Online Privacy and Information Disclosure by Consumers

Shota Ichihashi*

June 26, 2018

Abstract

I study the welfare and price implications of consumer privacy. A consumer discloses information to a multi-product seller, which learns about his preferences, sets prices, and makes product recommendations. While the consumer benefits from accurate product recommendations, the seller may use the information to price discriminate. I show that the seller prefers to commit to not price discriminate to encourage information disclosure. However, this commitment hurts the consumer, who could be better off by precommitting to withhold some information. In contrast to single-product models, equilibrium is typically inefficient even if the consumer can disclose any information about his preferences.

1 Introduction

This paper studies the welfare and price implications of consumers' privacy in online marketplaces, which are first-order issues in the Internet economy. Online sellers can observe detailed informa-

*Stanford University, Department of Economics, Email: shota2@stanford.edu. I am grateful to Paul Milgrom, Gabriel Carroll, Matthew Gentzkow, and Ilya Segal for their unfailing support and guidance. I would like to thank Mohammad Akbarpour, Susan Athey, Luca Braghieri, Isaías Chaves, Weixin Chen, Piotr Dworczak, Takako Fujiwara-Greve, Pedro Gardete, Michihiro Kandori, Fuhito Kojima, Shengwu Li, Sitian Liu, Qihong Liu, Giorgio Martini, Guido Martirena, Shunya Noda, Takuo Sugaya, and participants in various seminars. I am grateful for the financial support from the Yoshida Scholarship Foundation and the Stanford Graduate Fellowship. The author declares that he has no relevant material or financial interests that relate to the research described in this paper.

tion about consumers, such as their browsing histories, purchases, and characteristics; however, consumers can often affect whether and to what extent this information is revealed. For instance, they can disable cookies in order not to disclose their web-browsing activities, or they can use their social networking accounts to log in to online shopping websites. For policymakers, information disclosure by consumers is an important consideration in formulating policies concerning online privacy.

In this paper, I focus on the following economic trade-off: The benefit for consumers to disclose information is that sellers can recommend or advertise appropriate products. The cost is that sellers may use this information to price discriminate. For instance, Amazon, Netflix, Spotify, and other e-commerce sellers use consumers' personal data to offer product recommendations, which help consumers discover items that they might not have found otherwise. However, these firms could potentially use such information to obtain estimates of consumers' willingness to pay and, in turn, set prices on this basis.

I study a simple model capturing this trade-off. The model consists of a monopolistic seller of K products and a consumer with unit demand. The consumer is initially uninformed of his value of each product. At the beginning of the game, he chooses a disclosure rule, which determines what the seller learns about his value for each product. After learning about the valuations, the seller recommends one of the products. Finally, the consumer observes the value and the price of the recommended product and decides whether to buy it.

One novel aspect of my analysis is to consider two versions of the model that differ in whether the seller can price discriminate. Under the *discriminatory pricing regime*, the seller sets prices after observing the information disclosed by the consumer. Under the *nondiscriminatory pricing regime*, the seller posts a price for each product before observing the information. Considering two pricing regimes enables us to study the interaction between the seller's ability to price discriminate and the consumer's disclosure incentive.

Information disclosure is modeled in a similar way to Bayesian persuasion (Kamenica and Gentzkow, 2011): Without observing his valuations, the consumer chooses what information to be disclosed to the seller. The idea is that while it is difficult for consumers themselves to determine which item in a huge set of available products is most appropriate for them, sellers can often do this using consumers' personal data. For instance, sellers might analyze browsing histories by using their knowledge of the products' characteristics, the prior experiences of other consumers, and their computing power. These enable sellers to map a given consumer's personal data into estimates of his willingness to pay, even though the consumer himself cannot evaluate all products in the market.

I obtain three main findings, which sharply contrast with the classical theory of price discrimination and have implications for understanding observed facts, privacy regulations, and the theoretical literature of information disclosure. First, the seller is better off under nondiscriminatory pricing. The seller's inability to price discriminate encourages the consumer to disclose information, which makes recommendations more accurate and increases revenue. This result gives a potential economic explanation of an observed puzzle: "The mystery about online price discrimination is why so little of it seems to be happening" (Narayanan, 2017). Namely, price discrimination by online sellers seems to be uncommon despite their potential ability to use consumers' personal data to obtain estimates of their willingness to pay and, in turn, vary prices on this basis to capture more of the surplus.[1]

Second, the consumer is worse off under nondiscriminatory pricing. When the seller can price discriminate, the consumer decides what information to reveal taking into account how information affects prices. In contrast, under nondiscriminatory pricing, the consumer discloses much information to obtain better recommendations without worrying about price discrimination. However, expecting this greater level of disclosure and resulting accurate recommendations, the seller prefers to set a high price for each product upfront.[2] As a result, the consumer discloses more information and obtains a lower payoff under nondiscriminatory pricing.

Third, equilibrium is often inefficient even if the consumer can disclose any information about his valuations. This contrasts with the single-product case of Bergemann, Brooks, and Morris (2015), in which equilibrium is efficient under discriminatory pricing. In my model, discriminatory pricing discourages information disclosure and makes product mismatch more likely. The proof is based on a "constrained" Bayesian persuasion problem, in which the consumer chooses

[1] There have been several attempts by researchers to detect price discrimination by e-commerce websites. For instance, Iordanou et al. (2017) examine around two thousand e-commerce websites and they "conclude that the specific e-retailers do not perform PDI-PD (personal-data-induced price discrimination)."

[2] Formally, I show that if the consumer reveals more information about which product is more valuable, the valuation distribution for the recommended product shifts in the sense of a lower hazard rate, which gives the seller an incentive to charge a higher price. Note that the first-order stochastic shift is not sufficient to conclude that the seller prefers to set high prices.

a disclosure rule subject to the constraint that the outcome is efficient. This proof strategy could be useful for analyzing complex information design problems where it is difficult to characterize optimal information structures.

The main insights are also applicable to offline transactions. For example, consider a consumer looking for a car. He may talk to a salesperson and reveals some information—such as his lifestyle, favorite color, and preferences for fuel efficiency versus horsepower—which is indicative of his tastes; even his clothes and smartphone may reveal his preferences. Based on this information and her knowledge about available cars, the salesperson gives recommendations. On the one hand, the consumer benefits from the recommendations because he can avoid extra search and test-driving. On the other hand, disclosing too much information may put him in a disadvantageous position in price negotiation, because knowing that he loves a particular car, the salesperson would be unwilling to compromise on prices. To consumers, this paper provides a relevant trade-off regarding how much and what kind of information to reveal. To dealers, it gives a normative prescription for pricing strategy—giving salespersons discretion in pricing could hurt revenue, because it gives consumers an incentive to mask some information crucial to improving the match quality between consumers and products.

The remainder of the paper is organized as follows. In Section 2, after discussing related work, I present the baseline model. I also provide a second interpretation of the model as information disclosure by a continuum of consumers. In Section 3, I restrict the consumer to choosing from a simple class of disclosure rules and show that the seller is better off and the consumer is worse off under nondiscriminatory pricing. Section 4 allows the consumer to choose any disclosure rule. I show that equilibrium is typically inefficient and use the inefficiency results to establish the main result. This section also shows that nondiscriminatory pricing can be more efficient. In Section 5, I discuss several extensions including markets for personal data. Section 6 concludes.

1.1 Related Work

My work relates to two strands of literature: The literature on information design and that on the economics of privacy. In terms of modeling, one related work is Bergemann, Brooks, and Morris (2015), who consider a single-product monopoly pricing in which a monopolist has additional information about a consumer's valuation. I consider a multi-product seller with product recommendations, which renders information useful not only for price discrimination but also for improving product match quality. These new elements overturn the welfare consequences of information disclosure and discriminatory pricing. In contrast to Bergemann et al. (2015), who characterize the entire set of attainable surplus, I restrict attention to the consumer disclosing information to maximize his own payoff. Part of my analysis employs their "greedy algorithm," which generates a consumer-optimal information disclosure rule given any prior valuation distribution.

My work also relates to the economics of privacy literature. As a growing number of transactions are based on data about consumers' behavior and characteristics, recent strands of literature have devoted considerable attention to the relationship between personal data and intertemporal price discrimination (Acquisti and Varian, 2005; Conitzer et al., 2012; Fudenberg and Tirole, 2000; Fudenberg and Villas-Boas, 2006; Taylor, 2004; Villas-Boas, 1999, 2004). In these models, sellers learn about a consumer's preferences from his purchase record, which arises endogenously as a history of a game. A consumer's attempt to hide information is often formulated as delaying purchase or erasing purchase history. In my model, the consumer is endowed with his personal data at the outset.

Several papers, such as Conitzer et al. (2012) and Montes et al. (2017), examine consumers' endogenous privacy choices. Braghieri (2017) studies a consumer search model in which a consumer can choose to be targeted by revealing his horizontal taste to sellers instead of engaging in costly search. In the model, targeting plays a similar role to information disclosure in this paper: It enables consumers to find their favored products at low cost, but it can also hurt them because of discriminatory pricing. A similar trade-off arises in De Corniere and De Nijs (2016), who study a platform's choice of disclosing consumers' preferences to advertisers.

This paper differs from these works at least in two ways. First, in addition to discriminatory pricing, I consider sellers who cannot price discriminate. Comparing these two scenarios helps us explain, for instance, why sellers might want to commit to not price discriminate, and why price discrimination can lower efficiency. Second, in contrast to the literature where a consumer's privacy choice is typically full or no disclosure, I assume that he can choose how much and what kind of information to disclose. This enables me to study the relationship between pricing regimes and types of information revealed in equilibrium.

Beyond the context of online disclosure, this paper relates to voluntary information disclosure in bilateral transactions (Glode et al., 2016). Also, as information disclosure with commitment can be interpreted as a combination of information gathering and truthful disclosure, my work also relates to information gathering by buyers before trade (Roesler, 2015; Roesler and Szentes, 2017).

Finally, several papers, such as Calzolari and Pavan (2006a,b), and Dworczak (2017), study the privacy of agents in mechanism design problems. In their models, a principal can commit to a mechanism which elicits an agent's private type and a disclosure rule which reveals information about an outcome of the mechanism to other players. Relative to these works, the consumer in my model has more commitment power regarding what information to provide, and the seller has less commitment power in determining allocation and pricing.

2 Baseline Model

There is a monopolistic seller of $K \in \mathbb{N}$ products with the set of products denoted by $\mathcal{K} = \{1, \ldots, K\}$. There is a single consumer with unit demand, in that he eventually consumes one of K products or nothing. The consumer's value for product k, denoted by u_k, is drawn independently and identically across $k \in \mathcal{K}$ according to probability distribution x_0 supported on $V \subset \mathbb{R}_+$.[3] For the moment, I do not impose any assumptions on x_0, except that all the relevant integrals exist. Let $u := (u_1, \ldots, u_K)$ denote the vector of valuations.

The consumer's preferences are quasi-linear: If he buys product k at price p, his ex post payoff is $u_k - p$. If he buys no products, he obtains a payoff of zero. The seller's payoff is its revenue. The consumer and the seller are risk-neutral.

