Personalized Pricing and Privacy Choice*

Andrew Rhodes[†] Jidong Zhou[‡]

September 2021

(Very Preliminary Draft with Incomplete References)

Abstract

We study the impact of personalized pricing (or perfect price discrimination) in a general oligopoly model. Existing research based on the Hotelling model suggests that competitive personalized pricing intensifies competition, harms firms and benefits consumers. This result extends to our general model if the market is fully covered (i.e., all consumers buy). However, if the market is not fully covered, this result can be completely reversed—competitive personalized pricing can benefit firms and harm consumers. Nevertheless, in the long run with free entry, personalized pricing leads to the socially optimal market structure and so favors consumers. We also study consumer privacy choice in the context of personalized pricing. Due to an externality across consumers, too many consumers share their data relative to the consumer optimum, and more competition can harm consumers by increasing data sharing.

Key words: personalized pricing, competition, price discrimination, consumer privacy, consumer data

*We are grateful to Barry Nalebuff, Ben Polak, and Patrick Rey for helpful comments.

[†]Toulouse School of Economics. andrew.rhodes@tse-fr.eu

[‡]Yale School of Management. jidong.zhou@yale.edu

1 Introduction

Personalized pricing—the practice of charging different consumers different prices according to their willingness-to-pay—is rapidly attracting the attention of both practioners and policymakers. Once considered to be "of academic interest only" (Pigou, 1920), personalized pricing—or first-degree price discrimination, as it is also known—is increasingly feasible thanks to advances in information technology. In particular, consumer-level "big data" is being collected and exchanged via brokers on an unprecedented scale, and is then being analyzed by evermore sophisticated pricing algorithms. This enables firms to estimate consumers' willingness-to-pay with increasing accuracy, and then target them with personalized offers.

Anecdotal evidence of personalized pricing is abundant (see, e.g., Section 2.4 in OECD, 2018),[1] and has been documented in a wide range of sectors such as retailing, travel, ridesharing, and personal finance. However, formal evidence is more limited (as we discuss in the literature review), because firms often have incentives to disguise personalized prices so as to avoid consumer backlash. For example, personalized prices are often implemented via targeted discounts sent to consumers by email or smartphone app (OFT, 2013).[2]

One reason that policymakers may worry about personalized pricing is fairness—some (possibly vulnerable) consumers may end up paying unfairly higher prices than others. Another concern, which is more related to its economic effect, is that personalized pricing may harm overall consumer surplus—or in the language of the Council of Economic Advisers (2015), it "transfers value from consumers to shareholders." These often go hand-in-hand with broader concerns about consumer privacy. Policies such as the EU's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) have been touted as ways to protect privacy-conscious consumers, whilst at the same time allowing other consumers to benefit from personalized offers.

The goal of this paper is twofold. First, we examine the welfare effects of personal-

[1] For more examples, as well as the history of personalized pricing, see https://bit.ly/3A4Rk10 and https://bit.ly/38Ygzq6.

[2] Shiller (2021) argues that firms can also conceal personalized pricing via "sticky targeting pricing," whereby a seller fixes a new price for all consumers (not just the targeted consumer) over a short period of time. More broadly, for things such as financial services, the already-personalized nature of product offerings may make personalized prices hard to detect.

ized pricing in a general oligopoly model, aiming to reconcile two opposing views from the prior literature as detailed below. In particular, we highlight the importance of economic factors such as cost conditions and the degree of competition (which influence the extent of market coverage) in shaping the welfare impact of personalized pricing. Second, we study consumer privacy choice in the context of personalized pricing. In particular, we are interested in understanding whether consumer privacy policies, which enable consumers to control their personal information, can reach an optimal balance between protecting privacy-conscious consumers whilst allowing others to make use of personalized pricing.

Section 2 reviews two well-known benchmarks in the literature. In the monopoly case personalized pricing enables more consumers to buy a product, but it also helps the firm extract *all* the surplus from consumers. As a result, relative to uniform pricing, personalized pricing benefits the firm but hurts consumers. In the arguably more realistic case of competition, however, Thisse and Vives (1988) show that in the classic Hotelling model, competitive personalized pricing lowers the price *every* consumer pays. (Each firm offers low prices to target consumers who prefer its rival's product. This, in turn, also forces the rival to drop its prices for its customer base.) Therefore, going from monopoly to duopoly completely reverses the impact of personalized pricing; it now harms firms and benefits consumers.

