Online Privacy and Information Disclosure by Consumers

Shota Ichihashi*

March 13, 2019

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 use information for pricing 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, total surplus may be lower if the seller can base prices on information.

1 Introduction

I study the welfare and price implications of consumers' privacy in online marketplaces. Online sellers can observe detailed information about consumers, such as their browsing histories, pur-

*Bank of Canada, Email: shotaichihashi@gmail.com. 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.

chases, and characteristics; however, consumers can often affect whether and to what extent this information is revealed. For instance, they can disable cookies to hide their web-browsing activities, or they can use their social networking accounts to log in to online shopping websites. For policymakers, information revelation by consumers is an important consideration in formulating policies concerning online privacy.

The paper considers the following economic trade-off: The benefit for consumers of disclosing information is that sellers can recommend or advertise products that are directly relevant. 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 sellers could also 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, the consumer chooses a disclosure rule, which determines what the seller learns about the consumer's value for each product. After learning about the values, the seller recommends one of K products. Finally, the consumer observes the value and the price of the recommended product and decides whether to buy it.

In the model, the consumer discloses information without observing his values, which is in line with Bayesian persuasion (Kamenica and Gentzkow, 2011). In online marketplaces, it is often difficult for consumers to determine which item in the set of available products is most appropriate for them; however, sellers can often do this using personal data. For example, sellers might analyze browsing histories by using their knowledge of the products' characteristics, the prior experiences of other consumers, and their computing power. Sellers can then map a given consumer's data into estimates of his values for products. Then, even though the consumer himself cannot evaluate all products, his privacy choices affect what sellers can learn about the values.

A novel aspect of the paper is that I consider two settings that differ in the timing at which the seller sets prices. Under the *discriminatory pricing* regime, the seller sets prices after learning about the consumer's values; under the *nondiscriminatory pricing* regime, the seller sets prices upfront without observing the information. A comparison of the two pricing regimes enables us to study the value of the seller's *commitment to not use the consumer's information for pricing.*

In the classical theory of (third degree) price discrimination, it is profitable for a seller to set prices on the basis of information about consumers' willingness to pay. Also, whereas the impact of price discrimination on consumers is generally ambiguous, consumers are worse off if the seller can tailor prices on fine-grained information about consumers' values.

My model generates starkly different predictions. The first main finding is that *the seller is better off by committing to not use the consumer's information for pricing.* The key is the consumer's endogenous disclosure: By making such a commitment, the seller can induce the consumer to disclose more information, which enables the seller to make more accurate recommendations. This shifts up the consumer's demand for recommended products and increases revenue. Such a commitment has an obvious downside that the seller cannot tailor prices on information. I provide conditions on the set of available disclosure rules under which the seller's benefit from accurate recommendation dominates the potential loss. For example, the result holds if the consumer can disclose *any* information about his value of each product. The result gives a potential explanation of an observed puzzle: "The mystery about online price discrimination is why so little of it seems to be happening" (Narayanan, 2013). Namely, price discrimination by online sellers seems to be uncommon despite their potential ability to use consumers' personal data to tailor prices.[1]

The second main finding is that *the consumer is worse off if the seller commits to not use information for pricing*. A key observation is that, if the seller does not make such a commitment, then the consumer can induce the seller to set lower prices by withholding information about which product is most valuable to the consumer. Although this leads to less accurate product recommendation, the consumer's gain from low prices can exceed the loss from potential product mismatch. In contrast, if the seller commits to prices in advance, then the consumer misses the opportunity to influence prices by strategically conceal information. This makes the consumer worse off compared to without the seller's commitment.

Finally, I show that equilibrium is inefficient under a mild distributional assumption, regardless of whether the seller makes the aforementioned commitment. The result contrasts with the single-product case of Bergemann et al. (2015), in which equilibrium is fully efficient whenever

[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)."

the seller can base prices on information. In my model, the seller's ability to tailor prices on information discourages the consumer's disclosure, and this leads to inefficiency due to inaccurate recommendation. The proof is based on a "constrained" Bayesian persuasion problem, in which the consumer chooses 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 various offline transactions. For example, consider a consumer looking for a car. The consumer may talk to a salesperson and reveals some information—such as 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. My result points out that eliminating this trade-off is precisely what car dealers should do: By committing to prices upfront, car dealers can encourage consumers to disclose information; at the same time, dealers can set a relatively high price for each product to extract surplus created by accurate recommendations. Thus, the result clarifies one benefit of "no-haggle" policies adopted by firms such as Fiat, Tesla, and Toyota (Zeng et al., 2013, 2016).

The remainder of the paper is organized as follows. In Section 2, after discussing related work, I present the baseline model. 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 price discrimination can hurt efficiency. In Section 5, I discuss several extensions including markets for personal data. Section 6 concludes.

1.1 Related Work

This paper relates to three strands of literature: The literature on monopoly price discrimination, the economics of privacy, and information design. In terms of modeling, one related work is Bergemann et al. (2015), who consider a single-product monopoly pricing in which a seller has additional information about a consumer's value. I consider a multi-product seller with product recommendations, which renders information useful not only for pricing but also for improving product match quality. These components lead to different welfare consequences of information disclosure and price discrimination. In contrast to Bergemann et al. (2015), who characterize the entire set of attainable surplus, I focus on the case where the consumer discloses information to maximize his own payoff.

Within the literature of monopoly pricing, this paper relates to works on third degree price discrimination. Starting with the work of Pigou (1920), the literature examines what happens to consumer surplus, producer surplus, and total surplus as a market is segmented.[2] Relative to this literature, my model differs in two ways. First, market segmentation is endogenous. This is especially relevant when a seller uses consumer information to segment a market but consumers can influence how much information is available to the seller. Second, the seller can use consumer information to recommend products. In this case, finer segmentation enables the seller to recommend a more relevant product to consumers in each segment, and this can shift up demands of all consumers. These two points are crucial to why the seller can be better off if she cannot tailor prices to segments.

The economic forces of this paper are reminiscent of those in the durable goods monopoly problem, such as Stokey (1979, 1981). In that setting, a seller prefers to commit to not price discriminate based on the timing of purchase. This is because if the seller engages in such discrimination, consumers would delay the timing of their purchase so as to lower profit overall. In my model, the seller prefers to commit to not lower prices even if the consumer discloses information that makes it (ex post) optimal for the seller to do so. That way, the consumer will disclose as much information as possible about which product is most valuable. Despite this high level connection, I argue that this paper provides novel insights. First, the effect of sellers' commitment power in how consumers reveal information is fundamental and economically important. This economic force

[2] For more recent papers, see, for example, Aguirre et al. (2010), Chen and Schwartz (2015), and Cowan (2016).

relates not only to the recent discussion on consumer privacy but also to broader settings such as markets for cars and houses as well as bargaining between workers and employers.[3] Second, the result that the seller's commitment may increase total welfare is unique to the present model: In the durable goods monopoly problem, the seller's inability to commit to prices leads to efficient outcomes if the length of the time period is short.

The paper also relates to the economics of privacy literature. As a growing number of transactions are based on data about consumers' behavior and characteristics, recent papers 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.

Hidir and Vellodi (2018) consider a model in which a buyer communicates his preferences with the seller, who uses the information to tailor product offerings and prices. Their focus and formulation differ from mine in at least two ways. First, their main focus is to understand the consumer's trade-off between better product match and lower prices in the face of price discrimination. In contrast, on top of this trade-off, I ask what could lead sellers to commit to not use consumer data for price discrimination. Thus, Hidir and Vellodi (2018) primarily focus on discriminatory pricing whereas I compare two pricing regimes. Second, the intended applications are different. The main application of Hidir and Vellodi (2018) is a situation where a consumer looks for a product knowing which product he wants to buy. This is suitable, for example, when a consumer performs a search query on an e-commerce website. In contrast, I consider a situation where consumers make privacy choices without knowing exactly which product he wants and how his data will be used by sellers. For example, when a consumer decides whether or not to reveal his browsing activities (by accepting cookies), he may not have a particular product in mind; however, the consumer expects that by sharing more data, he will likely see more relevant products as product recommendations or targeted ads. Consqeuntely, these papers formulate information disclosure differently: Hidir and Vellodi (2018) adopt a cheap talk setting, whereas I employ a formulation à la Bayesian persuasion,

[3]See the discussion in Section 2.

where a consumer decides what information to disclose without knowing values.[4]

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 taste to sellers. Targeting helps consumers find appropriate products at low cost, but it can hurt them because of price discrimination. A similar trade-off arises in De Corniere and De Nijs (2016), who study a platform's choice of disclosing consumers' preferences to advertisers. In contrast to these papers, I consider the seller's commitment of not using information for pricing. Also, these papers assume that a consumer reveals either full or no information, but I consider a consumer who can choose a general experiment of his values. This enables me to study what kind of information is 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 and a disclosure rule. A mechanism elicits an agent's private type and a disclosure rule 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 some non-degenerate probability distribution supported on a compact set $V \subset \mathbb{R}_+$.[5] Let $u := (u_1, \ldots, u_K)$ denote the vector of values.

[4]Because of their cheap talk setting, Hidir and Vellodi (2018) offer a careful analysis of the multiplicity of equilibria, which is not a focus of this paper.

[5]See Remark 2 for how the results extend to correlated values.

The consumer's preferences are quasi-linear: If he buys product k at price p, his ex post payoff is $u_k - p$. Otherwise, the payoff is zero. The seller's payoff is her 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 $\mathcal{D}$.[6] 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 all probability distributions over M. After the consumer chooses a disclosure rule, 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, $\mathcal{D}$ corresponds to the set of consumers' privacy choices, such as whether or not to share one's browsing history. As in Section 4, if $\mathcal{D}$ consists of all disclosure rules, information disclosure takes the form of Bayesian persuasion (Kamenica and Gentzkow, 2011).

Next, I describe the seller's pricing. I consider two games that differ in the timing at which the seller sets prices. 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.[7] Note that in this case, the seller not only does not base prices on a realized message m but also does not base prices on a disclosure rule ϕ.[8] As in the introduction, nondiscriminatory pricing captures the seller's commitment of not using information to set prices. (The next subsection discusses why I focus on these particular regimes.)

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,

[6]I will impose more structure on $\mathcal{D}$ in Sections 3 and 4.

[7]Alternatively, I can 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 the equilibrium if the consumer can only reveal information about which product has the highest value as in Section 3. In contrast, it could change the equilibrium if the consumer can disclose information in an arbitrary way as in Section 4. 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.

[8]For example, an e-commerce firm that adopts this regime sets prices based on neither browsing histories nor whether consumers share their browsing histories.