At the beginning of the game, before observing u, the consumer chooses a *disclosure rule* (M, ϕ) from an exogenously given set $\mathcal{D}$. Each element of $\mathcal{D}$ is a pair of a message space M and a function $\phi : V^K \to \Delta(M)$, where $\Delta(M)$ is the set of the probability distributions over M. After the consumer chooses a disclosure rule (M, ϕ), Nature draws $u \in V^K$ and a message $m \in M$ according to $\phi(\cdot|u) \in \Delta(M)$. In the application of online disclosure, $\mathcal{D}$ consists of consumers' privacy choices, such as whether to share one's browsing history or not. As in Section 4, if $\mathcal{D}$ consists of all disclosure rules, information disclosure takes the form of Bayesian persuasion

[3] See Remark 5 for how the results extend to correlated values.

studied by Kamenica and Gentzkow (2011). Hereafter, I sometimes write a disclosure rule as ϕ instead of (M, ϕ).

Next, I describe the seller's pricing. I consider two games that differ in whether the seller can price discriminate on the basis of information. Under the *discriminatory pricing regime*, the seller sets the price of each product *after* observing a disclosure rule (M, ϕ) and a realized message m. Under the *nondiscriminatory pricing regime*, the seller sets the price of each product *simultaneously* with the consumer's choice of a disclosure rule (M, ϕ).[4] Note that under nondiscriminatory pricing, the seller not only does not base prices on a realized message m but also does not base prices on a disclosure rule ϕ.[5]

Under both pricing regimes, after observing a disclosure rule (M, ϕ) and a realized message m, the seller recommends one of K products. The consumer observes the value and price of the recommended product and decides whether to buy it.

The timing of the game under each pricing regime, summarized in Figure 1, is as follows. First, the consumer chooses a disclosure rule $(M, \phi) \in \mathcal{D}$. Under the nondiscriminatory pricing regime, the seller simultaneously sets the price of each product. Then Nature draws the consumer's valuations u and a message $m \sim \phi(\cdot|u)$. After observing (M, ϕ) and m, the seller recommends a product. Under the discriminatory pricing regime, the seller sets the price of the recommended product at this point. Finally, the consumer decides whether to buy the recommended product.

My solution concept is subgame perfect equilibrium that satisfies three conditions. First, the seller breaks a tie in favor of the consumer whenever it is indifferent. Second, under nondiscriminatory pricing, I focus on equilibrium in which each product has the same price. Third, if there are multiple equilibria which give the consumer identical expected payoff, I focus on an equilibrium which gives the seller the highest payoff. The third condition eliminates the multiplicity of equilibria due to the consumer's indifference among disclosure rules, which is not the main focus

[4] I can alternatively assume that under nondiscriminatory pricing, the seller sets prices first, and after observing them, the consumer chooses a disclosure rule. This assumption does not change equilibrium if the consumer can only reveal information about which product has the highest valuation as in Section 3. In contrast, it could change equilibrium if the consumer can disclose information in an arbitrary way as in Section 4, because the seller may set different prices for different products to induce an asymmetric disclosure rule. (An example is available upon request.) However, the main result continues to hold: The seller is better off and the consumer is worse off under nondiscriminatory pricing. This is because the seller setting prices strictly before the consumer only increases the seller's revenue under nondiscriminatory pricing.

[5] For example, if e-commerce firms adopt this regime, they set prices based on neither browsing history nor whether consumers share their browsing history.

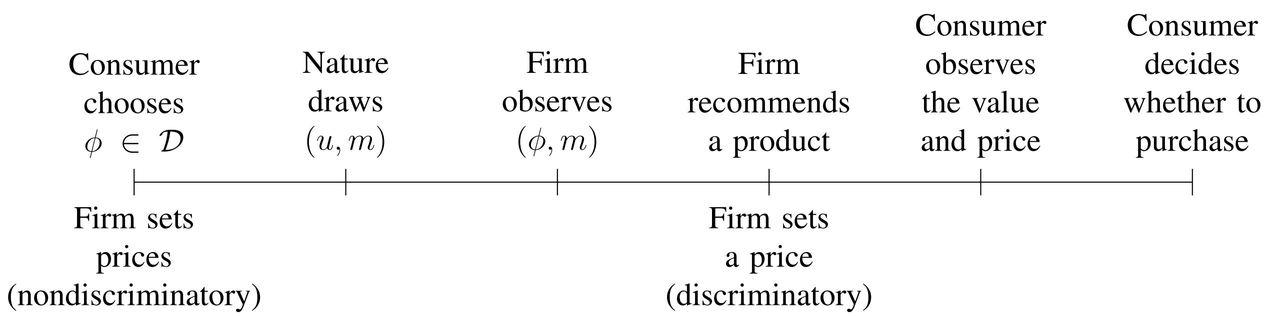

Figure 1: Timing of moves under each pricing regime

of the paper. Hereafter, "equilibrium" refers to SPE with these conditions.

Remark 1 (Other Applications). As I discuss in the introduction, there are many applications beyond online privacy choices. Consider markets for cars, houses, and financial products, in which the variety of available products is huge. In these markets, consumers often reveal information to sellers to obtain product recommendations, which enable consumers to focus on a small subset of products; however, sellers may also base prices on the information. The model captures the interaction between consumers' incentives to reveal information and sellers' pricing strategies in those markets.

Indeed, the application is not even restricted to ordinary buyer-seller interactions. Consider the following situation, which is mathematically equivalent to the baseline model: An employer assigns his worker one of K tasks, the completion of which delivers a fixed value to the employer. The worker can disclose information about cost c_k that he incurs to complete each task k. In this case, two pricing regimes could correspond to whether the wage is set contingent on the revealed information.

Remark 2 (Discussion of Modeling Assumptions). The model departs from the economics of privacy literature by being agnostic about what personal data a consumer is endowed with and what privacy choices he has. Instead, the consumer's choice set is defined as a set of Blackwell experiments about his valuations. This formulation calls for several implicit assumptions; for example, the consumer understands how his privacy choice affects the firm's posterior belief. While such an assumption might be restrictive, it enables us to draw general insights on consumers' informational incentives in online and offline transactions without referring to specific disclosure technologies.

Relatedly, it is crucial to my results that the consumer chooses a disclosure rule before observing his valuations u. This would be suitable, for instance, if the consumer is not informed of the existence or characteristics of products, but understands that his personal data enable the seller to learn about which product is valuable to him. It would also be natural to assume that the consumer cannot manipulate message realizations ex post, as consumers or regulators typically set disclosure rules up front and incentives to distort or misrepresent one's browsing history or characteristics seem to be less relevant. In Section 5, I provide a microfoundation for this idea in a model of two-sided private information, where the consumer is informed of his subjective taste and the seller is informed of its products' characteristics.

There are also two substantial assumptions on the recommendation and purchasing decision. First, the seller recommends one product and the consumer decides whether to buy it. Although this assumption would be empirically false, it parsimoniously captures situations in which the variety of available products is large and sellers make product recommendations to help consumers focus on a strict subset of the products.[6] The main insights continue to hold if a consumer can examine more than one but strictly less than K products.

Second, the consumer observes his willingness to pay for the recommended product when he decides whether to buy it. One way to interpret this assumption is that the consumer does not know what products exist, and has not thought about how much he would be willing to pay for each possible bundle of characteristics; however, once he is shown a particular product and sees its characteristics, he is able to compute a value for it. In practice, the assumption is reasonable if a consumer can learn the value after the purchase and return it for a refund whenever the price exceeds the value.

Finally, it is *not* without loss of generality to assume that production costs are equal across products. (Assuming that they are equal, it is without loss to normalize them to zero.) For example, if the seller can produce product 1 more cheaply, it has a greater incentive to recommend product 1 even if it is less valuable to the consumer than other products. Correspondingly, heterogeneous production costs are likely to affect the consumer's incentive to disclose information. I leave this

[6]Several papers, such as Salant and Rubinstein (2008) and Eliaz and Spiegler (2011), formulate consumers' limited attention in a similar manner.

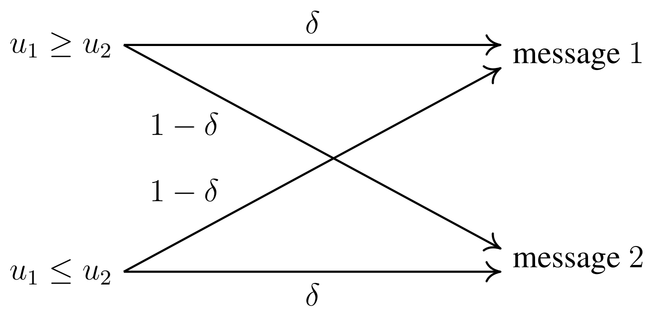

Figure 2: Disclosure rule for $\delta \in [1/2, 1]$

extension for future research.

3 Restricted Information Disclosure

In this section, I simplify the baseline model as follows. First, assume that the seller sells two products ($K = 2$). Then, let $\mathcal{D} = [1/2, 1]$ and call each $\delta \in [1/2, 1]$ a *disclosure level*. Each δ corresponds to the following disclosure rule $(\{1, 2\}, \phi_\delta)$: $\phi_\delta(\cdot|u_1, u_2)$ draws message $k \in \{1, 2\}$ with probability δ whenever $u_k = \max(u_1, u_2)$ and $u_1 \neq u_2$; if $u_1 = u_2$, ϕ_δ uniformly randomizes between messages 1 and 2. Figure 2 depicts the disclosure rule corresponding to δ. Note that the greater δ is, the more informative ϕ_δ is in the sense of Blackwell.

Remark 3. Assuming $K = 2$ is to simplify the analysis. For instance, one can obtain similar results with the following assumption: There are K products, and ϕ_δ sends message k with probability δ if $k \in \arg\max_{k \in \mathcal{K}} u_k$, and it sends any message $k \in \{1, \ldots, K\}$ with probability $\frac{1-\delta}{K}$. I can also relax the assumption that, at the most informative disclosure rule ($\delta = 1$), a message realization is a deterministic function of the valuations (except for $u_1 = u_2$). A similar result holds, for instance, if the consumer can choose from any garblings of the disclosure rule that sends message k whenever $k \in \arg\max_k(u_k + \varepsilon_k)$ where ε_k is IID across k. Section 5 more generally defines a disclosure rule that reveals horizontal information and shows that the identical results hold if the consumer can choose any garblings of such a disclosure rule.

3.1 Equilibrium Analysis

The analysis consists of three steps. First, I show that greater disclosure levels lead to more accurate product recommendations. Second, under discriminatory pricing, disclosure also leads to higher prices for recommended products. Finally, I combine these lemmas and prove the main result: The seller is better off and the consumer is worse off under nondiscriminatory pricing.

The following lemma states that, to maximize revenue, the seller recommends the product that the consumer is more likely to prefer. See Appendix A for the proof.

Lemma 1. *Fix a pricing regime. Take any equilibrium and consider a subgame in which the consumer has chosen a disclosure level $\delta > 1/2$. Then, the seller recommends product $k \in \{1, 2\}$ after observing message k.*

The equilibrium behavior in Lemma 1 encourages the consumer to disclose information, as it increases the chance that the seller recommends the most valuable product. Indeed, disclosure level δ is precisely the probability that the consumer is recommended the preferred product.

Now, how does information disclosure affect product prices? Let F^{MAX} and F^{MIN} denote the cumulative distribution functions of $\max(u_1, u_2)$ and $\min(u_1, u_2)$, respectively.[7] Conditional on a disclosure level δ and a realized message $k \in \{1, 2\}$, the value distribution of the recommended product (i.e. product k) is given by $\delta F^{MAX} + (1 - \delta)F^{MIN}$. Given this value distribution, the firm sets an optimal price.

To study how the optimal price depends on δ, I show that F^{MAX} is greater than F^{MIN} in the *hazard rate order*: F^{MAX} has a lower hazard rate than F^{MIN}. This is stronger than the first-order stochastic dominance, which has no implications on the behavior of the monopoly price.[8] The following definition does not require distributions to have densities.