The first contribution of this paper is to reconcile these two opposing views. In Section 3 we introduce a discrete-choice model which nests both monopoly and Hotelling as special cases. In particular, there is an arbitrary number of (single-product) firms, and consumers' valuations for their products are drawn from a (symmetric) joint distribution. Consumers either buy one of the products or take an outside option. Our model is based on Perloff and Salop (1985), but is more general because it allows for correlated valuations and partial market coverage. Under uniform pricing firms cannot use information about individual consumers' preferences, and so offer all consumers the same price. Under personalized pricing firms know each consumer's valuations for all the products, and make personalized offers accordingly.

Section 4 compares market performance in these two regimes. Under a mild regularity condition, we show that the result in Thisse and Vives (1988) that all personalized prices are lower than the uniform price fails: consumers who strongly prefer one product over others pay more than the uniform price, whilst those who regard products (at least the best two) as close substitutes pay less. Nevertheless, if the market is fully covered under uniform pricing, competitive personalized pricing does lower industry profit and increase aggregate consumer surplus under a log-concavity condition. (The log-concavity condition, or equivalently the increasing-hazard-rate condition, ensures that there are relatively few consumers with strong preferences compared to those with weak preferences.) We therefore significantly generalize the welfare result in Thisse and Vives (1988).

We then show, however, that if the market is not fully covered, the impact of personalized pricing can be totally reversed. In particular, competitive personalized pricing can increase industry profit but lower consumer surplus, as in the monopoly case. This always happens—irrespective of the number of firms in the market—when the production cost is sufficiently high, such that relatively few consumers purchase under either pricing regime. Indeed, when product valuations are independent and follow an exponential distribution, it happens whenever marginal cost is such that the market is not fully covered under uniform pricing. Using numerical examples, we also show that the welfare impact of personalized pricing follows a cut-off rule for common distributions such as the extreme value (which generates the logit model) and the Normal (which generates the probit model). Specifically, for a given number of firms, when the production cost is sufficiently low the impact is similar to Thisse and Vives (1988), when the cost is sufficiently high the impact is similar to monopoly, and when the cost is intermediate personalized pricing benefits both consumers and firms.[3] Similarly, for a fixed production cost, with few firms the impact is like in monopoly, with many firms the impact is similar to Thisse and Vives (1988), and for an intermediate number of firms industry profit and consumer surplus both increase.

The intuition for why competitive personalized pricing can benefit firms and harm consumers is as follows. First, consider the case where all consumers value each product above marginal cost. Here, partial coverage arises when the uniform price excludes some low-valuation consumers from the market. Personalized pricing brings these consumers into the market, but since they have low valuations the positive effect on their surplus is relatively small. On the other hand, consumers who bought under uniform pricing have a high valuation for at least one product, and so amongst them there are relatively more

[3]Note that if uniform pricing leads to only partial coverage, personalized pricing increases total surplus by expanding the market. This explains why profit and consumer surplus can both increase.

strong-preference consumers—meaning that personalized pricing can raise the average price they pay. When this happens, personalized pricing can make consumers worse off overall, even though it expands demand. Second, consider the case where for each product some consumers have a valuation below marginal cost. Now each firm faces a new "monopoly segment" of consumers who value its product above cost but value all other products below cost. This is an additional force for personalized pricing to harm consumers.