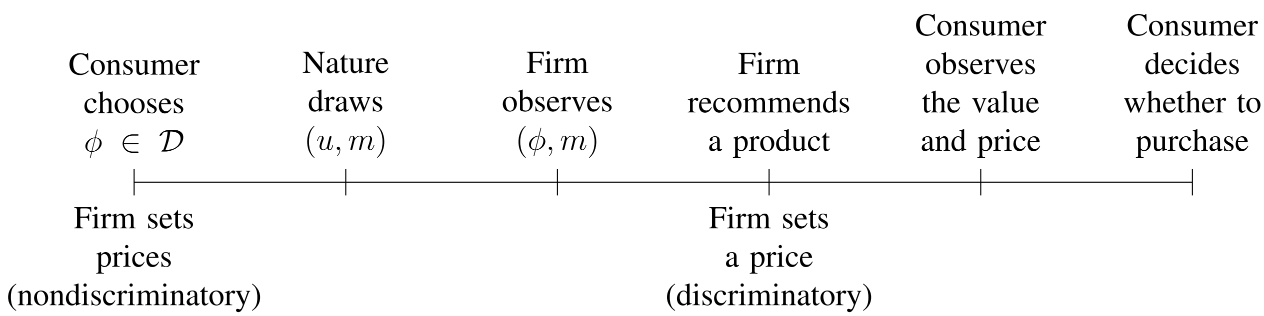

Figure 1: Timing of moves under each pricing regime

the seller simultaneously sets the price of each product. Then Nature draws the consumer's values 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 pure-strategy perfect Bayesian equilibrium (PBE) with four restrictions. First, at each information set where the seller recommends a product (and sets a price under discriminatory pricing), the seller forms her belief about values u according to the prior distribution, disclosure rule (M, ϕ), realized message m, and Bayes' rule.[9] Second, the seller breaks a tie in favor of the consumer whenever she is indifferent among multiple prices or recommendations. Third, under nondiscriminatory pricing, I focus on equilibrium in which each product has the same price. Fourth, if there are still multiple equilibria that give the consumer identical expected payoffs (after applying previous restrictions), I focus on equilibria that maximize the seller's payoff. This condition eliminates the multiplicity of equilibria due to the consumer's indifference among disclosure rules, which is not the main focus of the paper. Hereafter, "equilibrium" refers to PBE that satisfies these restrictions.

As 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 and 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

[9] In particular, this uniquely pins down the seller's beliefs after the consumer deviates from the equilibrium disclosure rule.

to reveal information and sellers' pricing strategies in those markets.

Indeed, the application is not even restricted to buyer-seller interactions. Consider the following situation: An employer assigns her 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. For instance, he might communicate what kind of tasks he is good at. The employer wants to maximize the value from a task minus wage payment, whereas the worker wants to maximize wage minus costs of completing a task. It is easy to show that this model is mathematically equivalent to the baseline model. In this case, two pricing regimes correspond to whether the wage is set contingent on the revealed information.

2.1 Discussion of Modeling Assumptions

Before proceeding to the analysis, I provide a detailed discussion of modeling assumptions.

Pricing Regimes

It would be illustrative to interpret the two pricing regimes in terms of the seller's commitment power. First, discriminatory pricing corresponds to the case where the seller has no commitment power. This is a tractable benchmark in which the consumer faces the trade-off between accurate recommendations and low prices. Second, nondiscriminatory pricing captures the situation where the seller *commits to not use the consumer's information to set prices.* Note that this is not the only form of commitment that yields a greater revenue than discriminatory pricing. For example, the seller could do even better if she could commit to any contingent schedule of prices as a function of disclosure rules and realized messages.[10] I exclude this kind of commitment power and focus on the commitment of not using information for pricing. This is because the latter seems more practical and relevant to policymakers or firms that consider whether to regulate or engage in personalized pricing.

Information Disclosure

The model is agnostic about what personal data and privacy choices the cosnumer has. Instead, I formulate the consumer's choice set as a set of Blackwell experiments about his values. This

[10]If the seller could commit to a mechanism that maps a disclosure rule and a realized message into prices, then the seller would extract full surplus with a mechanism that sets price $+\infty$ for all products if and only if the consumer does not disclose full information.

formulation calls for several implicit assumptions; for example, the consumer understands how his privacy choice affects the seller'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 values. 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 his value of each product. 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 the products' characteristics. It would also be natural to assume that the consumer cannot manipulate message realizations ex post, as consumers or regulators typically set disclosure rules upfront and incentives to distort or misrepresent one's browsing history or characteristics seem to be less relevant.

Product Recommendation and Purchase Decision

There are also two substantial assumptions on product recommendation and purchasing decision. First, the seller recommends a single product and the consumer decides whether to buy it. In other words, the consumer cannot purchase products not recommended by the seller. This formulation captures situations where there are constraints on how many products can be marketed to a given consumer. Such constraints are natural if the variety of available products is large but the consumer has a limited ability to assess products because of his time or cognitive constraints.[11] It is important to note that this assumption is not a result of the seller's revenue-maximization. In other words, the seller is assumed to recommend a single product (as opposed to recommending multiple products) not because it is optimal for her to do so, but because such an assumption concisely captures the consumer's limited attention. Indeed, if the consumer could evaluate all products and choose which product buy, then the seller would prefer to offer and price all products.[12] The model

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

[12] If the consumer can evaluate all products and the seller can offer all products, then we obtain the following results: Under nondiscriminatory pricing, the consumer always buys the highest value product $k^* \in \arg\max_k u_k$, and information disclosure and recommendation play no role. In particular, even if the seller could restrict the number of products offered to the consumer, she prefers to offer all products to ensure that the consumer can find the highest value product. Under discriminatory pricing, if the consumer can only disclose information about product rankings as in Section 3, then the equilibrium coincides with nondiscriminatory pricing. For a general set of disclosure rules $\mathcal{D}$,

excludes such a mechanism to incorporate the consumer's limited attention.

Second, the consumer observes the value of 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.

Production Costs

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.

3 Restricted Information Disclosure

To illustrate the main idea in a simple way, this section imposes the following structure on the baseline model. First, assume that the seller sells two products ($K = 2$). Then, identify $\mathcal{D}$ with $[1/2, 1]$. Each $\delta \in [1/2, 1]$ is called a *disclosure level*, which represents the amount[13] of information that the consumer discloses about *which product is more valuable*. As in Figure 2, each δ corresponds to a disclosure rule that draws either message 1 or 2: It draws message $m \in \{1, 2\}$ with probability δ if product m is strictly more valuable than the other product, and it draws messages 1 and 2 with equal probability if products 1 and 2 have the same value. A greater δ implies that the seller can learn more accurately about which product has a higher value.

3.1 Equilibrium Analysis

I first clarify the consumer's economic trade-off between accurate recommendations and lower prices. Then, I show that the seller prefers to commit to prices upfront and this commitment

the consumer might reveal some information to induce the seller to set lower prices when he has low values.

[13] Formally, a greater δ is associated with a more Blackwell-informative disclosure rule.

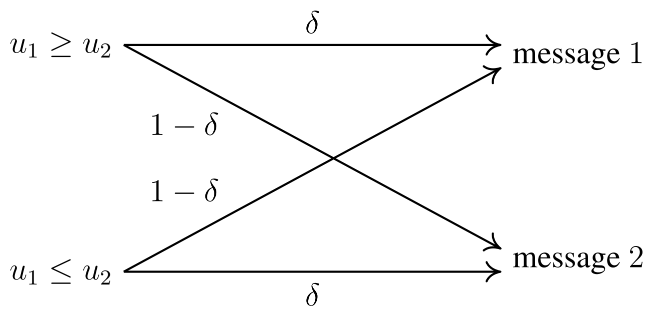

Figure 2: Disclosure rule for $\delta \in [1/2, 1]$

lowers the consumer's welfare.

First, consider the seller's recommendation strategy. For a given price, the seller prefers to recommend the product that is more likely to have a higher value, because it maximizes the probability of purchase. This leads to the following lemma. See Appendix A for the proof.

Lemma 1. *Fix a pricing regime and take any equilibrium. If the consumer chooses a disclosure level* $\delta \in (1/2, 1]$*, then after observing message* $k \in \{1, 2\}$*, the seller recommends product* k*.*

Lemma 1 implies that, in equilibrium, a disclosure level (δ) is equal to the probability of the seller's recommending the most valuable product (see Figure 2). Thus, greater disclosure makes it more likely that the consumer sees his preferred product.

Next, consider how information disclosure affects prices. From the consumer's perspective, disclosure has no impact on prices under nondiscriminatory pricing. Thus, I focus on discriminatory pricing for now. Let F^{MAX} and F^{MIN} denote the cumulative distribution functions of $\max(u_1, u_2)$ and $\min(u_1, u_2)$, respectively.[14] If the consumer chooses a disclosure level δ and message $k \in \{1, 2\}$ is realized, then the consumer's value for the recommended product (i.e., product k) is drawn from $\delta F^{MAX} + (1-\delta)F^{MIN}$. Thus, to study how information affects prices, it is crucial to understand how $\delta F^{MAX} + (1-\delta)F^{MIN}$ depends on δ.

I show that greater disclosure leads to a lower hazard rate of the value distribution of the recommended product. For the sake of generality, I employ the following definition, which does not require distributions to have densities.[15]

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

[15]I employ the hazard rate order to clarify how information disclosure affects prices. First-order stochastic dom-

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 $z \in (-\infty, \max(s_1, s_0))$.[16] Here, s_0 and s_1 are the right endpoints of the supports of G_0 and G_1, respectively.[17]

The next lemma follows from Theorem 1.B.26 of Shaked and Shanthikumar (2007).

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 some product at price p. Conditional on this event, how likely is the consumer to stop buying the product if the seller marginally increases the price by ε? If the product is the consumer's preferred one, whose value is $\max(u_1, u_2)$, then 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 $\min(u_1, u_2)$, then he stops buying whenever one of u_1 and u_2 is below $p+\varepsilon$. Thus, the consumer is more likely to stop buying the less preferred product than the more preferred product. This implies that the value distribution F^{MAX} has a lower hazard rate than F^{MIN}.

As the intuition suggests, the hazard rate order relates to demand elasticity. As in Bulow and Roberts (1989), the demand curve for a CDF F is given by $D(p) = 1 - F(p)$, and thus the demand elasticity is $-\frac{d \log D(p)}{d \log p} = \frac{f(p)}{1-F(p)}p$. Then, a CDF F_1 is greater than a CDF 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 F_0. Thus, Lemma 2 states that the consumer's demand for the more preferred product is less elastic.