Definition 1. Let G_0 and G_1 be two CDFs. G_1 is greater than G_0 in the *hazard rate order* if $\frac{1-G_1(z)}{1-G_0(z)}$ increases in $z \in (-\infty, \max(s_1, s_0))$.[9] Here, s_0 and s_1 are the right endpoints of the supports of G_0

[7] In this paper, I define CDFs as left-continuous functions. Thus, for instance, $F^{MAX}(p)$ is the probability of $\max(u_1, u_2)$ being *strictly* lower than p.

[8] For example, suppose that distribution F_0 puts equal probability on values 1 and 3 and that distribution F_1 puts equal probability on 2 and 3. Though F_1 first-order stochastically dominates F_0, the monopoly price under F_0 is 3, while the one under F_1 is 2.

[9] $a/0$ is taken to be equal to $+\infty$ whenever $a > 0$.

and G_1, respectively.[10]

The proof of the next lemma is in Section 5, where I prove the same result for a more general formulation of "horizontal information," the information that is useful for accurate recommendations.

Lemma 2. *F^{MAX} is greater than F^{MIN} in the hazard rate order.*

The intuition is as follows. Suppose that the consumer prefers to buy a product at price p. Conditional on this event, if the seller marginally increases the price by ε, how likely the consumer is to stop buying? If the product is his preferred one so that the value is $\max(u_1, u_2)$, he stops buying only when both u_1 and u_2 are below $p+\varepsilon$; if the product is his less preferred one so that the value is $\min(u_1, u_2)$, he stops buying whenever one of u_1 and u_2 is below $p+\varepsilon$. Thus, the consumer is less likely to stop buying after observing a marginal price increment, if the recommended product is his preferred one. This implies that the value distribution F^{MAX} has a lower hazard rate than F^{MIN}.

This intuition suggests that the hazard rate order relates to price elasticity of demands. Indeed, for two CDFs F_1 and F_0, F_1 is greater than F_0 in the hazard rate order if and only if the "demand curve" for F_1 has a lower price elasticity of demand than the one for F_0. Here, as in Bulow and Roberts (1989), the demand curve for F is given by $D(p) = 1 - F(p)$, and thus the demand elasticity is $-\frac{d \log D(p)}{d \log p} = \frac{f(p)}{1-F(p)}p$. Thus, the previous lemma states that the consumer's demand for the more preferred product is less elastic.

We are ready to prove the key comparative statics: More information disclosure leads to higher prices for recommended products under discriminatory pricing.The result holds for any prior distribution x_0.

Lemma 3. *Consider discriminatory pricing. Let $p(\delta)$ denote the equilibrium price for the recommended product given a disclosure level δ.[11] Then, $p(\delta)$ is increasing in δ.*

[10] If G_0 and G_1 have densities g_0 and g_1, the above definition is equivalent to

$$\frac{g_0(z)}{1-G_0(z)} \geq \frac{g_1(z)}{1-G_1(z)}, \forall z \in (-\infty, \max(s_1, s_0)).$$

[11] The lowest optimal price $p(\delta)$ exists and is uniquely defined given the firm's tie-breaking rule for recommendation and pricing.

Proof. By Lemma 1, the seller recommends the best product with probability δ. Then, at price p, trade occurs with probability $1 - \delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)$. Thus, the set of optimal prices is $P(\delta) := p[1 - \delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)]$. To show that $p(\delta) = \min P(\delta)$ is increasing in δ, note that

$$\begin{aligned}&\log p[1 - \delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)] - \log p[1 - \delta' F^{MAX}(p) - (1-\delta')F^{MIN}(p)] \\ &= \log \frac{1 - \delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)}{1 - \delta' F^{MAX}(p) - (1-\delta')F^{MIN}(p)}. \end{aligned} \tag{1}$$

By Theorem 1.B.22 of Shaked and Shanthikumar (2007), if $\delta > \delta'$, $\delta F^{MAX} + (1-\delta)F^{MIN}$ is greater than $\delta' F^{MAX} + (1-\delta')F^{MIN}$ in the hazard rate order. Then, (1) is increasing in p. This implies that $\log p[1-\delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)]$ has increasing differences in (p, δ). By Topkis (1978), $P(\delta)$ is increasing in the strong set order.[12] Therefore, $p(\delta)$ is increasing in δ. □

Combining the previous lemmas, I obtain the main result.

Theorem 1. *Each pricing regime has a unique equilibrium. Moreover, the seller obtains a higher payoff and the consumer obtains a lower payoff under nondiscriminatory pricing than under discriminatory pricing.*

Proof. If the consumer is recommended his preferred and less preferred products at price p, his expected payoffs are $u^{MAX}(p) := \int_p^{+\infty}(v-p)dF^{MAX}(v)$ and $u^{MIN}(p) := \int_p^{+\infty}(v-p)dF^{MIN}(v)$, respectively.

Consider nondiscriminatory pricing. Let p^* denote the equilibrium price for each product.[13] If the consumer chooses δ, his expected payoff is

$$\delta u^{MAX}(p^*) + (1-\delta)u^{MIN}(p^*).$$

$\delta = 1$ maximizes this expression and is a unique disclosure level consistent with our equilibrium constraint. Thus, the equilibrium price is $p^* = p(1)$, and the consumer's payoff is $u^{MAX}(p(1))$.

[12] $A \subset \mathbb{R}$ is greater than $B \subset \mathbb{R}$ in the strong set order if $a \in A$ and $b \in B$ imply $\max(a,b) \in A$ and $\min(a,b) \in B$.
[13] As I focus on subgame perfect equilibrium in which the seller breaks tie in favor of the consumer, the seller always sets the lowest price among the optimal prices, which excludes the use of strictly mixed strategy in pricing.

Under discriminatory pricing, the consumer's payoff from disclosure level δ is

$$\delta u^{MAX}(p(\delta)) + (1-\delta)u^{MIN}(p(\delta)). \tag{2}$$

Thus, the equilibrium payoff is

$$\max_{\delta \in [1/2,1]} \delta u^{MAX}(p(\delta)) + (1-\delta)u^{MIN}(p(\delta)) \geq u^{MAX}(p(1)).$$

That is, the consumer is worse off under nondiscriminatory pricing.

Finally, consider the seller's payoff. As $\delta F^{MAX} + (1-\delta)F^{MIN}$ is stochastically increasing in δ, $p[1-\delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)]$ is increasing in δ for any p. This implies that the $\max_p p[1-\delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)]$ is maximized at $\delta = 1$: The seller is better off under nondiscriminatory pricing. $\square$

The intuition is as follows. Under nondiscriminatory pricing, the consumer fully reveals information to obtain good recommendations without worrying about price discrimination. Expecting this, the seller sets a high price for each product upfront. In contrast, under discriminatory pricing, the consumer chooses a disclosure level balancing the benefit of better product match and the cost of higher prices. As a result, the consumer withholds some information, which leads to lower prices and a higher payoff than under nondiscriminatory pricing.

Theorem 1 gives an economic explanation of the observed puzzle: Online sellers seem to not use individual data to price discriminate,[14] and consumers seem to casually share their information despite the growing concerns for personalized pricing.In light of the theorem, one may view the puzzle as sellers' strategic commitment to encourage disclosure and consumers' best response to nondiscriminatory pricing. However, the result also suggests that such a situation might not be desirable for consumers, because they could be better off were sellers less informed about their preferences.

[14]For empirical studies indicative of this, see the discussion in the introduction. For another instance, in 2000 Amazon CEO Jeff Bezos said, "We never have and we never will test prices based on customer demographics." (`http://www.e-commercetimes.com/story/4411.html`). Of course, there could be other explanations for firms not price discriminating. For instance, firms might be worried that price discrimination infuriates consumers, who have fairness concerns. My result suggests that even if firms are able to frame personalized pricing in a way that consumer backlash is less likely to occur, they may still want to refrain from price discrimination to limit consumers' (rational) incentives to game the system.

Theorem 1 also has policy implications: Consumers could benefit from regulations that restrict the amount of information sellers can expect to acquire. To see this, suppose that the seller can commit to nondiscriminatory pricing, under which the consumer chooses the greatest disclosure level. Relative to this situation, he is better off if a regulator restricts the set $\mathcal{D}$ of available disclosure levels to $[1/2, \delta^*]$, where δ^* is the equilibrium choice under discriminatory pricing (i.e., a maximizer of equation(2)). With this restriction, the seller is indifferent between two pricing regimes. The consumer chooses disclosure level δ^* and obtains a greater payoff than without the regulation.[15]

One may think that the seller preferring nondiscriminatory pricing is specific to the current assumption on the set of available disclosure rules, $\mathcal{D}$. On the one hand, this is partly true: $\mathcal{D}$ is a key parameter for my result. For example, if $\mathcal{D}$ consists only of the disclosure rule that reveals exact valuations, the seller is better off and the consumer is worse off under discriminatory pricing due to perfect price discrimination. On the other hand, the current restriction on $\mathcal{D}$ is not necessary for Theorem 1. First, $\mathcal{D}$ can be any subset of the disclosure levels $[1/2, 1]$. For example, we can set $\mathcal{D} = \{1/2, 1\}$, which consists of, say, enabling cookies ($\delta = 1$) and disabling cookies ($\delta = 1/2$). Thus, the (unrealistic) assumption that the consumer can choose from a continuum of disclosure level is immaterial. Second, the result is not specific to the assumption that the consumer can only disclose information about his ordinal preferences; in the next section, I establish essentially the same result assuming that the consumer can choose *any* disclosure rule.

Remark 4 (Limited Ability to Evaluate Products vs. Ability to Choose Disclosure Rules). One might wonder how many of consumers, who cannot examine all the available products in the market, have enough time to figure out optimal disclosure rules. I argue that these two assumptions do not contradict in most of the applications.

First of all, e-commerce firms, such as Amazon and eBay, obviously sell more products than one can exhaustively examine. Then, it has to be an institutional feature of these platforms to display only a small subset of the whole universe of products. Similarly, offline markets for cars or houses seem to have more products than one can exhaustively test-drive or visit, even if consumers restrict their attention to a particular brand or geographic area. In such cases, we cannot conclude

[15] One should note that such a regulation does not always reduce total surplus, because lower disclosure levels lead to lower prices which may increase the probability of trade.

that if consumers know how to use privacy tools or mask information from sellers, they should also be able to find relevant products without the help of search engines and recommendations.

Second, in some situations, it is relatively easy to figure out how to withhold information. For instance, on the Internet, it is increasingly common that pop-up windows ask users whether to enable cookies, likely because of the EU legislation on cookies. In offline markets, withholding information would be even easier—a consumer can disclose less by talking less about his tastes about cars, houses, and financial products.[16] In this case, consumers need not be highly sophisticated to figure out what "disclosure rules" are available to them.

3.2 Disclosure by a Continuum of Consumers and Theorem 1 as a Tragedy of the Commons

This subsection shows that we can interpret the current setting as a model with a continuum of consumers. This interpretation allows us to see Theorem 1 as a *tragedy of the commons* due to a negative externality associated with information sharing.

Formally, suppose that there is a continuum of consumers, each of whom discloses information as in the baseline model. The values are independent across consumers.[17] Under nondiscriminatory pricing, *after* observing the disclosure level and realized message of each consumer, the seller sets a single price for each product. Under discriminatory pricing, the seller can charge different prices to different consumers. Under both pricing regimes, the seller can recommend different products to different consumers.

