge, as in Bertrand competition.

[15]https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers/algorithms-how-they-can-reduce-competition-and-harm-consumers.

and the Netherlands) suggests that while people became familiar with the presence of GDPR, this has had little impact on how worried consumers remained about their online privacy.[16]. See also, Aridor et al. (2020) who find that while some consumers decided to opt out from being tracked, overall " trackability has *increased* [italics added] by 8 percent under GDPR," where trackability is defined as "the fraction of consumers whose identifier a website repeatedly observes in its data over some time period."

Our results imply a general divergence between the privacy attitudes and privacy choices of consumers. In our setting buyers will pool on the cheapest privacy platform (or the one with the greatest direct net benefit) irrespective of its informational content. This insight is then relevant in light of this great policy interest in regulating the ways firms need to ask for the consent of those whose information they collect, store and sell. Specifically, it suggests that if firms simply are required to give the option to consumers to protect their privacy this may be completely ineffective. Instead, direct regulation of the kind of information that can be collected may be a more effective way of safeguarding information that consumers would prefer to keep private. This might help inform such regulations as GDPR. By shifting the (cheapest for the buyer) default from 'opt out' to 'opt in,' or making it difficult to offer a somewhat more pleasant, e.g., faster access, consumer experience when choosing to share as opposed to conceal data, the regulation can shift the equilibrium from no privacy to the regulators' intended privacy — provided the platforms offer no subsidies. With subsidies, and assuming that the sellers are willing to subsidize opt-in sufficiently to cover the costs of an active choice, the old equilibrium (opt in) prevails with the buyers merely compensated for the active opt-in choice. Our logic also suggests that in some setting regulation that takes away the property rights of platform of the data collected, e.g., mandatory sharing of the data, may work since it breaks the link between the platforms' and the seller's payoffs.

More generally, our results contribute to the rationale of active regulation of privacy at the first place. This is in contrast with earlier accounts, such as Taylor (2004).[17] Specifically, Taylor (2004) finds that, in the presence of tracking technologies that allow sellers to infer consumers' preferences and sell such information to others who can then engage in price discrimination,

[16] e.g., https://www.surveymonkey.com/curiosity/g pr-one-year-out/

[17] It is also in contrast with implications of classic dynamic price discrimination given standard Coasian informational dynamics where the seller loses and the consumer gains when the seller tracks prior purchase decisions, e.g., Hart and Tirole (1988).

the usefulness of privacy regulatory protection depends on consumers' level of sophistication. Naive consumers do not anticipate a seller's ability to use any and every detail about their past interactions for price discrimination; consequently, in equilibrium, their surplus is captured by firms—unless privacy protection is enforced through regulation. Regulation, however, is not necessary if consumers are aware of how a company may use and sell their data and buyers can adapt their purchasing decisions accordingly, because it is in a company's best interest to protect customers' data (even if there is no specific regulation that forces it to do so). Our results instead imply that when consumers can express their privacy preferences, firms will greatly benefit from tracking and even if privacy protection remains a cheap option for consumers, e.g., ensured by a regulator, it leads to a potentially very significant reduction of consumer surplus.

References

[1] Acquisiti, A. and J. Grossklag (2005). "Privacy and Rationality." In: *Individual Decision Making IEEE. Sec Priv* 3: 24-30.

[2] Acquisiti, A., L. K. John, and G. Loewenstein (2013). What is Privacy Worth. *Journal of Legal Studies.*

[3] Acquisti, A., C. Taylor and L. Wagman (2016). The Economics of Privacy, Journal of Economics Literature.

[4] Ali, N., G. Lewis, and S. Vasserman (2021). "Voluntary Disclosure and Personalized Pricing." Working Paper.

[5] Aridor, G., Y-K Che, and Y. Salz (2020). "The Economic Consequences of Data Privacy Regulation: Empirical Evidence from GDPR" Working Paper

[6] Barth and DeJong (2017). "The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review." *Telematics and Informatics*, 1038-1058.

[7] Beresford, A., D. Kubler, and S. Preibusch (2012). "Unwillingness to Pay for Privacy: A Field Experiment." *Economics Letters* 117:25-27.

[8] Bergemann, D,, and S. Morris (2019). "Information Design: A Unified Perspective." *Journal of Economic Literature* 57(1): 44–95

[9] Board, S and J. Lu (2018). "Competitive Information Disclosure in Search Markets." *Journal of Political Economy* 126 (5), 1965-2010.

[10] Grossman, S., and O. Hart (1980). "Disclosure Laws and Takeover Bids." Journal of Finance 35 (2): 323-34.

[11] Gul, Faruk, Hugo Sonnenschein, and Robert Wilson. (1986). âFoundations of Dynamic Monopoly and the Coase Conjecture.â *Journal of Economic Theory* 39 (1): 155â90.