Recall that, by Lemma 1, a disclosure level δ is equal to the probability that the seller recommends the consumer's preferred product. Thus, a greater δ implies that the consumer is more likely to have less elastic demand for the recommended product. As a result, the seller prefers to set a higher price for it. The following result formalizes this intuition. Appendix B contains the proof.

inance, which is weaker than the hazard rate order, is not sufficient for this, because it has no implications on the behavior of the monopoly price. 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.

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

[17] Suppose that G_0 and G_1 have densities g_0 and g_1. By differentiating $\log \frac{1-G_1(z)}{1-G_0(z)}$ in z, we can show that the above definition is equivalent to

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

Lemma 3. *Consider discriminatory pricing. Let $p(\delta)$ denote the equilibrium price for the recommended product given a disclosure level δ.*[18] *Then, $p(\delta)$ is increasing in δ.*

It is crucial to note that $p(\delta)$ is also the equilibrium price under *nondiscriminatory pricing* when the seller anticipates that the consumer chooses δ. The reason is as follows. Under nondiscriminatory pricing, the seller sets a price of (say) product 1 to maximize the expected revenue conditional on recommending product 1, which is the only event where the price of product 1 affects revenue. This maximization problem is identical with the seller's pricing problem when she recommends product 1 after observing δ under discriminatory pricing, whose solution is $p(\delta)$.

This observation and the previous results lead to the main result of this section. Appendix C contains the proof.

Theorem 1. *In any equilibrium,*[19] *the seller obtains a higher payoff and the consumer obtains a lower payoff under nondiscriminatory pricing than under discriminatory pricing. Under nondiscriminatory pricing, the consumer chooses the highest disclosure level ($\delta = 1$) and is charged higher prices.*

The intuition is as follows. Under nondiscriminatory pricing, the consumer prefers the highest disclosure level $\delta = 1$, because more disclosure leads to better recommendations without affecting prices. Now, the seller anticipates the choice of $\delta = 1$ and resulting accurate recommedendations, which makes the consumer's demand less elastic. Then, the seller sets a price of $p(1)$ for each product upfront. In contrast, under discriminatory pricing, the consumer can influence prices by strategically concealing information. In particular, he can always choose $\delta = 1$ to induce the equilibrium outcome of nondiscriminatory pricing, because $p(\cdot)$ describes the optimal price under both pricing regimes. Thus, the consumer is never worse off under discriminatory pricing.

In contrast, the seller prefers committing to prices upfront. First, the optimal price depends on δ but not on a realized message. Then, the seller is indifferent between the two pricing regimes if the consumer chooses the same disclosure level δ. Moreover, the revenue is increasing in δ, because greater disclosure leads to more accurate recommendations and shifts up the consumer's demand

[18] Appendix H proves the existence of the lowest optimal price $p(\delta)$.

[19] The equilibrium payoffs of the consumer and the seller are unique. Indeed, given the equilibrium restriction, the equilibrium is unique up to which product the seller recommends when the consumer chooses $\delta = 1/2$.

for recommended products. Thus, the seller prefers nondiscriminatory pricing, which guarantees that the consumer chooses $\delta = 1$.

It is important to note that the first part of Theorem 1—the seller is better off and the consumer is worse off under nondiscriminatory pricing—does *not* use the comparative statics in Lemma 3. Indeed, the welfare comparison for the seller relies only on that the consumer chooses $\delta = 1$ under nondiscriminatory pricing (by Lemma 1), and the welfare comparison for the consumer depends on that the consumer under discriminatory pricing can replicate the equilibrium outcome of nondiscriminatory pricing by choosing $\delta = 1$. Nonetheless, the comparative statics on prices illuminates the economic force that makes it profitable for the consumer to withhold information under discriminatory pricing: The consumer can induce the seller to set lower prices by concealing some information, because a less informed seller, who can only make noisy recommendations, faces more elastic demand. Thus, withholding information can benefit the consumer through lower prices at the expense of recommendation accuracy.[20]

Theorem 1 gives an economic explanation of the observed puzzle: Online sellers seem to not use individual data to price discriminate,[21] 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 information disclosure, and consumers' best response to such a commitment. The theorem further suggests that such a situation might not be desirable for consumers, because they could be better off were sellers less informed about their preferences.

Theorem 1 also has policy implications: Consumers may benefit from regulations that restrict the amount of information sellers can expect to acquire. To see this, suppose that the seller can commit to prices, under which the consumer chooses the greatest disclosure level. Relative to this

[20]Theorem 1 does not tell us when the consumer *strictly* prefers to withhold information under discriminatory pricing. Unfortunately, I have not been able to identify a natural condition under which the seller is *strictly* better off and the consumer is strictly worse off under nondiscriminatory pricing. Later, I address this issue in two ways: I show that the strict welfare comparison holds under a mild distributional assumption if (i) the consumer can disclose any information about the values, or (ii) there are a sufficiently large number of products.

[21]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 can be other explanations for sellers not price discriminating. For instance, sellers may think that price discrimination would infuriate consumers who have fairness concerns. The rational explanation in Theorem 1 suggests that, even if sellers can frame personalized pricing in a way that consumer backlash is less likely to occur, sellers may still find it profitable to refrain from price discrimination.

situation, the consumer 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. With this restriction, the consumer chooses disclosure level δ^* and obtains a greater payoff than without the regulation.[22]

One may think that the main result is driven by the particular restriction on what information the consumer can reveal. This is partly true and actually an important insight: If the consumer's endogenous disclosure is crucial for the seller to give accurate recommendations, then the seller can be better off by committing to not use information for pricing. Focusing on disclosure rules parametrized by δ is a simple way to capture such a situation. In contrast, if endogenous disclosure is not important—for example, if the seller knows the consumer's willingness to pay at the outset—then the seller may prefer discriminatory pricing, which hurts the consumer.[23]

Nonetheless, the current restriction on $\mathcal{D}$ is one of various conditions under which the welfare comparisons in Theorem 1 hold. For example, $\mathcal{D}$ can be any set of disclosure levels such as $\mathcal{D} = \{0.5, 0.8\}$, which consists of, say, enabling cookies ($\delta = 0.8$) and disabling cookies ($\delta = 0.5$). Subsection 5.1 discusses more general conditions on $\mathcal{D}$ under which the welfare comparisons hold. In the next section, I establish the (strict) welfare comparison result assuming that the consumer can choose *any* disclosure rule.

3.2 Theorem 1 as a Tragedy of the Commons

We can interpret the current setting as a model with a continuum of consumers. This interpretation enables 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 unit mass of consumers, each of whom chooses a disclosure level. The value of each product is independent across consumers.[24] Under nondiscriminatory pricing, *after* observing the disclosure level and realized message of each consumer, the seller sets

[22]the seller is indifferent between two pricing regimes.

[23]We can capture such a situation by assuming that $\mathcal{D}$ consists only of the disclosure rule that reveals the exact value vector. That is, the consumer does not have a choice of concealing information. The preferences of the seller and the consumer over pricing regimes could be aligned in other contexts. For instance, if there is a single product and the value is drawn from $U[0, 1]$, then both the seller and the consumer prefer discriminatory pricing if the consumer can only disclose whether $v \geq 1/2$ or not.

[24]The independence of value 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.

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.

The equilibrium prediction in Theorem 1 persists: Under nondiscriminatory pricing, each consumer chooses the highest disclosure level and obtains a lower payoff, and the seller sets higher prices upfront. To see this intuitively, consider an equilibrium under nondiscriminatory pricing. Each consumer $i \in [0, 1]$ chooses a disclosure level δ_i *taking prices as given*, because the choice of a single consumer in a large population has no impact on prices. Thus, it is optimal for every i to choose $\delta_i = 1$, following which the seller sets a price of $p(1)$.

According to this interpretation, we can view Theorem 1 as a tragedy of the commons: If some (positive mass of) consumers disclose more information under nondiscriminatory pricing, then the seller prefers to increase prices as she can offer accurate recommendations to a greater fraction of consumers. But then, *all* consumers face higher prices. That is, greater disclosure by some consumers lowers the welfare of other consumers through higher prices. As consumers do not internalize this negative impact, they disclose full information, although they could be better off by collectively withholding information. This problem does not arise under discriminatory pricing, because each consumer internalizes the impact of disclosure on prices he has to pay. Appendix D formalizes this observation.

Remark 1 (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 many applications.

First, e-commerce firms, such as Amazon and eBay, 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. In such cases, we cannot conclude 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, due to recent legislation in the EU. 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.[25] In this case, consumers need not be highly sophisticated to figure out what "disclosure rules" are available to them.

Remark 2. Theorem 1 is robust to a variety of extensions.

Correlated Values: Even if u_1 and u_2 are correlated, the welfare comparison in Theorem 1 holds as long as (u_1, u_2) is drawn from an exchangeable distribution. The comparative statics on prices (Lemma 3) holds if this joint distribution has a multivariate hazard rate that 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.

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.[26] This does not change the conclusion that the consumer is worse off under nondiscriminatory pricing. Characterizing an equilibrium 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 for consumer i is $\Delta(\delta_i, \bar{\delta})$, which is an increasing function of i's disclosure level δ_i and the average disclosure level of the population $\bar{\delta} = \int_{i \in [0,1]} \delta_i di$. This captures the idea that the seller can learn about i's preferences from information disclosed by others. In this case, Theorem 1 continues to hold.

4 Unrestricted Information Disclosure

This section assumes that the consumer can disclose *any* information about the value vector $u = (u_1, \cdots, u_K)$. Formally, suppose that the seller sells $K \geq 2$ products, and $\mathcal{D}$ consists of all disclosure rules with finite message spaces.[27] In this "unrestricted" model, the consumer can disclose

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

[26] 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 $-c(\delta)$.

[27] Since I will assume that the value distribution has a finite support, this restriction is without loss of generality. Indeed, the consumer can maximize his payoff as long as he can choose any disclosure rule that has a message space

not only about which product is most valuable (as in the previous section) but also about the value of a particular product, such as whether u_k exceeds some deterministic threshold.

Two differences between the unrestricted model and the "restricted" model in Section 3 are worth mentioning. First, compared to the restricted model, the unrestricted model a priori "favors" discriminatory pricing in terms of revenue. This is because given general disclosure rules, discriminatory pricing often yields a higher revenue, as opposed to the restricted model where the two pricing regimes yield equal revenue for any fixed disclosure level δ.[28] Second, the unrestricted model has a theoretical connection to Bergemann et al. (2015). Their results imply that a single-product seller is indifferent between the two pricing regimes if information is disclosed to maximize consumer surplus, and that the equilibrium is efficient under discriminatory pricing. In contrast, I will show that the equilibrium is typically inefficient and the seller strictly prefers nondiscriminatory pricing.