Equilibrium prediction in Theorem 1 persists: Each consumer chooses the highest disclosure level under nondiscriminatory pricing, where the seller sets higher prices and consumers obtain lower payoffs. To see this intuitively, consider nondiscriminatory pricing and an equilibrium in which consumer $i \in [0, 1]$ chooses a disclosure level of δ_i. Note that δ_i maximizes each consumer's payoff when choosing a disclosure level taking prices as given, because the choice of a single consumer in a large population has no impact on prices.[18] This implies $\delta_i = 1$.

[16]The model of two-sided private information in Section 5 would capture this kind of information disclosure in offline transactions.

[17]The independence of valuation vectors across a continuum of consumers might raise a concern about the existence of a continuum of independent random variables. Sun (2006) formalizes the notion of a continuum of IID random variables for which the "law of large numbers" holds, which is all what I need.

[18]Appendix B formalizes this.

According to this interpretation, we can view Theorem 1 as a tragedy of the commons. If some consumers disclose more information, the seller prefers to set higher prices as it can offer accurate recommendations to a greater fraction of consumers. Because each product has a single price, these higher prices are shared by all consumers. To sum up, under nondiscriminatory pricing, information disclosure by some consumers lowers the welfare of other consumers through higher prices. As consumers do not internalize this negative impact, they prefer to fully reveal information and are collectively worse off. Appendix B formalizes this observation.

Remark 5. Theorem 1 is robust to a variety of extensions.

Correlated Values: If u_1 and u_2 are correlated, Theorem 1 holds as long as vector (u_1, u_2) is drawn from an exchangeable distribution whose multivariate hazard rate satisfies a condition in Theorem 1.B.29 of Shaked and Shanthikumar (2007). The condition ensures that $\max(u_1, u_2)$ is greater than $\min(u_1, u_2)$ in the hazard rate order, which is sufficient for the theorem.

Costly disclosure: Consumers may incur some intrinsic privacy costs by disclosing information. I can incorporate this by assuming that the consumer incurs a cost of $c(\delta)$ from a disclosure level δ.[19] This may affect equilibrium but does not change the main conclusion: The consumer discloses more information and is worse off under nondiscriminatory pricing. Finding an equilibrium under nondiscriminatory pricing requires a fixed-point argument, because the consumer may prefer different disclosure levels depending on prices he expects.

Informational Externality: In practice, online sellers may infer the preferences of some consumers from those of others. To incorporate this "informational externality," consider the model with a continuum of consumers, and assume that a "true" disclosure level of consumer i is given by $\Delta(\delta_i, \bar{\delta})$, an increasing function of i's disclosure level δ_i and the average disclosure level of the population $\bar{\delta} = \int_{i \in [0,1]} \delta_i di$. $\Delta(\delta_i, \bar{\delta})$ captures the idea that the seller can learn about i's preferences from information disclosed by others. In this case, I obtain essentially the same result as Theorem 1.

[19]More precisely, if the consumer chooses disclosure level δ and purchases product k at price p, his payoff is $u_k - p - c(\delta)$. If he buys nothing, the payoff is $-c(\delta)$.

4 Unrestricted Information Disclosure

In this section, I assume that the consumer can choose *any* disclosure rule and the seller sells $K \geq 2$ products. This "unrestricted model" enables us to study what kind of information is disclosed in equilibrium, and how it depends on pricing regimes. The previous model is not suitable for this purpose, as it assumes that the consumer reveals information about which product has a greater value, regardless of pricing regimes.

The unrestricted model is also useful for studying whether the main result persists in an environment that does not a priori favor nondiscriminatory pricing.[20] The previous assumption on $\mathcal{D}$ favors nondiscriminatory pricing, because the two pricing regimes yield equal revenue for any *fixed* δ. In general, for a fixed disclosure rule, the seller typically achieves a higher revenue under discriminatory pricing. For example, given a disclosure rule revealing u, the seller can extract full surplus only under discriminatory pricing.

The model has a theoretical connection to Bergemann et al. (2015). Their findings imply that a single-product monopolist is indifferent between the two pricing regimes if the information is disclosed to maximize consumer surplus; moreover, equilibrium is efficient under discriminatory pricing. I will show that introducing multiple products and recommendations substantively change the conclusion: equilibrium is typically inefficient and the seller strictly prefers nondiscriminatory pricing.

For ease of exposition, the prior distribution x_0 of the value of each product is assumed to have a finite support $V = \{v_1, \ldots, v_N\}$ with $0 < v_1 < \cdots < v_N$ and $N \geq 2$. For each $x \in \Delta(V)$, $x(v)$ denotes the probability that x puts on $v \in V$. Abusing notation slightly, let $p(x)$ denote the lowest optimal price given $x \in \Delta(V)$:

$$p(x) := \min \left\{ p \in \mathbb{R} : p \sum_{v \geq p} x(v) \geq p' \sum_{v \geq p'} x(v), \forall p' \in \mathbb{R} \right\}.$$

Note that $p(x)$ does not depend on K. To focus on the most interesting cases, I impose the following assumption. Loosely speaking, it requires that x_0 does not put too much weight on the lowest value of its support.

[20] It is important to note that the unrestricted model is not a general version of the previous restricted model but a version of the baseline model with a different assumption on $\mathcal{D}$.

Assumption 1. The lowest optimal price at the prior distribution strictly exceeds the lowest value of its support: $p(x_0) > v_1$.

As the consumer can access a rich set of disclosure rules, the equilibrium analysis is more involved than before; however, there turns out to be a clear relationship between pricing regimes and kinds of information that the consumer is willing to disclose. The next subsection illustrates this by showing that different pricing regimes exhibit different kinds of inefficiency in equilibrium. I use these results to show that, again, the seller is better off and the consumer is worse off under nondiscriminatory pricing.

4.1 Inefficiency of Equilibrium

In this model, an allocation can be inefficient in two ways. One is when the consumer decides not to buy any products; the other is when he buys some product other than the most valuable ones. The following result states that nondiscriminatory pricing leads to the first kind of inefficiency.

Proposition 1. *Consider nondiscriminatory pricing. In any equilibrium, the seller recommends the most valuable product with probability* 1*. However, trade fails to occur with a positive probability.*[21]

Proof. Take any equilibrium under nondiscriminatory pricing. Because prices are the same across products and do not depend on information disclosure, it is optimal for the consumer to disclose information so that the seller recommends the most valuable products with probability 1.[22] Now, given such a disclosure rule, consider the seller's pricing decision. When the seller recommends product k, the posterior belief for the value is equal to the distribution of $\max_{k \in \mathcal{K}} u_k$, because valuations are IID across products. Denoting the distribution by $x^{MAX} \in \Delta(V)$, we obtain

$$p(x_0) \sum_{v \geq p(x_0)} x^{MAX}(v) \geq p(x_0) \sum_{v \geq p(x_0)} x_0(v) > v_1,$$

[21] Note that I focus on subgame perfect equilibrium (SPE) in which the seller sets the same price for each product. Without this restriction, there could be a fully efficient SPE (An example is available upon request.) However, if the prior distribution has a density and satisfies Assumption 1, there is no efficient SPE. This is because in any SPE with efficient recommendations, the seller sets a price greater than $p(x_0)$ and trade may fail to occur with a positive probability.

[22] Such a disclosure rule is not unique; we can consider any disclosure rules weakly more informative than the one disclosing $\arg\max_k u_k$.

where the strict inequality follows from Assumption 1. Thus, the price for each product is strictly greater than v_1, and the consumer buys no products with a probability of at least $x(v_1)^K > 0$. $\square$

The intuition is analogous to Theorem 1: Under nondiscriminatory pricing, the consumer reveals information without worrying about price discrimination. This enables the seller to make accurate recommendations; however, because the seller cannot tailor prices to the consumer's willingness to pay, trade fails to occur if he has low values for all the products.

Does discriminatory pricing eliminate inefficiency? The next result shows that equilibrium continues to be inefficient: Although trade occurs whenever efficient, it is associated with product mismatch. The proof needs some work, which is contained in Appendix C. Below, I describe the basic idea of the proof.

Proposition 2. *Consider discriminatory pricing. In any equilibrium, trade occurs with probability* 1. *However, for generic*[23] *priors* x_0 *satisfying Assumption 1, in any equilibrium, the consumer purchases some products other than the most valuable ones with a positive probability.*

A rough intuition is as follows. Under discriminatory pricing, the seller can lower prices if it learns that the consumer has low valuations. As discounted offers also benefit the consumer, he is indeed willing to disclose such information. As a result, trade occurs even if the consumer has low values. In contrast, the seller's ability to price discriminate gives the consumer an incentive to withhold some information. What kind of information does the consumer want to mask? The second part of the proposition states that the information useful for product recommendations is exactly what the consumer prefers to obfuscate.

As discussed earlier, Proposition 2 is in contrast to the single-product case of Bergemann et al. (2015) in which a single disclosure rule maximizes both consumer and total surplus. In light of this, the result has the following takeaway. Consider a regulator or an Internet intermediary, who cares about consumers and wants to release their information to sellers in order to maximize consumer welfare. If information is relevant not only for pricing but also for product matching, such a regulator or an intermediary may have to balance enhancing consumer welfare and potential efficiency loss.

[23]The following is the formal description of genericity: Define $X_{>v_1}$ as the set of priors x_0 such that $p(x_0) > v_1$. "Generic priors x_0 satisfying Assumption 1" means that there is a Lebesgue measure-zero set $X_0 \subset \Delta(V)$ such that, for any $x_0 \in X_{>v_1} \setminus X_0$, any equilibrium has a positive probability of product mismatch.

Now, I sketch the proof of Proposition 2, relegating the details to Appendix C. For ease of exposition, I use the following terminologies.

Definition 2. An equilibrium is *vertically efficient* if trade occurs with probability 1. An equilibrium is *horizontally efficient* if the seller recommends the most valuable products with probability 1.

We can rephrase Proposition 2 as follows: Under discriminatory pricing, equilibrium is vertically efficient, but generically horizontally inefficient. The proof of vertical efficiency is by contradiction: If an equilibrium is vertically inefficient, we can modify it to another equilibrium where both the consumer and the seller are better off, which is a contradiction.[24]

To see this, suppose that in some equilibrium, the seller recommends product k at price p' that the consumer rejects with a positive probability. Modify the disclosure rule so that, on such an event, the consumer additionally reveals whether $u_k < p'$ or $u_k \geq p'$. (Note that this may reveal information about products other than k, because the equilibrium we initially choose may induce posteriors such that values are correlated across products.) If $u_k < p'$, the consumer and the seller are better off because the seller can either lower the price or recommend another product. The key is to prove that if $u_k \geq p'$, the seller continues to recommend the same product at the *same* price.[25] This shows that such a modification makes the consumer and the seller better off, which is a contradiction.[26]

Proving horizontal inefficiency is more involved at least for three reasons. First, disclosing more information (in the sense of Blackwell) may not lead to higher prices once we consider the full set of disclosure rules. Thus, we do not have simple comparative statics as in the restricted model. Second, it is challenging to characterize equilibrium disclosure rules, as we have to solve a multidimensional Bayesian persuasion problem, which is known to be difficult. Third, there can be multiple equilibria and we have to prove horizontal inefficiency for all of these.

[24]The vertical efficiency relates to Bergemann et al. (2015), which show that equilibrium is efficient if the seller sells a single product. In contrast to their work, which directly constructs a disclosure rule achieving an efficient outcome, my proof indirectly shows vertical efficiency, because it is difficult to characterize equilibrium.