[12] Gutman, I. Kremer, and A. Skrzypacz (2014). "Not Only What but Also When: A Theory of Dynamic Voluntary Disclosure," *American Economic Review*

[13] Kamenica, E., and M. Gentzkow (2011). "Bayesian Persuasion." *American Economic Review.* 101(6), 259–2615.

[14] Kokolakis, S. (2017). "Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon." *Computers & Security.* Pp. 122-134.

[15] Milgrom, Paul (1981) "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* 12, 380–391.

[16] Myerson, Roger B. (1981). "Optimal Auction Design". *Mathematics of Operations Research* 6 (1): 58–73.

[17] Stigler, G. J. (1980). "An Introduction to Privacy in Economics and Politics." *Journal of Legal Studies* 9, pp. 623–644

[18] Taylor C. (1999). "Time-on-the-Market as a Sign of Quality." *Review of Economic Studies* 66, 555-78.

[19] Taylor (2004). "Consumer Privacy and the Market for Customer Information." Rand Journal of Economics, 35(4) 631-650 (20 pages)

[20] Turow, J., King, J., Hoofnatle, C.J., et al. (2009). "Americans Reject Tailored Advertising and Three Activities That Enable It." Departmental Papers (ASC), 137.

[21] Rose, Ellen A. (2005). "Data Users versus Data Subjects: Are Consumers Willing to Pay for Property Rights to Personal Information?" *Proceedings of the 38th Annual Hawaii International Conference on System Sciences.* Washington, D.C.: IEEE Computer Society Press. doi:10.1109/HICSS.2005.184.

[22] Tsai, Janice Y., S. Egelman, L. Cranor, and Alessandro Acquisti. 2011. "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study." *Information Systems Research* 22:254–68.

6 Appendix

Proof. [Proof of Proposition 2] First note that an equilibrium in which all types pool on $m = 0$ exists. This is true because if any type would deviate to some choice m' the seller can attribute this to coming from $\overline{\theta}$ given the full-support assumption on each s_t. Since given such beliefs it is an equilibrium off-path strategy to then keep charging $p = \overline{\theta}$, no profitable deviation exists. It remains to prove that no different equilibrium exists. Because type $\underline{\theta}$ never obtains a surplus from trading in equilibrium, this type will choose investment $m = 0$. This argument rules out any pooling equilibrium on some investment choice $m' \neq 0$. Consider now the infimum θ of types that chooses some investment $m' \neq 0$. By also choosing m' type θ would not obtain a surplus from trading in equilibrium and would prefer to deviate to investment $m = 0$ because $c(m') > 0$. As θ is the infimum of types choosing m' for any $\epsilon > 0$ there must be type $\theta_\epsilon \in (\theta, \theta + \epsilon)$ that chooses m'. Let $\epsilon \in (0, c(a))$. Then type θ_ϵ profit from trading is smaller than ϵ as the price for buyers choosing m' is at best θ and this type would benefit by deviating to investment $m = 0$. The presence of this deviation rules out any semi-separating equilibrium. ■

Proof. [Proof of Proposition 3] As before, $\underline{\theta}$ must choose $x^* = \arg\min c(x, \underline{\theta})$. Suppose that some types choose an experiment $x' \neq x^*$. Consider again the infimum of the types choosing x^*. This type, or a type arbitrarily close to it, receives essential no surplus from trading. Since $c(x', \widehat{\theta}) > c(x^*, \widehat{\theta})$, the result follows from the previous argument. ■

Proof. [Proof of Proposition 4] Note first that irrespective of the provider's choice in equilibrium no type will buy offline. To show this, suppose that a set of types chose to shop offline. Take the infimum of these types. In equilibrium, such a type can have no surplus from buying offline given a profit maximizing seller. At the same time, this type loses benefit $b > 0$ net of the price of the platform. As long as the price is strictly below b, this type has a strict incentive to deviate, a contradiction. Similarly were the price strictly higher than b, no types would go through the platform, and thus the platform provider charges at most b. Finally, if the platform provider charged b and a buyer would go offline with positive probability then the platform could deviate by charging slightly less than b.

Suppose now that the platform provider offered two different platforms at different prices. By our previous argument, all types choose the cheaper platform. Hence, in equilibrium all platforms will be offered at price b. Furthermore, since the price is the same for all equilibrium platforms and the expected transfer the platform provider obtains from the seller is strictly increasing in the seller's expected revenue, in equilibrium the platform provider will maximize the seller's revenue. ■

Proof. [Proof of Proposition 5] By the same argument as in earlier environments studied, all buyer types buy from one of the cheapest platforms irrespective of their informational content. Thus, the platform providers are engaged in the Bertrand competition on subsidies and each provider whose platform is chosen with positive probability (also referred to as a winning provider) offers to the buyer the maximum of the subsidies that other providers are able to offer. By our assumption on the bargaining between platforms and the seller, the platforms that maximize the seller's revenue also maximize the platform provider's profits and their ability to subsidize the buyer; thus these platforms are chosen by the winning providers. ■