I assume that x_0, which denotes the prior value distribution of each product, has a finite support $V = \{v_1, \dots, v_N\}$ with $0 < v_1 < \cdots < v_N$ and $N \geq 2$. For any $x \in \Delta(V)$, $x(v)$ 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 case, I impose the following assumption. Loosely speaking, it requires that the prior x_0 does not put too much weight on the lowest value of its support.

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

As the consumer can access a rich set of disclosure rules, the analysis is more involved than before; however, there turns out to be a clear relationship between pricing regimes and the kinds of information disclosed by the consumer. The next subsection illustrates this by showing that

M such that $|M| \leq |V| \times K$, because the consumer can always "pool" two message realizations that lead to the same recommended product and the same price.

[28]For example, if the consumer chooses a disclosure rule that reveals exact value vector u, then the seller can extract full surplus only under discriminatory pricing.

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 equilibrium can be inefficient in two ways: One is when the consumer decides not to buy any products; the other is when the consumer buys some product other than the most valuable ones. In principle, these kinds of inefficiencies can coexist. However, I show that each pricing regime is associated with only one type of inefficiency. The following result states that nondiscriminatory pricing leads to the first type of inefficiency. The proof is in Appendix E.

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.*[29]

The intuition is as follows. Once the seller commits to prices, information disclosure only affects product recommendations. Thus, the consumer prefers to fully disclose the highest value product so that the seller can recommend the best product for sure. Given efficient recommendations, the consumer is less likely to have low values for the recommended product than he is under the prior value distribution. Anticipating this, the seller prefers to commit to prices strictly greater than $\min V$, whenever she prefers to do so at the prior value distribution. Thus, with a positive probability, the consumer's value for the best product falls below the price and trade does not occur.

The next result shows that discriminatory pricing exhibits a different kind of inefficiency: Trade occurs whenever efficient, but it is associated with product mismatch. The proof needs some work, which is contained in Appendix F. As the existence of an equilibrium is non-trivial, I separately prove it in Appendix H.

Proposition 2. *Consider discriminatory pricing. In any equilibrium, trade occurs with probability*

[29] Note that I focus on perfect Bayesian equilibrium (PBE) in which the seller sets the same price for each product. Without this restriction, there could be a fully efficient PBE where the seller sets different prices for different products.

2. *However, for generic*[30] *priors* x_0 *satisfying Assumption 1, in any equilibrium, the consumer purchases some products other than the most valuable ones with a positive probability.*

An intuition for the first part is as follows. Suppose that the seller charges a price that exceeds the consumer's value with a positive probability. Suppose that, on such an event, the consumer discloses whether his value exceeds the price. If the seller learns that the value falls below the original price, she prefers to revise the price downward or recommend another product, which benefits both the consumer and the seller relative to no trade. Importantly, on the complementary event where the seller learns that the value exceeds the price, she prefers to recommend the same product at the same price as before.[31] Overall, the consumer is always willing to disclose whether his value exceeds the price, and this ensures that trade occurs for sure.

The second part implies that any equilibrium is generically inefficient, because the consumer fails to purchase the highest value product with a positive probability. The interpretation is that the seller's ability to base prices on information incentivizes the consumer to withhold information about which product is most valuable. The proposition implies that, under the current distributional assumption, the consumer *always* finds it profitable to conceal some information. Later, I provide a detailed discussion of why this is the case.

Now, can we use this result in policy? It would need more works to derive concrete policy implications, but the proposition 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 improve welfare. The analysis suggests that releasing information about consumers who have low values (i.e., values below the monopoly price) for all products is good for the welfare of sellers and consumers. In contrast, a regulator or an intermediary should be careful about releasing information bundles that contain consumers who have low values for some products and high values for other products. Whereas releasing such information may increase total welfare, it can hurt consumers to the extent that the loss from higher prices dominates the benefit from the improved

[30] The following is the formal description of genericity: Fix V and define $X_{>v_1} \subset \Delta(V)$ as the set of priors x_0 in $\Delta(V)$ 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.

[31] This hinges on the following observation in the (single product) monopoly pricing problem: If the seller optimally sets a price given some value distribution, then she does not revise the price even after learning that the value exceeds the price. Indeed, if the value u is drawn from $\mathbf{P}(\cdot)$, the optimal price p^* maximizes $p \cdot \mathbf{P}(u \geq p)$. If the seller additionally learns that $u \geq p^*$, then her new problem is $\max_{p \geq p^*} p \cdot \frac{\mathbf{P}(u \geq p)}{\mathbf{P}(u \geq p^*)}$, which is clearly maximized at $p = p^*$.

match quality.

I sketch the proof of Proposition 2. 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, any equilibrium is vertically efficient and generically horizontally inefficient. The proof of vertical efficiency follows the previous intuition: If an equilibrium is vertically inefficient, we can modify it so that the consumer discloses the additional information of whether his value for the recommended product exceeds the price. This leads to another equilibrium where the consumer and the seller obtain greater payoffs than the original equilibrium, which is a contradiction.[32]

The second part of the proposition—horizontal inefficiency—is more challenging to prove 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 equilibria, as we have to solve a Bayesian persuasion problem (Kamenica and Gentzkow, 2011) with multidimensional state and action spaces, which is known to be difficult.[33] Third, there may be multiple equilibria and we want to prove that all of them are horizontally inefficient.

To prove horizontal inefficiency without characterizing equilibrium, I take the following two-step approach. First, solve a "constrained Bayesian persuasion" 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^*, turns out to be simpler than the "unconstrained" maximization problem that the consumer faces in equilibrium. Second, modify ϕ^* to create disclosure rule ϕ^I that leads to an inefficient allocation but gives the consumer a strictly greater payoff than ϕ^*. These two steps imply that any equilibrium is associated with inef-

[32]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 result, which directly constructs a disclosure rule achieving an efficient outcome, I indirectly show vertical efficiency, as it is difficult to characterize an equilibrium.

[33]Precisely, the model is slightly different from a Bayesian persuasion because the sender (consumer) also takes an action. However, we can regard the consumer's payoffs from the optimal purchase behavior as the receiver's payoffs, which depend only on the state (values) and the sender's action (seller's ecommendation and pricing).

ficient allocation. As we proved that any 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. (The first column shows possible value vectors, and each row shows the distribution over messages 1 and 2 given each value vector.) ϕ only discloses

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

Table 1: Disclosure rule ϕ revealing product ranking

which product is more valuable. Since such information is necessary and sufficient to achieve horizontally efficient allocations, any efficient disclosure rules must be weakly more informative than ϕ.[34]

I find ϕ^* by maximizing the consumer's payoff among disclosure rules weakly more informative than ϕ. Specifically, for each k, I first calculate the posterior distribution of u_k conditional on message $k \sim \phi(\cdot|u)$. Then, I apply Bergemann et al.'s (2015) *consumer surplus maximizing segmentation (CSMS)* to each posterior distribution. In the single-product case, a CSMS discloses information about the consumer's value to maximize consumer surplus. In the current context, applying a CSMS to each posterior enables the consumer to disclose information about the value of the best product to maximize his payoff. Also, a property of CSMS implies that the resulting disclosure rule ensures that trade occurs with probability 1. Thus, ϕ^* constructed in this way is a solution of the constrained problem.

Table 2 presents disclosure rule ϕ^* obtained in this way. The CSMS decomposes each message k (of ϕ) into messages $k1$ and $k2$.[35] 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.

[34] 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.

[35] In general, a CSMS is not unique. In the proof and in this example, I use a CSMS obtained by the *greedy algorithm* of Bergemann et al. (2015).

	$\phi^*(11\|u_1,u_2)$	$\phi^*(12\|u_1,u_2)$	$\phi^*(21\|u_1,u_2)$	$\phi^*(22\|u_1,u_2)$
$(2,2)$	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^*.

	$\phi^I(11\|u_1,u_2)$	$\phi^I(12\|u_1,u_2)$	$\phi^I(21\|u_1,u_2)$	$\phi^I(22\|u_1,u_2)$
$(2,2)$	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^* 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 a small probability $\varepsilon > 0$. This creates horizontal inefficiency, because once the "new" message 21 is realized, the seller recommends product 2 even though $(u_1,u_2)=(2,1)$ with a positive probability. However, this modification does not affect the consumer's payoff, since at message 12, the consumer continues to obtain a payoff of zero. Importantly, this modification relaxes the seller's incentive, as she now *strictly* prefers to set price 1 at message 21. Second, I further modify ϕ^* 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^*, he now obtains a strictly positive payoff when message 21 is realized. To sum up, ϕ^I leads to horizontal inefficiency but gives the consumer a strictly greater payoff than ϕ^*.

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

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

The following result, however, shows that the seller still prefers to commit to not use the consumer's information for pricing, and the commitment 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 $U_{ND} < U_D$.*

In the proof, I consider the disclosure rule that maximizes the consumer's payoff subject to the efficiency constraint, which is characterized in Proposition 2. I compare this efficient disclosure rule with the equilibrium disclosure rule of each pricing regime in terms of the welfare of the seller and the consumer. In this way, I can do welfare comparison without knowing how discriminatory pricing equilibrium looks like.

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^* denote the disclosure rule constructed in the proof of Proposition 2: ϕ^* 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^*. (This is because I construct ϕ^* from ϕ^H

using the Bergemann et al.'s (2015) consumer surplus maximizing segmentation.) In other words, under ϕ^*, the seller can achieve the highest revenue by posting price p^* upfront for all products. Denoting the optimal revenue under ϕ^H and ϕ^* by R_{ND} and R^* respectively, I obtain $R_{ND} = R^*$.

As ϕ^* 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^*. This also implies that the seller under discriminatory pricing is strictly worse off ($R_D < R^* = R_{ND}$). Finally, as the consumer's payoff is greater under ϕ^* than ϕ^H, $U_D > U_{ND}$ holds. □

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 because it makes trade more likely to occur. (Note that this effect was absent in Section 3.) The reason why this argument fails is that, if the consumer can choose any disclosure rule, then he can disclose partial information about values to increase the probability of trade without increasing the seller's payoff. In other words, the seller does not lose from not being able to base prices on information. As a result, the seller prefers nondiscriminatory pricing, which leads to more accurate recommendations.

4.3 Nondiscriminatory Pricing Can Enhance Efficiency

Which pricing regime achieves greater total surplus? The answer is not obvious because neither of the pricing regimes achieves full efficiency under Assumption 1. Indeed, the answer depends on the prior value distribution x_0 of each product and the number K of products. For example, 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 where it leads to a *strictly* greater total surplus, I assume that x_0 does not put too much weight on the highest value of its support V. In the following analysis, I do not impose Assumption 1.