[25]If the seller learns that the value exceeds p', it continues to post price p', as the seller's new objective function is $p \cdot \frac{\mathbf{P}(u_1 \geq p)}{\mathbf{P}(u_1 \geq p')}$, which has the same maximizer as the original objective $p \cdot \mathbf{P}(u_1 \geq p)$.

[26]The actual proof is a bit more subtle: First, I have to consider the case in which the consumer is exactly indifferent before and after the additional disclosure, in which case I use our equilibrium restriction. Second, I prove the existence of vertically efficient equilibrium, which is stronger than the claim that any equilibrium is vertically efficient (which can be vacuous).

The proof of horizontal inefficiency consists of two steps. First, solve a "constrained" Bayesian persuasion problem in which the consumer chooses a disclosure rule subject to the constraint that the resulting allocation is efficient (given the seller's optimal behavior). Characterizing such a disclosure rule, denoted by ϕ^G, turns out to be simpler than the "unconstrained" maximization problem which the consumer solves in equilibrium. Second, modify ϕ^G to create disclosure rule ϕ^I that leads to an inefficient allocation but gives the consumer a strictly greater payoff than ϕ^G. These two steps imply that any equilibrium is associated with inefficient allocation. As we proved that equilibrium is vertically efficient, it must be horizontally inefficient.

The following example illustrates these two steps.

Example 1. Suppose that $K = 2$, $V = \{1, 2\}$, and $(x_0(1), x_0(2)) = (1/3, 2/3)$.
Step 1: Consider disclosure rule ϕ in Table 1. (Each row is the distribution over messages 1 and 2 for each valuation vector.) ϕ only discloses which product is more valuable, which is the

	$\phi(1\|u_1, u_2)$	$\phi(2\|u_1, u_2)$
$(2, 2)$	$1/2$	$1/2$
$(2, 1)$	1	0
$(1, 2)$	0	1
$(1, 1)$	$1/2$	$1/2$

Table 1: Disclosure rule ϕ revealing ordinal ranking

information necessary to achieve horizontally efficient allocations. This implies that any efficient disclosure rules are weakly more informative than ϕ.[27]

I find ϕ^G by maximizing the consumer's payoff among disclosure rules weakly more informative than ϕ. Specifically, for each k, I calculate the posterior distribution of u_k conditional on message $k \sim \phi(\cdot|u)$, and apply Bergemann et al.'s (2015) "greedy algorithm" to it. In the single-product case, given any prior distribution, their greedy algorithm outputs a disclosure rule maximizing consumer surplus. In my multi-product case, applying it to each posterior belief induced by ϕ is equivalent to maximizing the consumer's payoff among the disclosure rules more informative than ϕ.

Table 2 presents disclosure rule ϕ^G obtained in this way. The greedy algorithm decomposes

[27] Precisely, I am restricting attention to "symmetric" disclosure rules such that permutating the indicies of the products do not change ϕ. In the proof, this is shown to be without loss of generality.

each message k (of ϕ) into messages $k1$ and $k2$.[28] The seller's best responses are as follows: After observing message $k1$ ($k = 1, 2$), the seller recommends product k at price 1, being indifferent between prices 1 and 2. After observing message $k2$ ($k = 1, 2$), the seller recommends product k at price 2.

	$\phi^G(11\|u_1, u_2)$	$\phi^G(12\|u_1, u_2)$	$\phi^G(21\|u_1, u_2)$	$\phi^G(22\|u_1, u_2)$
$(2, 2)$	0	$1/2$	0	$1/2$
$(2, 1)$	$1/4$	$3/4$	0	0
$(1, 2)$	0	0	$1/4$	$3/4$
$(1, 1)$	$1/2$	0	$1/2$	0

Table 2: Efficient disclosure rule ϕ^G.

	$\phi^G(11\|u_1, u_2)$	$\phi^G(12\|u_1, u_2)$	$\phi^G(21\|u_1, u_2)$	$\phi^G(22\|u_1, u_2)$
$(2, 2)$	0	$1/2$	ε'	$1/2 - \varepsilon'$
$(2, 1)$	$1/4$	$3/4 - \varepsilon$	ε	0
$(1, 2)$	0	0	$1/4$	$3/4$
$(1, 1)$	$1/2$	0	$1/2$	0

Table 3: Horizontally inefficient disclosure rule ϕ^I.

Step 2: I modify ϕ^G twice to create ϕ^I in Table 3: First, at $(u_1, u_2) = (2, 1)$, ϕ^I sends message 21 instead of 12 with probability $\varepsilon > 0$. For a small ε, this modification does not reduce the consumer's payoff: At the new message 12, the consumer continues to obtain a payoff of zero; at the new message 21, the seller continues to recommend product 2 at price 1. However, this modification relaxes the seller's incentive, as it now *strictly* prefers to set price 1 at message 21. Second, I further modify ϕ^G so that, at $(u_1, u_2) = (2, 2)$, ϕ^I sends message 21 instead of 22 with a small probability $\varepsilon' > 0$. This strictly increases the consumer's payoff: At $(u_1, u_2) = (2, 2)$, where the consumer obtains a payoff of zero at the original ϕ^G, he now obtains a strictly positive payoff when message 21 is realized. Because the seller recommends product 2 when $(u_1, u_2) = (2, 1)$ with a positive probability , ϕ^I is inefficient.

Finally, I discuss how the proof strategy extends to general parameters (K and x_0). Generalizing Step 1 is straightforward. For Step 2, first, I prove that disclosure rule ϕ^G obtained in Step 1 (generically) sends messages m_0 and m_1 with the following properties: Conditional on message

[28] As I discuss in the proof, the original greedy algorithm does not pin down the valuation distribution of the nonrecommended products. Table 2 is derived based on the procedure I define in the proof, which uniquely pins down the joint distribution of $(u_1, \ldots, u_K)$.

m_0, the consumer obtains a payoff of zero and has the lowest value v_1 for all the products that are not recommended; conditional on message m_1, the seller prefers to set the lowest possible price v_1, being indifferent to setting any prices in V. I modify ϕ^G so that it sends m_1 instead of m_0 with a small positive probability, in order to give the seller a strict incentive to set price v_1 at the new m_1. This does not lower the consumer's payoff. Finally, I use the seller's strict incentive to show that I can modify the disclosure rule to increase the consumer's payoff.

4.2 Firm Prefers Nondiscriminatory Pricing, Which Hurts Consumer

In the model of restricted disclosure, the seller is better off and the consumer is worse off under nondiscriminatory pricing for any prior value distributions (Theorem 1). An initial intuition might suggest that such a result no longer holds in the unrestricted disclosure model, because discriminatory pricing has a greater probability of trade (Proposition 2).

The following result, however, shows that the seller prefers to commit to not price discriminate, which hurts the consumer. To state the result, let R_{ND} and U_{ND} denote the equilibrium payoffs of the seller and the consumer under nondiscriminatory pricing. Similarly, let R_D and U_D denote the payoffs of the seller and the consumer, respectively, in any equilibrium under discriminatory pricing.

Theorem 2. *Suppose that the consumer can choose any disclosure rule and Assumption 1 holds. Generically, the seller is strictly better off and the consumer is strictly worse off under nondiscriminatory pricing: $R_{ND} > R_D$ and $U_{ND} < U_D$.*

As the proof shows, without Assumption 1, the result continues to hold with the strict inequalities replaced by weak ones.

Proof. Let ϕ^H denote any equilibrium disclosure rule under nondiscriminatory pricing, where the seller recommends the most valuable products (Proposition 1). Also, let ϕ^G denote the disclosure rule constructed in the proof of Proposition 2: ϕ^G maximizes the consumer's payoff among all the disclosure rules achieving efficient allocations. Under both disclosure rules, conditional on the event that the seller recommends product k, the value distribution of product k is equal to x^{MAX}, the distribution of $\max_k u_k$.

Let p^* denote the equilibrium price of each product under ϕ^H. One observation is that p^* also maximizes revenue under any posteriors drawn by ϕ^G. (This is because I construct ϕ^G from ϕ^H using the Bergemann et al.'s (2015) greedy algorithm.) In other words, under ϕ^G, the seller can achieve the highest revenue by posting price p^* upfront for all products. Denoting the optimal revenue under ϕ^H and ϕ^G by R_{ND} and R_G respectively, I obtain $R_{ND} = R_G$.

As ϕ^G is efficient, it can never consist of an equilibrium (Proposition 2). That is, the consumer's equilibrium payoff under discriminatory pricing (U_D) is strictly greater than the one from ϕ^G. Thus, the seller under discriminatory pricing is strictly worse off ($R_D < R_G = R_{ND}$). Finally, as the consumer's payoff is greater under ϕ^G than ϕ^H, I obtain $U_D > U_{ND}$. □

A rough intuition is as follows. As Proposition 2 shows, the consumer under discriminatory pricing obfuscates which product has the highest value. However, the seller might still benefit from discriminatory pricing as it makes trade more likely to occur when the consumer has low values. (Note that this effect is absent in the previous restricted disclosure model.) The reason why this argument fails is that, when $\mathcal{D}$ consists of all the disclosure rules, the consumer can disclose partial information about his willingness to pay to increase the probability of trade without increasing the seller's payoff. Thus, the seller prefers nondiscriminatory pricing, because while it leads to more accurate recommendations, the revenue loss from not being able to price discriminate is negligible.

4.3 Nondiscriminatory Pricing Can Enhance Efficiency

Which pricing regime is more efficient? The previous analysis suggests that nondiscriminatory and discriminatory pricing regimes have different advantages in increasing total welfare: Nondiscriminatory pricing leads to efficient recommendations, and discriminatory pricing leads to the greatest probability of trade; moreover, neither of them can achieve full efficiency under Assumption 1.

Indeed, which pricing regime is more efficient depends on the prior value distribution x_0 of each product and the number K of products. For instance, if $K = 1$, discriminatory pricing is always (weakly) more efficient.

The next result shows that, if there are a large number of products, nondiscriminatory pricing is more efficient. To focus on the interesting case in which it is *strictly* more efficient, I assume that x_0 does not put too much weight on the highest value $\max V$ of its support V. (x_0 is no longer required to have a finite support.)

Assumption 2. The optimal price at the prior distribution is strictly lower than the highest value of its support: $p(x_0) < \max V$.

Note that Assumption 2 holds whenever x_0 has a density. The proofs of the next proposition and lemma are in Appendix D.

Proposition 3. *Take any prior distribution x_0. Under nondiscriminatory pricing, as $K \to +\infty$, equilibrium total surplus converges to $\max V$. Under discriminatory pricing, if Assumption 2 holds, there is $\varepsilon > 0$ such that equilibrium total surplus is at most $v_N - \varepsilon$ for any K.*

The intuition is as follows. Under discriminatory pricing, with Assumption 2, the consumer can secure a positive information rent by hiding some information, which leads to inefficient recommendations. In contrast, under nondiscriminatory pricing, recommendations are efficient; furthermore, for a large K, the consumer's values for recommended products are almost degenerate at v_N. In this case, the seller can charge prices close to v_N upfront to extract most of the surplus, which in turn leads to almost efficient outcomes. Thus, if there are a large number of products, nondiscriminatory pricing enhances total welfare at the expense of consumer welfare.

Proposition 3 relies on the following lemma, which corresponds to the intuition.

Lemma 4. *Under nondiscriminatory pricing, as $K \to +\infty$, the seller's equilibrium payoff converges to $\max V$ and the consumer's equilibrium payoff converges to 0. Under discriminatory pricing, if Assumption 2 holds, there is $\underline{u} > 0$ such that the consumer's equilibrium payoff is at least $\underline{u}$ for any K.*

5 Extensions

5.1 Market for Personal Data

An institution within which consumers can sell their information has been paid attention as a market-based solution for privacy problems. In my model, such a "market for data" indeed improve the surplus of the consumer and the seller simultaneously.