We also examine the long-run impact of personalized pricing, by endogenzing the number of firms in the market. In particular, we suppose that firms choose whether or not to pay a fixed cost to enter, and then engage in price competition. We show that if the entry of a new product does not change consumers' preferences over existing products, then with personalized pricing the new entrant fully extracts the increase in match efficiency caused by its entry. Consequently, in the long run, personalized pricing leads to the socially optimal market structure. If we ignore integer constraints, this implies that (i) personalized pricing must benefit consumers in the long run relative to uniform pricing, and (ii) if the market is fully covered, uniform pricing leads to excessive entry because for a fixed number of firms it leads to higher profit.[4]

The second contribution of this paper is to study the interaction between personalized pricing and consumer privacy choice. In Section 5 we extend the model by allowing consumers to costlessly hide their data (e.g., due to privacy policies such as GDPR and CCPA). Consumers that choose to share their data incur a privacy cost, which reflects, e.g., their concerns about data security. We assume that consumers make their privacy choices before they learn their valuations for the products in the marketplace; this captures the idea that consumers usually need to accept or reject cookies on websites (such as media sites or blogs) that are not directly related to products which they may buy in the future. We also assume that firms offer a common list price to consumers who do not share their data, and then offer personalized discounts to consumers who share their data; this captures the idea that consumers can check a public list price before deciding whether or not to accept a targeted offer.

Using this framework, we identify a novel externality across consumers. Specifically,

[4]The excessive-entry result under uniform pricing is also proved in Anderson, de Palma, and Nesterov (1995) and Tan and Zhou (2021) when consumer valuations are independent across products. We prove this result in a more general setup and our proof is simpler.

when more consumers share their data, firms raise their public list price to facilitate making personalized offers, which then hurts consumers who pay the list price (including those who hide their data). Although a privacy policy such as GDPR benefits consumers, we demonstrate that because of the externality (i) too many consumers share their data relative to the consumer optimum, and (ii) more competition can harm consumers by inducing more consumers to share their data.

Related literature. The literature on price discrimination is extensive, but it mainly focuses on imperfect price discrimination.[5] One exception is the study of spatial price discrimination, where firms can charge customers in different locations different prices. An important contribution to this literature is Thisse and Vives (1988) (which can also be reinterpreted as a model of competitive personalized pricing).[6] They consider a two-stage game where firms first choose whether or not to price discriminate and then compete in prices. Using a Hotelling model with uniformly distributed consumers, they show that discriminatory pricing is a dominant strategy for each firm, and so the unique equilibrium features price discrimination. When firms have the same cost, as discussed earlier, they are trapped in a Prisoner's dilemma because every personalized price is below the uniform price.[7]

The Hotelling setup in Thisse and Vives (1988) has been widely used in the subsequent literature. For example, Shaffer and Zhang (2002) use it to study personalized pricing when one firm has a brand advantage over the other, while Chen and Iyer (2002) use it to study personalized pricing when firms first need to advertise to reach consumers. Montes, Sand-Zantman, and Valletti (2019) use it to study whether a monopolistic data intermediary should sell data to one or both competing firms who can use the data to conduct personalized pricing. Chen, Choe, and Matsushima (2020) use it to study consumer identity management which helps consumers avoid being ex-

[5] See the survey papers by Varian (1989), Armstrong (2007), Fudenberg and Villas-Boas (2007), and Stole (2007).

[6] Another related contribution is Anderson, de Palma, and Thisse (1989), who incorporate logit product differentiation into the Hotelling model. They show numerically that the welfare result in Thisse and Vives (1988) remains unchanged.

[7] When firms have different costs, the low-cost firm can earn more than under uniform pricing, but within the parameter range in Thisse and Vives (1988) industry profit is still lower and consumer surplus is still higher under discriminatory pricing.

ploited by firms via personalized pricing.[8] In all these studies, an implicit underlying assumption is that competitive personalized pricing in the benchmark case intensifies competition, harms firms and benefits consumers. Our paper shows that this is not necessarily true in a more general model which allows for partial market coverage.

Our paper is also closely related to Anderson, Baik, and Larson (2019) (ABL thereafter), who also use a general discrete-choice framework (but with full market coverage) to study competitive personalized pricing. One important difference is that our paper allows for partial market coverage, and emphasizes that this can qualitatively change the impact of personalized pricing. Another important difference is that in our paper firms can freely offer personalized prices, leading to a relatively simple pure-strategy pricing equilibrium; ABL, by contrast, assume that it is costly for firms to send targeted discounts, which leads to a mixed strategy equilibrium in both pricing and advertising.[9] (Our modeling assumption captures the idea that the cost of making personalized offers is mainly a fixed cost, due to investments in buying consumer data and developing pricing algorithms.) ABL also study consumer privacy choice, but they use a different timing and have a different focus—for example, we stress suboptimality of consumer privacy choices and the possibility that competition can have perverse effects on consumer surplus, issues which are not studied by ABL.