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

Note that Assumption 2 holds whenever x_0 has a density. The proof of the next proposition is in Appendix G.

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

The intuition is as follows. As K grows large, the consumer's value for the best product (i.e., $\max(u_1, \ldots, u_K)$) becomes nearly degenerate at $\max V$. This implies that, if the seller can always recommend the best product, then she can set prices near $\max V$ to achieve the revenue close to $\max V$. As a result, consumer surplus converges to zero and total surplus converges to $\max V$. This is indeed an equilibrium of nondiscriminatory pricing as it achieves the efficient recommendation (Proposition 1). In contrast, under discriminatory pricing (with Assumption 2), the consumer can secure a positive payoff that is independent of the number of products by disclosing no information. In other words, the consumer's equilibrium payoff never vanishes even if K grows large. This implies that the total surplus cannot approach the efficient level $\max V$.

5 Extensions

5.1 General Conditions for Welfare Ranking Results

This extension provides more general conditions on the set $\mathcal{D}$ of available disclosure rules under which the seller is better off and the consumer is worse off under nondiscriminatory pricing. Define ϕ^E as a disclosure rule that only discloses the highest value product: For any realized $u \in V^K$, ϕ^E draws message $k \in \arg\max_\ell u_\ell$ with probability $\frac{1}{|\arg\max_\ell u_\ell|}$. (Alternatively, the following results hold if I define ϕ^E as a disclosure rule that discloses the ranking of the products in terms of their values.) The following result provides a condition under which the seller prefers to commit to not use information for pricing. The condition is more general than the one in Section 3.

Proposition 4. *If ϕ^E belongs to $\mathcal{D}$ and is more informative than any other disclosure rules in $\mathcal{D}$, then the seller obtains a higher payoff under nondiscriminatory pricing than under discriminatory pricing.*

Proof. As in Theorem 1, under nondiscriminatory pricing, the consumer chooses ϕ^E and the seller sets a price of $\arg\max_p p[1 - F^{MAX}(p)]$ for each product. Now, consider discriminatory pricing.

Let $(M, \phi) \in \mathcal{D}$ denote an equilibrium disclosure rule, and take any message realization m drawn by ϕ. Conditional on m, the consumer's value distribution for the recommended product (say k^*) is lower than F^{MAX} in the sense of first order stochastic dominance. Indeed, since ϕ is less informative than ϕ^E, the value distribution of product k^* conditional on m is a convex combination of the distributions of $u_{k^*} | \{u_{k^*} = \max_k u_k\}$ and $u_{k^*} | \{u_{k^*} \neq \max_k u_k\}$. Thus, the seller obtains a weakly higher payoff under nondiscriminatory pricing. □

The next result provides a condition under which the consumer is worse off under the seller's commitment. The condition is more general than the ones in Sections 3 and 4. (However, it only states that the consumer *weakly* prefers discriminatory pricing, in contrast to Theorem 2.)

Proposition 5. *If $\phi^E \in \mathcal{D}$, then the consumer obtains a higher payoff under discriminatory pricing than under nondiscriminatory pricing.*

Proof. As in Theorem 1, under nondiscriminatory pricing, the consumer chooses ϕ^E. Under discriminatory pricing, the consumer can achieve the same outcome by choosing ϕ^E. Thus, the consumer is weakly better off under discriminatory pricing. □

5.2 Market for Personal Data

Institutions within which consumers can sell their information have been discussed as market-based solutions to address privacy problems. In my model, such a "market for data" could benefit both the seller and the consumer.

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

How does this "market for data" affect equilibrium outcomes? It has no impact under nondiscriminatory pricing because the consumer is willing to disclose full information without compensation. In contrast, under discriminatory pricing, the market for data benefits the seller without affecting the consumer. For example, if $\mathcal{D}$ contains disclosure rule ϕ^* that reveals values u, then 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. Recall that without compensation, the consumer typically hides some information, which leads to product mismatch (Proposition 2).

Importantly, in this new setting, not only the consumer but also the seller may prefer discriminatory pricing, because the market for data increases the seller's payoff under discriminatory pricing but has no impact under nondiscriminatory pricing. Thus, the market for data can align the preferences of the seller and the consumer over whether the seller uses information for pricing.

5.3 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. As discussed, this is 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 restricted model 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. Note that θ or π alone is not informative of product valuations, but (θ, π) is. This formulation captures situations in which sellers have to combine information abut the tastes of consumers and product characteristics in order to learn about preferences and give product recommendations.

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. The seller observes δ and a realized message, and then recommends a product. As before, I consider two pricing regimes. Note that, once the consumer observes θ, the game looks identical with the original restricted model. Thus, this setting leads to the same result as Theorem 1:

5.4 Alternative Interpretation: Online Advertising Platform

I 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 ad. Which ad 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 consumer sees the ad of product k. After seeing an ad, the consumer learns the value and price of the advertised product, and then 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.

This result suggests that platforms could benefit from committing not to disclose the value of δ to advertisers, because it implements nondiscriminatory pricing. For example, platforms might classify consumers into two segments, each of which consists of consumers who are more likely to prefer one product than the other. By committing to engage in coarse segmentation, platforms can encourage consumers to provide information.

6 Concluding Discussion

This paper studies consumers' privacy choices, their price implications, and welfare consequences. The key to the analysis is the following trade-off: A consumer may benefit from revealing information, because a seller can then offer product recommendations. However, the seller may also use the information to tailor prices.

The key finding, however, is to show that this trade-off may *not* arise in equilibrium. I identify a condition under which the seller prefers to commit to not use the consumer's information for pricing. The commitment encourages the consumer to disclose more information that is useful for accurate recommendations. Surprisingly, this commitment can hurt the consumer: The consumer could be better off if the seller had less information and could only make noisy recommendation. Moreover, in contrast to the standard single-product model, the seller's commitment can enhance total welfare at the expense of consumer welfare. These results give a potential explanation of a casual observation that online retailers seem to not use individual data to tailor prices, and consumers seem to share much information. Finally, the model of unrestricted disclosure reveals that even with fine-grained control of information, we cannot simultaneously achieve efficient price discrimination and efficient matching of products without sacrificing consumer welfare.

There are various directions for future research. For example, one might want to incorporate information sharing between sellers or the presence of data brokers, which is likely to add new policy implications. Also, it would be important to study how the market for personal data works beyond the context of product recommendations and pricing. Finally, this paper shows that there could be a divergence between individual and collective incentives to share information, even if types are independent across consumers. It would be fruitful to more generally study when this type of divergence occurs and what kind of policies can mitigate it.

References

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

Aguirre, Inaki, Simon Cowan, and John Vickers (2010), "Monopoly price discrimination and demand curvature." *American Economic Review*, 100, 1601–15.

Aliprantis, Charalambos D and Kim Border (2006), *Infinite dimensional analysis: a hitchhiker's guide*. Springer Science & Business Media.

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.

Chen, Yongmin and Marius Schwartz (2015), "Differential pricing when costs differ: a welfare analysis." *The RAND Journal of Economics*, 46, 442–460.

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.

Cowan, Simon (2016), "Welfare-increasing third-degree price discrimination." *The RAND Journal of Economics*, 47, 326–340.

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."

Hidir, Sinem and Nikhil Vellodi (2018), "Personalization, discrimination and information revelation."

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 (2013), "Online price discrimination: conspicuous by its absence."

Pigou, Arthur (1920), *The Economics of Welfare*. London: Macmillan.

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.

Stokey, Nancy L (1979), "Intertemporal price discrimination." *The Quarterly Journal of Economics*, 355–371.

Stokey, Nancy L (1981), "Rational expectations and durable goods pricing." *The Bell Journal of Economics*, 112–128.

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.

Zeng, Xiaohua, Srabana Dasgupta, and Charles B Weinberg (2013), "The competitive implications of a no-haggle pricing policy: The access toyota case." Technical report, Working paper, College of Business, City University of Hong Kong, Hong Kong .

Zeng, Xiaohua, Srabana Dasgupta, and Charles B Weinberg (2016), "The competitive implications of a no-haggle pricing strategy when others negotiate: Findings from a natural experiment." *International Journal of Research in Marketing*, 33, 907–923.

Appendix For Online Publication

A Proof of Lemma 1

Let F^{MAX} and F^{MIN} denote the CDFs of $\max(u_1, u_2)$ and $\min(u_1, u_2)$, respectively. Without loss of generality, suppose that message 1 is realized. If the seller recommends products 1 and 2, then the consumer draws values from $\delta F^{MAX} + (1-\delta)F^{MIN}$ and $\delta F^{MIN} + (1-\delta)F^{MAX}$, respectively. The former first order stochastically dominates the latter because $\delta > 1/2$ and F^{MAX} first order stochastically dominates F^{MIN}. Then, under both pricing regimes, recommending product 1 maximizes revenue given any prices. (Under nondiscriminatory pricing, I use the equilibrium restriction that all products have the same price.) Note that the seller is indifferent between recommending two products if the optimal price is equal to the lowest possible value ($\min V$). In this case, the equilibrium restriction requires that the seller recommends product 1 because it uniquely maximizes the consumer's payoff given that the value distribution is non-degenerate.

B Proof of Lemma 3

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 the optimal prices is $P(\delta) := \arg\max_p p[1 - \delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)]$. By the tie-breaking rule, the seller sets a price of $p(\delta) := \min P(\delta)$, which is well-defined as shown in Appendix H. I show that $p(\delta)$ is increasing in δ. Note that

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

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. Therefore, $p(\delta)$ is increasing in δ.

C Proof of Theorem 1

If the consumer is recommended his preferred and less preferred products at price p, the 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.[36] If the consumer chooses δ, then his expected payoff is $\delta u^{MAX}(p^*) + (1-\delta)u^{MIN}(p^*)$. This is maximized at $\delta = 1$, which is also a unique disclosure level consistent with my equilibrium

[36] Since the seller breaks tie in favor of the consumer, the seller sets the lowest optimal price with probability 1. This excludes the use of a strictly mixed strategy in pricing.

restriction. Anticipating $\delta = 1$, the seller sets the equilibrium price $p^* = p(1)$, and thus the consumer's payoff is $u^{MAX}(p(1))$.

Under discriminatory pricing, the consumer's payoff from δ is $\delta u^{MAX}(p(\delta)) + (1-\delta)u^{MIN}(p(\delta))$. Thus, the equilibrium payoff is

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

where the existence of a maximizer is shown in Appendix H. That is, the consumer is worse off under nondiscriminatory pricing, and he is charged higher prices by Lemma 3.