To see this, consider the following extension: At the beginning of the game, the seller can offer to buy information: Formally, the seller chooses a pair of a disclosure rule $\phi \in \mathcal{D}$ and a transfer $t \in \mathbb{R}$. Then, the consumer decides whether to accept it. If he accepts, he reveals valuations according to ϕ and receives t; if he rejects, he can choose any disclosure rule but receive no transfer. In either case, this stage is followed by a product recommendation and a purchasing decision. Again, I consider two pricing regimes that differ in whether the seller can base prices on information.

How does this "market for data" affect equilibrium outcomes? Under nondiscriminatory pricing, it has no impact because the consumer is willing to disclose full information without any compensation. In contrast, under discriminatory pricing, the market for data may (weakly) improve the surplus of everyone. For example, if $\mathcal{D}$ consists of disclosure rule ϕ^* fully revealing valuations u, the seller offers (ϕ^*, t), where t makes the consumer indifferent between accepting and rejecting the offer. In equilibrium, the consumer accepts the offer and the seller engages in perfect price discrimination with efficient recommendations.

Importantly, in the new setting, not only the consumer but also the seller may prefer discriminatory pricing, because the market for data increases the seller's payoff from discriminatory pricing without changing revenue from nondiscriminatory pricing. Thus, the market for data could align the preferences of the seller and the consumer over information can be used for pricing.

5.2 A Model of Two-Sided Private Information

It is crucial to my results that the consumer chooses a disclosure rule without observing the values of products. Intuitively, this would be suitable if the consumer is initially uninformed of product characteristics necessary to calculate his willingness to pay. I provide a microfoundation for this assumption, focusing on the model of restricted disclosure in Section 3.

For ease of exposition, suppose that there are two products labeled as 1 and -1. At the beginning of the game, the consumer privately observes his *taste* $\theta \in \{1, -1\}$. Also, the seller privately observes *product characteristics* $\pi \in \{1, -1\}$. Each pair of (θ, π) is equally likely. Given a realized (θ, π), the consumer draws values of products $\theta \cdot \pi$ and $-\theta \cdot \pi$ from (the distributions of) $\max\{u_1, u_2\}$ and $\min\{u_1, u_2\}$, respectively.

The game proceeds as follows. *After privately observing* θ, the consumer (publicly) chooses a disclosure level δ: With probabilities δ and $1-\delta$, messages θ and $-\theta$ are realized, respectively. We interpret disclosure rules as any statistical information about his tastes. The seller observes δ and a realized message, and then recommends a product. As before, we can define two pricing regimes analogously.

Note that θ or π alone is not informative of product valuations, but (θ, π) is. This formulation captures situations in which sellers combine information abut a consumer's tastes and product characteristics to learn about preferences and give product recommendations.

This setting produces the same result as Theorem 1: Under nondiscriminatory pricing, the consumer with θ chooses a disclosure level of 1, by which the seller can recommend the best product with probability 1. Under discriminatory pricing, the consumer obtains a greater payoff by choosing δ^*, the equilibrium disclosure level derived in Section 3.

5.3 General Formulation of Horizontal Information

The model of restricted disclosure in Section 3 assumes that the most informative disclosure rule in $\mathcal{D}$ reveals whether $u_1 > u_2$ or $u_1 < u_2$. Here, I relax this assumption by providing a more general formulation of "horizontal information," under which all the results in Section 3 hold. Hereafter, x_0 denotes the prior CDF for the value of each product.

Consider disclosure rule $(\bar{M}, \bar{\phi})$ with the following properties: $\bar{M} = \{1, 2\}$, $\bar{\phi}(1|u_1, u_2) = \bar{\phi}(2|u_2, u_1)$ for any $(u_1, u_2) \in V^2$, and $\int_{u_2 \in V} \bar{\phi}(1|u_1, u_2) dx_0(u_2)$ is strictly increasing in $u_1 \in V$. These are satisfied, for example, if a disclosure rule sends message 1 with probability $h(u_1 - u_2)$ where $h(\cdot)$ is strictly increasing and satisfies $h(x) + h(-x) = 1$. (Disclosure level $\delta = 1$ in Section 3 corresponds to h being a step function.) Intuitively, the more valuable product $k \in \{1, 2\}$ is, the more likely $\bar{\phi}$ sends message k.

Because of the symmetry, the posterior distribution of u_k conditional on message j depends only on whether $k = j$. Let F_1 and F_0 denote the posteriors of u_k conditional on $j = k$ and $j \neq k$, respectively. The following result extends Lemma 2.

Lemma 5. *F_1 is greater than F_0 in the hazard rate order.*

Proof. Because $\int_{u_2} \bar{\phi}(1|u_1, u_2)dx_0(u_2)$ is increasing in u_1, for any $u^+ \geq u^-$, I obtain

$$\frac{\int_{u_1>u^+}\int_{u_2} \phi(1|u_1,u_2)dx_0(u_2)dx_0(u_1)}{1-x_0(u^+)} \geq \frac{\int_{u_1>u^-}\int_{u_2} \phi(1|u_1,u_2)dx_0(u_2)dx_0(u_1)}{1-x_0(u^-)}$$
$$\iff \frac{1-F_1(u^+)}{1-x_0(u^+)} \geq \frac{1-F_1(u^-)}{1-x_0(u^-)}.$$

Replacing $\phi(1|u_1, u_2)$ by $\phi(2|u_1, u_2)$,

$$\frac{1-F_0(u^+)}{1-x_0(u^+)} \leq \frac{1-F_0(u^-)}{1-x_0(u^-)}.$$

These inequalities imply

$$\frac{1-F_1(u^+)}{1-F_0(u^+)} \geq \frac{1-F_1(u^-)}{1-F_0(u^-)}$$

whenever the fractions are well-defined. Therefore, F_1 is greater than F_0 in the hazard rate order. □

Note that the proof of Theorem 1 only uses the fact that the value of the preferred product is greater than that of the less preferred product in the hazard rate order. Thus, I can conduct the same analysis assuming that the most informative disclosure rule in $\mathcal{D}$ is $(\bar{M}, \bar{\phi})$.

5.4 Alternative Interpretation: Online Advertising Platform

We can rewrite the model of restricted disclosure in Section 3 as a game between a consumer, an online advertising platform (such as Google or Facebook), and two advertisers. Advertisers 1 and 2 sell products 1 and 2, respectively. The consumer makes a purchasing decision after seeing an advert. Which advert the consumer sees depends on the outcome of an ad auction run by the platform.

In this interpretation, first, the consumer chooses a disclosure level δ (e.g., whether to accept a cookie) and visits the platform. Each advertiser $k \in \{1, 2\}$ chooses a price of product k and a bidding rule $b_k : \{1, 2\} \to \mathbb{R}$. Here, $b_k(j)$ is the bid by advertiser k for the impression of the consumer with a realized message $j \in \{1, 2\}$. I assume that advertisers choose bidding rules after observing δ and a realized message. If advertiser k wins the auction, the advert of product k is shown to the consumer. The consumer sees an advert, learns the value and price of the advertised product, and decides whether to buy it.

I show that the same result as Theorem 1 holds. Suppose that the consumer chooses disclosure level δ. First, If advertisers can base product prices on disclosure levels, each advertiser chooses price $p(\delta)$ and bidding rule b_k where $b_k(k) = p(\delta)[1 - \delta F^{MAX}(p(\delta)) - (1 - \delta)F^{MIN}(p(\delta))]$ and $b_k(j) < b_k(k)$ for $j \neq k$. The platform sets reserve price $p(\delta)[1 - \delta F^{MAX}(p(\delta)) - (1 - \delta)F^{MIN}(p(\delta))]$ to extract full surplus from advertisers. Given these strategies, the consumer sees the ad of his preferred product with probability δ. Second, if advertisers have to set prices without observing δ, the consumer chooses disclosure level 1 and each advertiser sets price $p(1)$. Thus, the consumer's disclosure decision and its welfare and price implications are identical as before.

6 Concluding Discussion

This paper studies consumers' privacy choices and their welfare and price implications. The key of the analysis is the following trade-off: Consumers may benefit from revealing about themselves, because sellers can offer product recommendations, which help consumers focus on a smaller subset of the huge variety of products. However, sellers may also use the information to price discriminate. This trade-off would be present in many real-life settings such as online shopping and buying cars or financial products.

I consider a model in which a consumer discloses information about his valuations to a multi-product seller. The consumer does not yet know his vector of values for the products and can evaluate only a small number of products relative to the number of available products. As a result, the seller can use disclosed information not only to extract surplus through pricing, but also to create surplus through product recommendations. Here, not only the consumer but the seller encounters a trade-off: Given available information, the seller profits from price discrimination; however, being able to price discriminate could affect the consumer's incentive to provide information.

The paper's contributions are threefold. One is to give an economic explanation of a somewhat puzzling observation in the Internet economy: Firms seem to not use individual data to price discriminate, and consumers seem to casually share their information with online sellers. The model explains this phenomenon as sellers' strategic commitment and consumers' best response. I show that this outcome robustly arises in two settings that differ in the information-disclosure technologies available to consumers.

The second contribution is to provide a framework for use in the design of privacy regulations. For instance, the model shows that nondiscriminatory pricing and the resulting full information revelation are consumer-suboptimal. Restricting the amount of information sellers can possess could benefit consumers, even if consumers are rational and can decide on their own what information to disclose.

The third contribution is to expand the theory of information disclosure by consumers. The model of unrestricted disclosure reveals that even with fine-grained control of information, consumers or a regulator cannot simultaneously achieve efficient price discrimination and efficient matching of products without sacrificing consumer welfare.

There are various interesting directions for future research. For example, the models could be extended to consider information sharing between sellers or the presence of data brokers, which is likely to add new policy implications. Moreover, this paper highlights the value of the equilibrium analysis to study consumers' privacy choices in the Internet economy. It would also be fruitful to study how consumers' information disclosure collectively affects welfare in other aspects of online privacy.

Appendix For Online Publication

A Proof of Lemma 1

Without loss of generality, suppose that message 1 is realized. If the seller recommends products 1 and 2 to the consumer, he draws values from $\delta F^{MAX} + (1-\delta)F^{MIN}$ and $\delta F^{MIN} + (1-\delta)F^{MAX}$, respectively. The former is greater than the latter in the first-order stochastic dominance, because $\delta > 1/2$ and F^{MAX} first-order stochastically dominates F^{MIN}. This implies that, under both pricing regimes, it is optimal for the seller to recommend product 1 because it maximizes the probability of trade given any prices. Finally, the seller's tie-breaking rule implies that only this recommendation strategy satisfies my equilibrium constraints. The tie-breaking rule matters if, for any δ, an optimal price is equal to the lowest possible value $(\inf V)$.

B Presence of "Negative Externality" with a Continuum of Consumers

I show that in the alternative interpretation of the model, information disclosure by a positive mass of consumers lowers the welfare of other consumers. To see this, note that if each consumer i chooses a disclosure level δ_i and the seller sets price p for each product, then the total revenue is given by

$$\begin{aligned}&\int_{i\in[0,1]} p[1-\delta_i F^{MAX}(p)-(1-\delta_i)F^{MIN}(p)]di\\ =&p[1-\bar{\delta}F^{MAX}(p)-(1-\bar{\delta})F^{MIN}(p)]\end{aligned}$$

where $\bar{\delta}=\int_{i\in[0,1]}\delta_i di$ is the average disclosure level. This implies that the optimal price under nondiscriminatory pricing is given by $p(\bar{\delta})$. If a positive mass of consumers disclose more information, $\bar{\delta}$ increases. This increases $p(\bar{\delta})$ and decreases the payoffs of other consumers who have not changed disclosure levels.