There is also growing empirical research on personalized pricing. One strand attempts to find evidence of firms engaging in personalized pricing. Detecting personalized pricing is usually hard because sellers have incentives to disguise personalized offers, but nevertheless there is some suggestive evidence. For instance, Hannak et al. (2014) find evidence of some form of personalization on 9 out of 16 e-commerce sites in their study, while Aparicio et al. (2021) document evidence that increasing use of algorithmic pricing is associated with increasing price differentiation (for the same product at the same time but across different delivery zipcodes). The other strand of the empir-

[8] The interaction between list prices and personalized prices in Chen, Choe, and Matsushima (2020) is similar as in our extended model with privacy choice. But the details of their model and their main messages differ from ours. For example, in their model consumers' ability to hide their identity softens price competition, while in our model hiding consumer data intensifies competition.

[9] Such randomized offers cause consumers to sometimes buy the wrong product and so harm match efficiency. As a result, in ABL personalized pricing can make both firms and consumers worse off compared to uniform pricing. In contrast, in our model it is possible that both firms and consumers get better off under personalized pricing as it can expand demand.

ical literature assesses the impact of personalized pricing (see, e.g., Waldfogel (2015), Dube and Misra (2019), Shiller (2020), and Kehoe et al. (2020)). For instance, Shiller (2020) shows that if Netflix could use consumer information from rich web-browsing data to implement price discrimination, its profit could increase by about 13%, while the profit improvement would be tiny if it only relied on demographic information.

Our paper is also related to the burgeoning literature on consumer data and privacy (see, e.g., Acquisti et al. (2016), and Bergemann and Bonatti (2019) for surveys). First, we study consumer privacy choice in the context of personalized pricing. There are many papers on this topic, but they explore different questions, and different from us they typically assume that consumers know their product valuations when deciding whether to share their preference information.[10] We contend that privacy choice is not usually product-specific, because in many circumstances consumers have little idea or control over how (and for which products) their data will be used in the future. Second, our privacy choice model highlights an externality across consumers through the impact of data sharing on the list price other consumers may pay. This externality is different from the "social data externality" emphasized in a series of recent papers (Choi et al. (2019), Acemoglu et al. (2020), Bergemann et al. (2020), and Ichihashi (2021)). There the idea is that when consumers' preferences are correlated, one consumer's data sharing often diminishes the value of other consumers' data. This enables the data intermediary to acquire consumer data cheaply, resulting in too much data sharing.

2 Two Benchmarks: Monopoly and Hotelling

The impact of personalized pricing (or perfect price discrimination) under *monopoly* is straightforward. Suppose consumers wish to buy at most one unit of a product, and have heterogeneous valuations for it. Under uniform pricing, the firm sets a standard monopoly price. Consumers who value the product more than the monopoly price buy and obtain positive surplus; all other consumers are excluded from the market. Under

[10] See, e.g., Anderson, Baik, and Larson (2019), Ali, Lewis, and Vasserman (2020), Chen, Choe, and Matsushima (2020), and Ichihashi (2020). For instance, Section 4 of Ali et al. (2020) shows, in a similar oligopoly discrete-choice model but with full market coverage, that if consumers can disclose preference information via *private* firm-specific messages, there exists a partial revelation equilibrium in which personalized pricing benefits each consumer relative to uniform pricing. Hence the Thisse and Vives (1988) welfare result holds with endogenous information disclosure.

personalized pricing, each consumer with a valuation above marginal cost is offered a personalized price exactly equal to their valuation, and they all buy. As a result, total surplus is maximized but it is fully extracted by the monopolist. Personalized pricing therefore increases total welfare and firm profit but reduces consumer surplus.