Next, consider the seller's payoff. Since F^{MAX} first order stochastically dominates F^{MIN}, it holds $F^{MAX}(p) \leq F^{MIN}(p)$, and thus $\delta F^{MAX}(p) + (1-\delta)F^{MIN}(p)$ is decreasing in δ for any p. Thus, $p[1 - \delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)]$ is increasing in δ for any p. Then, $\max_p p[1 - \delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)]$ is maximized at $\delta = 1$. Therefore, the seller is better off under nondiscriminatory pricing.

D 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 $p(\bar{\delta})$, where $p(\cdot)$ is defined in Lemma 3. Now, if a positive mass of consumers choose strictly greater disclosure levels, then $\bar{\delta}$ increases. This increases $p(\bar{\delta})$ and decreases the payoffs of other consumers who have not changed disclosure levels.

In contrast, under discriminatory pricing, consumer i is charged a price of $p(\delta_i)$ for recommended products. Thus, each consumer's problem is identical with the one in the original formulation. Thus, consumers disclose less information and are better off under discriminatory pricing.

E Proof of Proposition 1

Take any equilibrium under nondiscriminatory pricing. Since prices are fixed and the same across products, it is optimal for the consumer to disclose information so that the seller recommends the most valuable products with probability 1.[37] Now, the seller sets a price of product k' to maximize the expected revenue conditional on recommending product k'. Conditional on this event, the seller's posterior belief for $u_{k'}$ is equal to the distribution of $\max_{k \in \mathcal{K}} u_k$, denoted by x^{MAX}. Note that the argument does not depend on k' because values are IID across products. Then, it holds that

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

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 $x(v_1)^K > 0$.

F Proof of Proposition 2

Proposition 2 follows from a series of lemmas. Lemma 4 proves vertical efficiency. Lemma 5 proves that any equilibrium is horizontally inefficient whenever prior x^0 is such that the distribution of the highest value product has a unique monopoly price. Lemma 6 proves that this condition is true for generic priors satisfying Assumption 1.

Lemma 4. *Any equilibrium is vertically efficient.*

Proof. Take any disclosure rule (M^*, ϕ^*) that leads to a vertically inefficient allocation given the seller's best response and the consumer's optimal purchase decision with the tie-breaking. (Hereafter, I omit the caveat "given the tie-breaking rule.") Then, ϕ^* draws a posterior $x \in \Delta(V^K)$ at which trade fails to occur with positive probability.[38] Without loss of generality, suppose that given

[37]For the sake of completeness, I present an example. Consider disclosure rule (ϕ^*, M^*) such that $M^* = \mathcal{K}$ and $\phi^*(k|u) = \frac{1}{|\arg\max_{k \in \mathcal{K}} u_k|} \mathbf{1}_{\{k \in \arg\max_{k \in \mathcal{K}} u_k\}}$. Namely, ϕ^* is a symmetric disclosure rule that reveals the name of the most valuable product. An equilibrium disclosure rule is not unique; we can consider any disclosure rules weakly more informative than (ϕ^*, M^*).

[38]Because $|V^K| < +\infty$, without loss of generality, I can assume $|M^*| < +\infty$. Then, each message is realized

x, the seller recommends product 1 at price v_n. Consider the following disclosure rule ϕ^{**}: On top of the information that ϕ^* discloses, ϕ^{**} also discloses $u_1 \geq v_n$ or $u_1 < v_n$ whenever posterior x is realized.

I show that ϕ^{**} yields a weakly greater consumer surplus and a strictly greater total surplus than ϕ^* does. Let x^+ and $x^- \in \Delta(V^K)$ denote the posterior beliefs of the seller when the consumer discloses $u_1 \geq v_n$ and $u_1 < v_n$ (following x), respectively. Note that for some $\alpha \in (0,1)$, $x = \alpha x^+ + (1-\alpha)x^-$. First, consider the consumer's payoff and total surplus conditional on x^-. The consumer obtains a greater payoff under ϕ^{**} than under ϕ^* because consumer surplus is zero under ϕ^*. Total surplus is strictly greater under ϕ^{**} because trade occurs with a positive probability under ϕ^{**} but occurs with probaiblity zero under ϕ^*. Second, I show that, conditional on x^+, the consumer's payoff and total surplus remain the same before and after the seller learns $u_1 \geq v_n$. To begin with, I show that the seller continues to recommend product 1 at price v_n at x^+. Suppose to the contrary that the seller strictly prefers to recommend product m at price v_ℓ where $(m,\ell) \neq (1,n)$. Let $x_1^+ \in \Delta(V)$ and $x_m^+ \in \Delta(V)$ be the marginal distributions of u_1 and u_m given x^+, respectively. Because the seller strictly prefers recommending product m at price v_ℓ to recommending product 1 at price v_n, we get

$$v_\ell \sum_{j=\ell}^{N} x_m^+(v_j) > v_n \sum_{j=n}^{N} x_1^+(v_j),$$

which implies

$$v_\ell \sum_{j=\ell}^{N} \left[\alpha x_m^+(v_j) + (1-\alpha)x_m^-(v_j)\right] \geq v_\ell \sum_{j=\ell}^{N} \alpha x_m^+(v_j) > v_n \sum_{j=n}^{N} \alpha x_1^+(v_j) = v_n \sum_{j=n}^{N} \left[\alpha x_1^+(v_j) + (1-\alpha)x_1^-(v_j)\right]. \tag{2}$$

The last equality follows from $x_1^-(v) = 0$ for any $v \geq v_n$. Inequality (2) contradicts that the seller prefers to recommend product 1 at price v_n at x. Therefore, the seller continues to recommend the same product at the same price between x and x^+. Thus, overall, disclosing $u_1 \geq v_n$ or $u_1 < v_n$ at x leads to a weakly greater consumer surplus and a strictly greater total surplus.

To show that any equilibrium is vertically efficient, take any equilibrium disclosure rule ϕ^*.

with a positive probability from the ex-ante perspective. This implies that there is an ex-ante positive probability event such that some posterior $x \in \Delta(V^K)$ is realized and trade fails to occur.

Appendix H proves the existence of an equilibrium. Suppose to the contrary that ϕ^* is vertically inefficient. Then, I can apply the modification described above to create ϕ^{**}. ϕ^{**} gives the consumer a weakly greater payoff than ϕ^*. Because ϕ^* is optimal for the consumer, he is indifferent between ϕ^* and ϕ^{**}. However, ϕ^{**} yields a strictly greater total surplus, which implies that the seller strictly prefers ϕ^{**}. This contradicts the tie breaking rule, which requires that ϕ^* maximizes the seller's payoffs among all consumer-optimal disclosure rules. Therefore, ϕ^* is vertically efficient. □

Lemma 5. *Suppose that prior x^0 satisfies Assumption 1 and there is a unique monopoly price given value distribution F^{MAX}. Here, F^{MAX} is a CDF of $\max(u_1, \dots, u_K)$ where each u_k is an IID draw from x^0. Then, any equilibrium is horizontally inefficient.*

Proof. I construct a disclosure rule that maximizes the consumer's ex ante expected payoff among $\mathcal{E} \subset \mathcal{D}$, where $\mathcal{E}$ is the set of all disclosure rules that lead to efficient recommendations given the optimal behavior of each player. Take any disclosure rule $\phi \in \mathcal{E}$. Since the recommended product always belongs to $\arg\max_{\ell \in \mathcal{K}} u_\ell$ with (ex ante) probability 1, the consumer's value of the recommended product (unconditional on which product is recommended) is drawn according to F^{MAX}. This implies that under ϕ, the seller can obtain a revenue of at least $\underline{R} := \max_{p \in V} p[1 - F^{MAX}(p)]$ by setting a price of $\arg\max_{p \in V} p[1 - F^{MAX}(p)]$ for all realized posteriors. Thus, ϕ^* maximizes the consumer's payoff among $\mathcal{E}$ whenever it achieves an efficient allocation and gives the seller a payoff of $\underline{R}$.

To begin with, consider disclosure rule $\phi^E \in \mathcal{E}$ such that for any realized $u \in V^K$, ϕ^E draws message $k \in \arg\max_\ell u_\ell$ with probability $\frac{1}{|\arg\max_\ell u_\ell|}$. Two remarks are in order. First, u_k is distributed according to F^{MAX} conditional on message k. Second, the seller prefers to recommend product k after observing message k no matter what additional information she learns, because she can maximize the probability of trade by recommending product k for any given price.

Next, I create $\phi^* \in \mathcal{E}$ by modifying ϕ^E as follows: For each $k \in \mathcal{K}$, conditional on that message k is realized under ϕ^E, ϕ^* discloses additional information about u_k according to a *consumer surplus maximizing segmentation (CSMS)* characterized by Bergemann et al. (2015).[39] In our

[39] In a single product monopoly pricing, a consumer surplus maximizing segmentation is equivalent to a disclosure rule that has the following property. First, at each realized posterior, the seller is willing to set the price equal to the minimum of its support, which implies that the trade occurs for sure. Second, at each posterior, the seller is indifferent between charging the minimum of each posterior and charging the monopoly price for the prior.

context, the information disclosed according to (any) CSMS ensures that the trade occurs with probability 1 whereas the seller's resulting revenue is $\underline{R}$. Thus, under ϕ^*, the seller recommends the highest value product and the trade occurs with probability 1, whereas the seller's revenue is $\underline{R}$. Thus, ϕ^* maximizes the consumer's payoff among $\mathcal{E}$.

Hereafter, I focus on a particular ϕ^* where the additional information about the highest value product is disclosed according to a CSMS constructed by the *greedy algorithm* in Bergemann et al. (2015). This has the following implication. Let $\{x^k_{S_1}, \ldots, x^k_{S_L}\}$ denote the set of posteriors induced by ϕ^* conditional on message k. Without loss of generality, we can regard $x^k_{S_1}, \ldots, x^k_{S_L}$ as realized messages. Let us also regard each $x^k_{S_\ell}$ as a marginal distribution of u_k instead of a joint distribution of $(u_1, \ldots, u_K)$. The greedy algorithm guarantees that each $x^k_{S_\ell}$ has support $S_\ell \subset V$, $S_1 \subset S_2 \subset \cdots \subset S_L = V$, and the set of all optimal prices against $x^k_{S_\ell}$ is S_ℓ. Moreover, it holds that $S_1 = \{v^*\}$ with $v^* > v_1$. To see this, note that $|S_1| \geq 2$ implies that two prices in S_1 are optimal against all posteriors in $\{x^k_{S_1}, \ldots, x^k_{S_L}\}$, which in turn implies that these prices are optimal under F^{MAX} due to the linearity of the expected revenue in the value distribution. This contradicts that there is a unique optimal price under F^{MAX}. Thus, $|S_1| = 1$, which implies that $S_1 = \{v^*\}$. $v^* > v_1$ follows from the proof of Proposition 1.