C Proof of Proposition 2

First, I show that there is a vertically efficient equilibrium. Take any (M^*, ϕ^*) which leads to a vertically inefficient allocation given the seller's best response. Let $x \in \Delta(V^K)$ denote a realized posterior at which trade may not occur.[29] Without loss of generality, suppose that product 1 is recommended at price v_ℓ at x. I show that there is another disclosure rule ϕ^{**} which gives a weakly greater payoff than ϕ^* to the consumer and achieves a strictly greater total surplus. Suppose that ϕ^{**} discloses whether the value for product 1 is weakly greater than v_ℓ or not whenever posterior x realizes, in addition to the information disclosed by ϕ^*. Let x^+ and $x^- \in \Delta(V^K)$ denote the posterior beliefs of the seller after the consumer discloses that the value for product 1 is weakly above and strictly below v_ℓ, respectively. Then, $x = \alpha x^+ + (1-\alpha)x^-$ holds for some $\alpha \in (0,1)$.

I show that ϕ^{**} weakly increases the payoff of the consumer. First, conditional on the event that the value is below v_ℓ, the consumer gets a greater payoff under ϕ^{**} than under ϕ^* because the consumer obtains a payoff of zero under ϕ^*. Second, I show that, conditional on the event that the

[29]Because $|V^K| < +\infty$, without loss of generality, I can assume $|M^*| < +\infty$. Then, each message is realized with a positive probability from the ex-ante perspective. Thus, there must be a posterior $x \in \Delta(V^K)$ which is realized with a positive probability and trade may fail to occur given x.

value is weakly above v_ℓ, the seller continues to recommend product 1 at price v_ℓ. To show this, suppose to the contrary that the seller strictly prefers to recommend another product m at price v_k. If $m = 1$, v_k is different from v_ℓ. Let $x_1^+ \in \Delta(V)$ and $x_m^+ \in \Delta(V)$ be the marginal distributions of u_1^i and u_m^i given x^+, respectively. Because the seller prefers recommending a product m at price v_k to recommending a product 1 at price v_ℓ, I obtain

$$v_k \sum_{j=k}^{K} x_m^+(v_j) > v_\ell \sum_{j=\ell}^{K} x_1^+(v_j),$$

which implies

$$v_k \sum_{j=k}^{K} \alpha x_m^+(v_j)+(1-\alpha)x_m^-(v_j) \geq v_k \sum_{j=k}^{K} \alpha x_m^+(v_j) > v_\ell \sum_{j=\ell}^{K} \alpha x_1^+(v_j) = v_\ell \sum_{j=\ell}^{K} \alpha x_1^+(v_j)+(1-\alpha)x_1^-(v_j).$$

The last equality follows from $x_1^-(v) = 0$ for any $v \geq v_\ell$. This contradicts that the seller prefers to recommend product 1 at price v_ℓ at x.

Consider the following mapping $\Phi : \mathcal{D} \to \mathcal{D}$: given any disclosure rule $\phi \in \mathcal{D}$, Φ chooses a posterior belief x induced by ϕ at which trade fails to occur with a positive probability. If there are more than one such belief, Φ chooses the posterior belief corresponding to the lowest price and the smallest index $k \in \mathcal{K}$.[30] $\Phi(\phi)$ is a disclosure rule which discloses whether the value for the recommended product is weakly greater than the price or not whenever posterior x is realized, in addition to the information disclosed by ϕ.

To show that there exists a vertically efficient equilibrium, take any equilibrium disclosure rule ϕ_0. Define $\Phi^1(\phi_0) = \Phi(\phi_0)$ and $\Phi^{n+1}(\phi_0) = \Phi(\Phi^n(\phi_0))$ for each $n \geq 1$. Because $|V^K| < +\infty$, there exists n^* such that $\Phi^{n^*} = \Phi^{n^*+1}$. Define $\phi^* := \Phi^{n^*}(\phi_0)$. By construction, ϕ^* gives a weakly greater payoff to a consumer than ϕ_0. Thus, it is an equilibrium under discriminatory pricing. Moreover, at each realized posterior, trade occurs with probability 1. Therefore, ϕ^* is a vertically efficient equilibrium.

Given our equilibrium notion, any equilibrium is vertically efficient. Indeed, if the consumer is indifferent between ϕ^* and ϕ^{**}, then the seller is strictly better off under ϕ^{**}. This implies that ϕ^*

[30] If this does not pin down a posterior uniquely, then I define Φ so that it first modifies ϕ by merging multiple beliefs at which the same product is recommended at the same price.

does not meet our equilibrium notion.

Next, I show that for a generic prior, any equilibrium is horizontally inefficient whenever $p(x_0) > v_1 := \min V$. While the consumer has private type u drawn from $x_0 \times \cdots \times x_0 \in \Delta(V^K)$, for the ease of exposition, I interpret the model as having the total mass one of consumers with mass $\prod_{k=1}^{K} x^*(u_k)$ having a valuation vector $u = (u_1, \ldots, u_K) \in V^K$.

Let $\mathcal{E} \subset \mathcal{D}$ denote the set of disclosure rules which lead to an efficient allocation for some best response of the seller. Take any disclosure rule $\phi^E \in \mathcal{E}$. Under ϕ^E, if the seller prefers to recommend product k, then $k \in \arg\max_{\ell \in \mathcal{K}} u_\ell$. Thus, if both ϕ^E and $\hat{\phi}^E$ achieve an efficient allocation, they only differ in terms of which product is recommended to consumers who have the same valuation for more than one product. I show that without loss of generality, I can focus on disclosure rules that recommend each product in $\arg\max u_k$ with equal probability whenever $|\arg\max u_k| \geq 2$.

To show this, take any $(M, \phi) \in \mathcal{E}$. Let $P \subset \mathcal{K}^{\mathcal{K}}$ be the set of the permutations of $\{1, \ldots, K\}$. Define ϕ^E as the following disclosure rule. First, ϕ^E publicly draws a permutation $\tau \in P$ uniformly randomly. Second, ϕ^E discloses information according to $\phi(u_{\tau(1)}, \ldots, u_{\tau(K)}) \in \Delta(M)$ for each realization $(u_1, \ldots, u_K)$. Then, from the ex-ante perspective, the consumer is recommended a product $k \in \arg\max u_j$ with probability $\frac{1}{|\arg\max u_j|}$.

I further modify ϕ^E to obtain $\phi^G \in \mathcal{E}$ which maximizes the consumer's payoff among the disclosure rules which lead to an efficient allocation. First, ϕ^G decomposes the prior x^* into K segments so that $x^* = x_1 + \cdots + x_K$. (As a disclosure rule, $\frac{1}{\sum_{n=1}^{N} x_k(v_n)} x_k \in \Delta(V^K)$ is a posterior belief that ϕ^G draws) Each x_k consists of $\frac{1}{|\arg\max u_j|} \cdot \prod_{j=1}^{K} x^*(u_j) \cdot 1_{\{k \in \arg\max u_j\}}$ mass of consumers with $u \in V^K$. Now, I apply the following procedure to each segment x_k. Without loss of generality, I explain the procedure for x_1. I apply the "greedy algorithm" in Bergemann et al. (2015) to x_1 with respect to the value for product 1 so that I can decompose x_1 into $x_1 = \alpha_1 x_1^{S_1} + \cdots + \alpha_{N_1} x_1^{S_{N_1}}$. Here, $S_1 = V$ and $S_{n+1} \supset S_n$ for $n = 1, \ldots, N_1 - 1$. Moreover, the marginal distribution of each $x_1^{S_n}$ with respect to u_1 is supported on $S_n \subset V$, and the seller is indifferent between charging any price for product 1 inside the set S_n if the value for product 1 is distributed according to $x_1^{S_n}$. In contrast to Bergemann et al. (2015), the consumer's type is K-dimensional. Thus, directly applying the algorithm does not pin down the valuation distribution for product $k \neq 1$ in each segment $x_1^{S_n}$. To pin down the distribution of values for product $k \neq 1$ in each $x_1^{S_n}$, I assume the following: whenever the algorithm picks consumers from x_1 to construct $x_1^{S_n}$, it picks consumers whose value for product 2 is lower. If this does not uniquely pin down the valuation vector to pick, it picks consumers whose value for product 3 is lower, and so on. In this way, the algorithm pins down a unique segmentation.

Consumer surplus under ϕ^G is weakly greater than under ϕ^E. This is because the segmentation created by the greedy algorithm maximizes consumer surplus and that the valuation distribution of each recommended product is identical between ϕ^E and ϕ^G. Also, under ϕ^G, the seller is willing to recommend product k to consumers in $x_k^{S_n}$ because $x_k^{S_n}$ only contains consumers such that $u_k \geq u'_k$ for any $k' \in \mathcal{K}$.

Next, I show the following: there exists a set $D \subset \Delta(V)$ satisfying the following: D has Lebesgue measure zero in $\mathbb{R}^N$, and for any prior $x_0 \in \Delta(V) \setminus D$, all consumers in $x_1^{S_{N_1}}$ constructed by the last step of the algorithm have the same value for product k. The proof of this part consists of two steps.

In the first step, take any subsets of V as $S_1 \supset S_2 \supset \cdots \supset S_{N_1}$ such that $|S_{N_1}| \geq 2$. Then, define

$$Y(S_1, \ldots, S_{N_1}) := \left\{ y \in \mathbb{R} : y = \sum_{n=1}^{N_1} \alpha_n x_1^{S_n}, \exists (\alpha_1, \ldots, \alpha_{N_1}) \in \Delta^{N_1 - 1} \right\}$$

where Δ^{N_1-1} is the $(N_1 - 1)$-dimensional unit simplex. Because $|S_{N_1}| \geq 2$, $N_1 \leq N - 1$. Thus, $Y(S_1, \ldots, S_{N_1})$ is a subset of at most $N - 1$ dimensional subspace, which has Lebesgue measure zero in $\Delta(V) \subset \mathbb{R}^N$. Define $\mathcal{S}$ as

$$\mathcal{S} = \{(S_1, \ldots, S_{N_1}) : \exists N_1 \in \mathbb{N}, V \supset S_1 \supset S_2 \supset \cdots \supset S_{N_1}, |S_{N_1}| \geq 2\}.$$

Let $\mathcal{Q}$ be the set of $x \in \Delta(V)$ such that consumers in $x_k^{S_n}$ constructed in the last step of the algorithm have different values for product k. I can write it as $\mathcal{Q} = \cup_{(S_1, \ldots, S_{N'}) \in \mathcal{S}} Y(S_1, \ldots, S_{N'})$. Because $|\mathcal{S}| < +\infty$ and each $Y(S_1, \ldots, S_{N_1})$ has measure zero, $\mathcal{Q}$ has Lebesgue measure zero as well.