I modify ϕ^* to create a horizontally inefficient ϕ^I that yields a strictly greater consumer surplus than ϕ^* does. For the ease of exposition, I use the following terminologies. First, I regard a distribution $x \in \Delta(V^K)$ as consisting of a unit mass of consumers, where mass $x(u)$ of consumers have valuation vector u. Second, I call any set of (a continuum of) consumers as a "segment."

To construct ϕ^I, I make three observations. First, a positive mass of consumers in $x^1_{S_1}$ have value v^* for product 1 and the lowest possible value $v_1 < v^*$ for product 2. Call this mass of consumers "segment (v^*, v_1)." Second, a positive mass of consumers in $x^1_{S_1}$ have value v^* for both products 1 and 2. Call these consumers "segment (v^*, v^*)."

First, I take a small but positive (say ε_1) mass of segment (v^*, v_1) from $x^1_{S_1}$ and pool this segment with $x^2_{S_L}$.[40] Let $\hat{x}^2_{S_L}$ denote the posterior following this pooling. For a sufficiently small $\varepsilon_1 > 0$, at $\hat{x}^2_{S_L}$, the seller recommend product 2, and it *strictly* prefers to set price v_1 for product 2. The reason is as follows. Under the original posterior $x^2_{S_L}$, it is optimal for the seller to recommend

[40] In terms of a disclosure rule, this means that I modify ϕ^* so that it draws message $x^2_{S_L}$ with probability $\varepsilon_1 > 0$ not only following message 2 but also when ϕ^* draws segment (v^*, v_1) in $x^1_{S_1}$.

product 2 at any price in V because $S_L = V$. After the modification, $\hat{x}^2_{S_L}$ contains a strictly greater mass of consumers who have value v_1 for product 2 (i.e., segment (v^*, v_1)). Thus, the seller strictly prefers to set price v_1 for product 2. Moreover, for a small $\varepsilon_1 > 0$, the seller does not strictly prefer to recommend other products. Indeed, if the seller recommended product $k \neq 2$ at $x^2_{S_L}$, then she would strictly prefer to set price v_1. Thus, for a small $\varepsilon_1 > 0$, the seller's pricing incentive does not change. This implies that the optimal revenue from recommending other products (at the new posterior $\hat{x}^2_{S_L}$) is v_1, which is no greater than the revenue from recommending product 2. Importantly, this modification does not change the consumer's payoff, because consumers in segment (v^*, v_1) obtain zero payoffs under $x^1_{S_1}$. Let ϕ^H denote the resulting disclosure rule.

Finally, I modify ϕ^H by pooling a small but positive (say ε_2) mass of segment (v^*, v^*) in $x^1_{S_1}$ with $\hat{x}^2_{S_L}$. Let $\tilde{x}^2_{S_L}$ denote the posterior following this pooling. If ε_2 is small, the seller continues to recommend product 2 at price v_1 under $\tilde{x}^2_{S_L}$, because it strictly prefers to set price v_1 for product 2 at $\hat{x}^2_{S_L}$. This modification strictly increases the consumer's payoff relative to ϕ^*, because consumers in segment (v^*, v^*) obtain a positive payoff $v^* - v_1$ while they obtain to a payoff of zero under ϕ^*.

Note that the resulting disclosure rule ϕ^I strictly increases the consumer's ex ante expected payoff, but it leads to inefficient recommendation at $\tilde{x}^2_{S_L}$. This implies that any equilibrium is horizontally inefficient, because for any disclosure rule leading to horizontal efficiency, the consumer can find more profitable disclosure rules that lead to horizontal inefficiency. □

Lemma 6. *Fix V and define $X_{>v_1} \subset \Delta(V)$ as the set of priors $x_0 \in \Delta(V)$ such that $p(x_0) > v_1$. There is a Lebesgue measure-zero set $X_0 \subset \Delta(V)$ such that, for any $x_0 \in X_{>v_1} \setminus X_0$, the induced F^{MAX} has a unique monopoly price (i.e., F^{MAX} satisfies the condition of Lemma 5.)*

Proof. First, let D_2 denote the set of all distributions $x \in \Delta(V)$ such that there are two or more optimal prices. I show that D_2 has measure zero. Let $D_{v,v'}$ denote the set of all distributions in $\Delta(V)$ such that both prices v and v' are optimal. I can write $D_{v,v'} = \{x \in \Delta(V) : v \sum_{v_n \geq v, v_n \in V} x(v_n) = v' \sum_{v_n \geq v', v_n \in V} x(v_n)\}$. $D_{v,v'}$ is a subset of $N-1$-dimensional hyperplane, which has measure zero in $\mathbb{R}^N$. Thus, $D_2 = \cup_{(v,v') \in V^2} D_{v,v'}$ has measure zero.

Consider a function φ which maps any distribution $x = (x_1, \ldots, x_N) \in \Delta(V)$ to the distribution of $\max(u_1, \ldots, u_K)$, where each u_k is an IID draw from x. 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 any measure-zero set by φ has measure zero. In particular, $X_0 := \varphi^{-1}(D_2)$ has measure zero. Clearly, X_0 has the desired property. □

G Proof of Proposition 3

Proposition 3 relies on the following lemma.

Lemma 7. *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, then there is $\underline{u} > 0$ such that the consumer's equilibrium payoff is at least $\underline{u}$ for any K.*

Proof. By the same argument as Proposition 1, the seller under nondiscriminatory pricing recommends the most valuable product with probability 1. Let F denote the CDF of the value for each product (induced by x_0). Take any $\varepsilon > 0$. Suppose that the seller sets $p = \max V - \varepsilon/2$ for each product upfront. As $K \to +\infty$, 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 $\max V - \varepsilon$ 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, observe that the consumer can choose to disclose no information and obtain a payoff of $\int_{p(x_0)}^{\max V} v - p(x_0) dF(v) > 0$, which is positive and independent of K. □

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

I show that total surplus under discriminatory pricing is uniformly bounded away from $\max V$. 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 $\varepsilon p + (1 - \varepsilon) \max V < \max V$. 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 seller'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, as in Lemma 7. □

H Existence of Equilibrium under Discriminator Pricing

I prove the existence of an equilibrium under discriminatory pricing. Recall that for nondiscriminatory pricing, I have proved the existence by constructing an equilibrium.

Restricted Model

Claim 1. In the restricted model, there exists an equilibrium under discriminatory pricing.

The result follows from two lemmata.

Lemma 8. *Given a disclosure level δ, the lowest optimal price $p(\delta)$ exists and is lower semi-continuous*[41] *in δ.*

Proof. Define $G(p) := \delta F^{MAX}(p) + (1-\delta)F^{MIN}(p)$. Recall that I define a CDF as a left-continuous function. This implies that G is lower semi-continuous. Indeed, for any p^* and $\varepsilon > 0$, take $\delta > 0$ so that for all $p \in (p^* - \delta, p^*]$, $|G(p) - G(p^*)| < \varepsilon$. Because $G(p)$ is increasing in p, for all $p \in (p^* - \delta, p^* + \delta)$, $G(p) > G(p^*) - \varepsilon$.

Because G is lower semi-continuous, $1 - G$ is upper semicontinuous, and thus $p[1 - G(p)]$, which is a product of two nonnegative upper semicontinuous functions, is upper semi-continuous in p.[42] This implies that $P(\delta) := \arg\max_{p \in V} p[1 - G(p)]$ is nonempty and compact (Theorem 2.43 of Aliprantis and Border (2006)). Thus, $p(\delta) := \min P(\delta)$ exists.

Next, I show that $p(\delta)$ is lower semi-continuous in δ. If $p(\delta)$ is not lower semi-continuous at some δ^*, there is $\varepsilon > 0$ such that we can construct a sequence $\delta_n \to \delta^*$ so that $p(\delta_n) < p(\delta^*) - \varepsilon$ for all n. Then, we can find a convergent subsequence of $(p(\delta_n))_n$ because $p(\delta_n) \in V$ and V is compact. Without loss of generality, assume that $(p(\delta_n))_n$ itself converges, so that there exists $p^* = \lim_n p(\delta_n) < p(\delta^*)$. Now, define $Y(p, \delta)$ as

$$Y(p,\delta) := p\left[1 - \delta F^{MAX}(p) - (1-\delta)F^{MIN}(p)\right] - p(\delta^*)\left[1 - \delta F^{MAX}(p(\delta^*)) - (1-\delta)F^{MIN}(p(\delta^*))\right].$$

[41] Depending on the context, I use one of the following two (equivalent) conditions as a definition of lower semicontinuity. Given a (first countable) topological space X and $f : X \to \mathbb{R}$, f is lower semicontinuous if $\{x \in X : c \geq f(x)\}$ is closed for each $c \in \mathbb{R}$. Equivalently, f is lower semicontinuous if $x_n \to x$ implies $\liminf_n f(x_n) \geq f(x)$. We obtain the definition(s) for upper semicontinuity by replacing $\geq$ and $\liminf$ with $\leq$ and $\limsup$, respectively. For the equivalence of the two conditions, see Lemma 2.42 of Aliprantis and Border (2006).

[42] To see this, if $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ are nonnegative and upper semicontinuous, for any $x_n \to x$, we obtain $\limsup_n f(x_n)g(x_n) \leq \limsup_n f(x_n) \limsup_n g(x_n) \leq f(x)g(x)$. Thus, fg is upper semicontinuous.

Because $p(\delta_n)$ is optimal given δ_n, it holds $Y(p(\delta_n), \delta_n) \geq 0$. Also, $Y(p, \delta)$ is upper semicontinuous in (p, δ).[43] This implies that $Y^* := \{(p, \delta) : Y(p, \delta) \geq 0\}$ is closed. Thus, $(p^*, \delta^*) = \lim_n (p(\delta_n), \delta_n) \in Y^*$, or equivalently,

$$\begin{aligned}&p^* \left[1 - \delta^* F^{MAX}(p^*) - (1-\delta^*) F^{MIN}(p^*)\right] \\ \geq &p(\delta^*) \left[1 - \delta^* F^{MAX}(p(\delta^*)) - (1-\delta^*) F^{MIN}(p(\delta^*))\right],\end{aligned}$$

which implies $p^* \in P(\delta^*)$. This contradicts $p(\delta^*) = \min P(\delta^*)$ because $p^* < p(\delta^*)$. □

Lemma 9. $\delta u^{MAX}(p(\delta)) + (1-\delta) u^{MIN}(p(\delta))$ *is upper semi-continuous in* δ.

Proof. Note that $u^{MAX}(p) = \int_p^{+\infty} (x-p) dF^{MAX}(x) = \int_p^{+\infty} 1 - F^{MAX}(x) dx$ is continuous and decreasing in p. Because $p(\delta)$ is lower semi-continuous, $u^{MAX}(p(\delta))$ is upper semi-continuous in δ. (To see this, if g is continuous and decreasing and f is lower semi-continuous, then for $x_n \to x$, we have $\lim_{k\to+\infty} \sup_{n>k} g(f(x_n)) = \lim_{k\to+\infty} g(\inf_{n>k} f(x_n)) = g(\lim_{k\to+\infty} \inf_{n>k} f(x_n)) \leq g(f(x))$. The last inequality is from the lower semi-continuity of f.) Similarly, we can show that $u^{MIN}(p(\delta))$ is upper semi-continuous. Therefore, $\delta u^{MAX}(p(\delta)) + (1-\delta) u^{MIN}(p(\delta))$ is upper semi-continuous in δ. □

Proof of Claim 1. $\delta u^{MAX}(p(\delta)) + (1-\delta) u^{MIN}(p(\delta))$ is upper semi-continuous, and the set of disclosure levels, $[0, 1]$, is compact. Thus, $D^* := \arg\max_{\delta \in [0,1]} \delta u^{MAX}(p(\delta)) + (1-\delta) u^{MIN}(p(\delta))$ is nonempty and compact (Theorem 2.43 of Aliprantis and Border (2006)). Thus, $\delta^* := \max D^*$ combined with the optimal on-path and off-path actions of the seller and the consumer consists of an equilibrium. □

Unrestricted Model

Claim 2. In the unrestricted model, there exists an equilibrium under discriminatory pricing.

[43] This follows from the fact that a product of two nonnegative upper (lower) semicontinuous functions and a sum of two upper (lower) semicontinuous functions are upper (lower) semicontinuous. Note that $F^{MAX}(p)$ and $F^{MIN}(p)$ are lower semicontinuous in p, which implies that $\delta F^{MAX}(p) + (1-\delta) F^{MIN}(p)$ is lower semicontinuous. Thus, $p\left[1 - \delta F^{MAX}(p) - (1-\delta) F^{MIN}(p)\right]$ is upper semicontinuous. Because the second term of $Y(p, \delta)$, which is $p(\delta^*)\left[1 - \delta F^{MAX}(p(\delta^*)) - (1-\delta) F^{MIN}(p(\delta^*))\right]$, is continuous in (p, δ), overall, $Y(p, \delta)$ is upper semicontinuous.

Proof. I prepare some notations. Let $A := \mathcal{K} \times V$ denote the seller's (finite) action space, i.e., the set of pairs of recommended products and prices. When the seller recommends product $k \in \mathcal{K}$ at price $p \in V$, I say that the seller chooses $a = (k, p) \in A$. Given $(a, b) \in A \times \Delta(V^K)$, let $U(a, b)$ and $R(a, b)$ denote the expected payoffs of the consumer and the seller, respectively, when the seller chooses a, the consumer's valuations are drawn according to b, and the consumer takes an optimal purchase decision (breaking ties in favor of the seller). Given the seller's belief $b \in \Delta(V^K)$, let $a(b) \in A$ denote the seller's optimal recommendation and pricing that break ties in favor of the consumer. ($a(b)$ exists because A is finite.)

I prove the existence of an equilibrium (disclosure rule) following the standard procedure. The first step shows that the consumer's payoff is upper semi-continuous in disclosure rules. The second step proves that the set of all disclosure rules is compact. I use weak* topology in $\Delta(\Delta(V^K))$.

Consider an information set where the seller chooses a product to recommend and a price. Let $b \in \Delta(V^K)$ denote the seller's belief about the valuation vector. If the seller and the consumer take optimal actions following the information set, the consumer's expected payoff is given by $U(a(b), b)$. I show that $U(a(b), b)$ is upper semi-continuous in $b \in \Delta(V^K)$. Suppose to the contrary that there exists $\varepsilon > 0$ and $(b_n)_{n=1}^{+\infty} \subset \Delta(V^K)$ such that $\lim_n b_n = b$ but $U(a(b_n), b_n) \geq U(a(b), b) + \varepsilon$ for all n. Because A is finite, we can choose a subsequence $(b_{n(m)})_{m=1}^{+\infty}$ so that for some $a' \in A$, $a(b_{n(m)}) = a'$ for all m. Without loss of generality, assume that $a(b_n) = a'$ for all n. Note that $R(a', b_n) \geq R(a(b), b_n)$ because $a' = a(b_n)$ is optimal for the seller given its belief b_n. Note also that $R(a, b)$ is continuous in b with a fixed a. Indeed, suppose that a is such that the seller recommends product k at price p, where the consumer's value for product k is distributed according to $b^k = (b_1^k, \ldots, b_N^k) \in \Delta(V)$ under $b \in \Delta(V^K)$. Then, $R(a, b) = p \cdot \sum_{\ell=1}^{N} \mathbf{1}_{\{v_\ell \geq p\}} b_\ell^k$, which is continuous in b. ($\mathbf{1}_{\{v_\ell \geq p\}}$ is the indicator function that takes values 1 and 0 if $v_\ell \geq p$ and $v_\ell < p$, respectively.) Given the continuity of $R(a, b)$ in b, $R(a', b_n) \geq R(a(b), b_n)$ for all n implies $R(a', b) \geq R(a(b), b)$. Thus, a' is optimal for the seller given b. This implies that $U(a(b), b) \geq U(a', b)$ by the seller's tie-braking rule. Also, $U(a(b_n), b_n) \geq U(a(b), b) + \varepsilon$ for all n implies $U(a', b) \geq U(a(b), b) + \varepsilon$, because $a(b_n) = a'$, and $U(a, b) = \sum_{\ell=1}^{N} \mathbf{1}_{\{v_\ell \geq p\}} (v_\ell - p) b_\ell^k$ is continuous in b. However, these two inequalities lead to $U(a(b), b) \geq U(a', b) \geq U(a(b), b) + \varepsilon$, which is a contradiction. Thus, $U(a(b), b)$ is upper semi-continuous in $b \in \Delta(V^K)$. By Theorem 15.5 of Aliprantis and Border (2006), $\int_{\Delta(V^K)} U(a(b), b) d\tau(b)$ is upper semi-continuous in

$\tau \in \Delta(\Delta(V^K))$ when $\Delta(\Delta(V^K))$ is endowed with weak* topology. This completes the first part.

Next, I show that the set $\mathcal{D}$ of all disclosure rules is weak* compact. To do so, let $b_0 := x_0 \times \cdots \times x_0$ denote the prior distribution of valuation vector. Also, given a disclosure rule $\phi \in \mathcal{D}$, I use $\hat{\phi} \in \Delta(\Delta(V^K))$ to mean the distribution over posterior beliefs about valuation vector u induced by ϕ and b_0. Moreover, let $\hat{\mathcal{D}} := (\hat{\phi})_{\phi \in \mathcal{D}} \subset \Delta(\Delta(V^K))$. Proposition 1 ((iii) $\Rightarrow$ (ii)) of Kamenica and Gentzkow (2011) implies that if the consumer's payoff $\int_{\Delta(V^K)} U(a(b), b) d\tau(b)$ is maximized at some $\tau^* \in \hat{\mathcal{D}} = \left\{\tau \in \Delta(\Delta(V^K)) : \int_{\Delta(V^K)} b d\tau(b) = b_0\right\}$, then there is a disclosure rule (with a finite message space) that maximizes his payoff among all available disclosure rules. Now, $\Delta(\Delta(V^K))$ is weak* compact because $\Delta(V^K)$ is compact (e.g., Theorem 15.11 of Aliprantis and Border (2006)). Also, $\hat{\mathcal{D}}$ is closed. To see this, take a sequence of disclosure rules τ_ℓ that converges to τ in weak* topology and satisfies $\int_{\Delta(V^K)} b\tau_\ell(b) = b_0$ for all ℓ. For each product k and value v_n, let b_n^k denote the probability that the value of product k is v_n under b. Because the projection map that takes b to b_n^k is continuous and bounded, $\int_{\Delta(V)} b_n^k \tau_\ell(b) \to \int_{\Delta(V)} b_n^k \tau(b)$ as $\ell \to +\infty$. Thus, $\int_{\Delta(V^K)} b\tau(b) = b_0$, which implies $\tau \in X$. Because the limit of any convergent sequence in $\hat{\mathcal{D}}$ belongs to $\hat{\mathcal{D}}$, $\hat{\mathcal{D}}$ is closed. Thus, $\hat{\mathcal{D}}$, which is a closed subset of a compact set $\Delta(\Delta(V^K))$, is weak* compact.

Finally, in equilibrium, the consumer solves $\max_{\tau \in \hat{\mathcal{D}}} \int_{\Delta(V^K)} U(a(b), b) d\tau(b)$. $\int_{\Delta(V^K)} U(a(b), b) d\tau(b)$ is upper semi-continuous in τ and $\hat{\mathcal{D}}$ is compact in weak* topology in $\Delta(\Delta(V^K))$. Thus, the set $\mathcal{D}^*$ of maximizers is nonempty and weak* compact. Finally, as I prove below, the seller's expected payoff $\int_{\Delta(V^K)} R(a(b), b) d\tau(b)$ is upper semi-continuous in τ. Therefore, $\max_{\tau \in \mathcal{D}^*} \int_{\Delta(\Delta(V^K))} R(a(b), b) d\tau(b)$ has a maximizer. Any maximizer τ^*, combined with the optimal on-path and off-path behavior of the seller and the consumer, consists of an equilibrium.

To see that $\int_{\Delta(\Delta(V^K))} R(a(b), b) d\tau(b)$ is upper semicontinuous, I show that $R(a(b), b)$ is upper semicontinuous in b. Suppose to the contrary that there exists $\varepsilon > 0$ and $(b_n)_{n=1}^{+\infty} \subset \Delta(V^K)$ such that $\lim_n b_n = b$ but $R(a(b_n), b_n) \geq R(a(b), b) + \varepsilon$ for all n. Because A is finite, we can choose a subsequence $(b_{n(m)})_{m=1}^{+\infty}$ so that for some $a' \in A$, $a(b_{n(m)}) = a'$ for all m. As $R(a, b)$ is continuous in b, we obtain $R(a', b) \geq R(a(b), b) + \varepsilon$. However, this contradicts, $R(a(b), b) \geq R(a', b)$. Thus, $R(a(b), b)$ is upper semi-continuous in b, and thus $\int_{\Delta(\Delta(V^K))} R(a(b), b) d\tau(b)$ is upper semi-continuous by Theorem 15.5 of Aliprantis and Border (2006). $\square$