In the second step, to show that there exists D with the desired property, consider a function φ which maps any prior $x \in \Delta(V)$ to the valuation distribution of product k conditional on the event product k is recommended under ϕ^E. Because the distribution does not depend on k, I consider

$k = 1$ without loss of generality. φ is written as follows.

$$\varphi(x) = K \cdot \begin{pmatrix} \frac{1}{K} x_1^K \\ x_2 \sum_{\ell=0}^{K-1} x_1^{K-1-\ell} x_1^\ell \cdot \frac{1}{\ell+1}\binom{K-1}{\ell} \\ x_3 \sum_{\ell=0}^{K-1} (x_1 + x_2)^{K-1-\ell} x_3^\ell \cdot \frac{1}{\ell+1}\binom{K-1}{\ell} \\ \vdots \\ x_N \sum_{\ell=0}^{K-1} (x_1 + \cdots + x_{N-1})^{K-1-\ell} x_N^\ell \cdot \frac{1}{\ell+1}\binom{K-1}{\ell} \end{pmatrix}.$$

φ is infinitely differentiable and its Jacobian matrix J_φ is a triangular matrix with the diagonal elements being positive as long as $x_n > 0$ for each $n = 1, \ldots, N$. Thus, $J_\varphi(x)$ has full rank if x is *not* in a measure-zero set

$$\{(x_1, \ldots, x_N) \in \Delta(V) : \exists n, x_n = 0\}. \tag{3}$$

By Theorem 1 of Ponomarev (1987), $\varphi : \mathbb{R}^N \to \mathbb{R}^N$ has the "0-property": the inverse image of measure-zero set by φ has measure zero. In particular, $D := \varphi^{-1}(\mathcal{Q})$ has measure zero. Thus, there exists a measure-zero set D such that for any $x \in \Delta(V) \setminus D$, all consumers in $x_k^{S_n}$ constructed in the last step of the algorithm have the same value for product k.

Consider the algorithm applied to product k. Recall that $x_k^{N_k}$ is the segment created at the last step. As I have shown, generically, all consumers in $x_k^{N_k}$ have the same value for product k. Let v^* denote the value. In equilibrium, consumers in $x_k^{N_k}$ obtain a payoff of zero given the firm's optimal price v^*. Moreover, if the optimal price at the prior is strictly greater than v_1 (i.e., $p(x_0) \geq v_2$), then $v^* > v_1$. Indeed, if $v^* = v_1$, then $v_1 \in S_n$ for $n = 1, \ldots, N_1$. This implies that v_1 is an optimal price for each $x_1^{S_n}$ and thus for $x_1 = \sum_{n=1}^{N_1} \alpha_n x_1^{S_n}$, which is a contradiction. To sum up, except for a Lebesgue measure zero set of priors, if the optimal price is strictly greater than v_1 under the prior, then consumers in $x_k^{N_k}$ obtain a payoff of zero given the firm's optimal price strictly above v_1.

Now, I modify ϕ^G to create a horizontally inefficient ϕ^I that yields consumer surplus strictly greater than ϕ^G, which completes the proof. To simplify the exposition, for any $S \subset \mathcal{K}$, let $v^*_S \in V^K$ denote a vector whose coordinate for each $k \in S$ is v^* and other coordinates are v_1. First, I replace ε mass of $v^*_{\{2\}}$ in the segment $x_2^{S_1}$ for product 2 created by the first step of the algorithm (applied for product 2) by the same probability mass of $v^*_{\mathcal{K}}$ in the segment $x_2^{S_{N_2}}$. Now, this does not affect consumer surplus generated from product 2. However, I now have ε mass of $v^*_{\{2\}}$ remaining. I pool this ε mass of $v^*(2)$ with segment $x_1^{S_1}$. Let $\hat{x}_1^1$ denote the segment created in this way. First, under $\hat{x}_1^1$, price v_1 is uniquely optimal because I add a positive mass of consumers having value v_1 to x_1^1, and price v_1 is optimal for x_1^1. Second, the seller is willing to recommend product 1 for $\hat{x}_1^1$, as long as ε is small. This follows from the fact that the seller strictly prefers to set price v_1 if the seller recommended product $k \neq 1$ for x_1^1. Indeed, at $x_1^{S_1}$, the firm's optimal price is v_1 no matter which product it recommends. While the seller is indifferent between recommending any prices of product 1, consumers who have value v_1 for all products but product 1 reject any price strictly greater than v_1. (Note that such consumers must be in the segment in $x_1^{S_1}$.) Thus, for product 2, price v_1 is uniquely optimal at x_1.

Because the seller strictly prefers to recommend product 1 at price v_1 compared to any other choices, for some $\delta > 0$, I can bring mass δ of $v^*_{\mathcal{K}}$ from $x_1^{S_{N_1}}$ who originally receives zero payoff. Let $\tilde{x}_1^1$ denote the segment created in this way. As long as δ is small, at $\tilde{x}_1^1$, the seller still recommends product 1 at price v_1. This strictly increases consumer surplus because consumers who obtain zero payoff at segment x_1^1 now obtain a positive payoff at $\tilde{x}_1^1$ without changing surplus accruing to other consumers. However, the resulting allocation is inefficient.

Therefore, for any disclosure rule which leads to an efficient allocation, there exists a horizontally inefficient disclosure rule which gives a strictly greater consumer surplus. This completes the proof.

D Proof of Lemma 4 and Proposition 3

Proof of Lemma 4. For each K, the consumer chooses $\delta = 1$ in the unique symmetric equilibrium under nondiscriminatory pricing because disclosure does not affect prices and increases his payoff through a better recommendation. Let F denote the CDF of the value for each product. Take any $\varepsilon > 0$. Suppose that the seller sets a nondiscriminatory price of $b - \varepsilon/2$ for each product. For a sufficiently large K, the probability $1 - F(p)^K$ that the consumer buys the recommended product goes to 1. Thus, there is $\underline{K}$ such that the seller's revenue is at least $b - \varepsilon$ if $K \geq \underline{K}$. This implies that the consumer's payoff is at most ε for any such K. This completes the proof of the first part.

To see that the consumer can always guarantee some positive payoff $\underline{u}$ under discriminatory pricing with Assumption 2, observe that the consumer can choose to disclose no information and obtain a payoff of $\int_{p_0}^{\max V} v - p_0 dF(v)$ where $p_0 < \max V$ is the optimal price given no disclosure, which is independent of K. $\square$

Proof of Proposition 3. First, the result under nondiscriminatory pricing follows from the previous result, as total surplus is greater than the firm's revenue.

Second, I show that total surplus under discriminatory pricing is uniformly bounded away from b. Suppose to the contrary that for any $n \in \mathbb{N}$, there exists K_n such that when the seller sells K_n products, some equilibrium under discriminatory pricing achieves total surplus of at least $\max V - \frac{1}{n}$. Then, I can take a subsequence $(K_{n_\ell})_\ell$ such that $K_{n_\ell} < K_{n_{\ell+1}}$ for any $\ell \in \mathbb{N}$. Next, I show that for any $p < \max V$ and $\varepsilon < 1$, there exists $\ell^* \in \mathbb{N}$ such that for any $\ell \geq \ell^*$,

$$\mathbf{P}_\ell(\text{the consumer's value for the recommended product} \geq p) \geq \varepsilon. \tag{4}$$

where $\mathbf{P}_\ell(\cdot)$ is the probability measure on the consumer's value for the recommended product in equilibrium of K_{n_ℓ}-product model. To show inequality 4, suppose to the contrary that there is some (p, ε) and a subsequence $(K'_m)_m$ of $(K_{n_\ell})_\ell$ such that the inequality is violated. Then, given any K'_m in this subsequence, the total surplus is at most $p\varepsilon + \max V(1-\varepsilon) < \max V$. This contradicts that the equilibrium total surplus converges to $\max V$ as $K'_m \to +\infty$.

Now, I use inequality 4 to show that the firm's equilibrium revenue converges to $\max V$ along $(K_{n_\ell})_\ell$. Take any $r < \max V$. If the seller sets price $\frac{r+\max V}{2}$, then for a sufficiently large ℓ, the consumer accepts the price with probability greater than $\frac{2r}{r+\max V} < 1$. That is, for a large ℓ, the seller's expected revenue exceeds r. Since this holds for any $r < \max V$, the seller's revenue converges to $\max V$ as $\ell \to +\infty$. This contradicts that the consumer's payoff is bounded from below by a positive number independent of K, which is shown in Lemma 4. $\square$

References

Acquisti, Alessandro and Hal R Varian (2005), "Conditioning prices on purchase history." *Marketing Science*, 24, 367–381.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris (2015), "The limits of price discrimination." *The American Economic Review*, 105, 921–957.

Braghieri, Luca (2017), "Targeted advertising and price discrimination online."

Bulow, Jeremy and John Roberts (1989), "The simple economics of optimal auctions." *Journal of Political Economy*, 97, 1060–1090.

Calzolari, Giacomo and Alessandro Pavan (2006a), "Monopoly with resale." *The RAND Journal of Economics*, 37, 362–375.

Calzolari, Giacomo and Alessandro Pavan (2006b), "On the optimality of privacy in sequential contracting." *Journal of Economic theory*, 130, 168–204.

Conitzer, Vincent, Curtis R Taylor, and Liad Wagman (2012), "Hide and seek: Costly consumer privacy in a market with repeat purchases." *Marketing Science*, 31, 277–292.

De Corniere, Alexandre and Romain De Nijs (2016), "Online advertising and privacy." *The RAND Journal of Economics*, 47, 48–72.

Dworczak, Piotr (2017), "Mechanism design with aftermarkets: Cutoff mechanisms."

Eliaz, Kfir and Ran Spiegler (2011), "Consideration sets and competitive marketing." *The Review of Economic Studies*, 78, 235–262.

Fudenberg, Drew and Jean Tirole (2000), "Customer poaching and brand switching." *RAND Journal of Economics*, 634–657.

Fudenberg, Drew and J Miguel Villas-Boas (2006), "Behavior-based price discrimination and customer recognition." *Handbook on Economics and Information Systems*, 1, 377–436.

Glode, Vincent, Christian C Opp, and Xingtan Zhang (2016), "Voluntary disclosure in bilateral transactions."

Iordanou, Costas, Claudio Soriente, Michael Sirivianos, and Nikolaos Laoutaris (2017), "Who is fiddling with prices?"

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian persuasion." *American Economic Review*, 101, 2590–2615.

Montes, Rodrigo, Wilfried Sand-Zantman, and Tommaso M Valletti (2017), "The value of personal information in markets with endogenous privacy."

Narayanan, Arvind (2017), "Online price discrimination: conspicuous by its absence."

Ponomarev, Stanislav P (1987), "Submersions and preimages of sets of measure zero." *Siberian Mathematical Journal*, 28, 153–163.

Roesler, Anne-Katrin (2015), "Is ignorance bliss? rational inattention and optimal pricing."

Roesler, Anne-Katrin and Balázs Szentes (2017), "Buyer-optimal learning and monopoly pricing." *American Economic Review*, 107, 2072–2080.

Salant, Yuval and Ariel Rubinstein (2008), "(a, f): choice with frames." *The Review of Economic Studies*, 75, 1287–1296.

Shaked, Moshe and George Shanthikumar (2007), *Stochastic orders*. Springer Science & Business Media.

Sun, Yeneng (2006), "The exact law of large numbers via fubini extension and characterization of insurable risks." *Journal of Economic Theory*, 126, 31–69.

Taylor, Curtis R (2004), "Consumer privacy and the market for customer information." *RAND Journal of Economics*, 631–650.

Topkis, Donald M (1978), "Minimizing a submodular function on a lattice." *Operations Research*, 26, 305–321.

Villas-Boas, J Miguel (1999), "Dynamic competition with customer recognition." *The Rand Journal of Economics*, 604–631.

Villas-Boas, J Miguel (2004), "Price cycles in markets with customer recognition." *RAND Journal of Economics*, 486–501.