The other well-known case is the linear Hotelling *duopoly* model studied in Thisse and Vives (1988). Suppose consumers are uniformly distributed along a unit-length Hotelling line. Suppose the two firms have cost normalized to zero, with firm 1 located at the leftmost point on the line, and firm 2 located at the rightmost point. A consumer located at $x \in [0,1]$ on the line values firm 1's product at $v_1 = V - x$ and firm 2's product at $v_2 = V - (1-x)$, where V is assumed to be large enough that the market is fully covered in equilibrium. Under uniform pricing, firms set the standard Hotelling price of 1. Under personalized pricing, the firms compete for each consumer individually. Consumers with location $x < 1/2$ prefer product 1, while consumers with location $x > 1/2$ prefer product 2. Firms therefore engage in asymmetric Bertrand competition, with equilibrium price schedules given respectively by

$$\begin{array}{l} p_1(x) = v_1 - v_2 = 1 - 2x \;\; \text{and} \;\; p_2(x) = 0 \;\; \text{for} \;\; x \in [0, \frac{1}{2}] \\ p_1(x) = 0 \;\; \text{and} \;\; p_2(x) = v_2 - v_1 = 2x - 1 \;\; \text{for} \;\; x \in [\frac{1}{2}, 1] \end{array}, \tag{1}$$

where $p_i(x)$ is the price offered by firm $i = 1, 2$ to the consumer at x, and each consumer buys her preferred product. Importantly, each consumer pays (weakly) less under personalized pricing because $p_i(x) \leq 1$. As a result—and in contrast to monopoly—personalized pricing harms firms and benefits consumers. (Under both uniform and personalized pricing the market is fully covered and consumers buy their preferred product, so personalized pricing has no impact on total welfare.)

However, the observation that each personalized price is lower than the uniform price can easily be overturned. To see this, suppose instead that consumers are distributed along the Hotelling line according to a symmetric and strictly log-concave (so single-peaked) density. Discriminatory prices are the same as in (1), but the uniform price, which equals 1 over the density of consumers at $x = 1/2$, is now strictly below 1. As a result, consumers near the two ends of the line (i.e., those with relatively strong preferences) now pay more under personalized pricing, while consumers near the middle of the line (i.e., those with relatively weak preferences) still pay less.[11] The impact of

[11] Armstrong (2007) makes the same point by considering a specific non-uniform distribution.

personalized pricing on industry profit and (aggregate) consumer surplus is then less clear. In the next section, we develop a more general oligopoly model—which includes the Hotelling model as a special case—and investigate to what extent competitive personalized pricing is overall pro-competitive.

3 A General Oligopoly Model

There are n firms, each supplying a differentiated product at constant marginal cost c. There is also a unit mass of consumers, and each consumer wishes to buy at most one of the products. If a consumer buys nothing she obtains an outside option with zero surplus.[12] Let $\mathbf{v} = (v_1, ..., v_n) \in \mathbf{R}^n$ denote a consumer's valuations for the n products. In the population $\mathbf{v}$ is distributed according to a symmetric joint CDF $\tilde{F}(\mathbf{v})$, with corresponding density function $\tilde{f}(\mathbf{v})$. (This implies that there are no systematic quality differences across products.) Let F and f be respectively the common marginal CDF and density function of each v_i, and let $[\underline{v}, \overline{v}]$ be its support. (We need $c < \overline{v}$ to have an active market.) To ease the exposition, we assume that $\tilde{F}$ has full support on $[\underline{v}, \overline{v}]^n$, but this is not crucial for the main results.

Note that although we allow a consumer's valuations for different products to be correlated,[13] sometimes we focus on the IID case where the v_i's are independent across products (which is the leading case in the literature on random-utility oligopoly models).

We consider two different pricing regimes. Under uniform pricing, firms set the same price for every consumer (either because they have no data on consumer preferences, or are forbidden from using it). Under personalized pricing, firms perfectly observe each consumer's vector of valuations $\mathbf{v} = (v_1, ..., v_n)$, and offer them a personalized price.

Notation. It will be convenient to introduce the following notation. Let $G(\cdot|v_i)$ and $g(\cdot|v_i)$ be respectively the CDF and density function of $\max_{j \neq i}\{v_j\}$ conditional on v_i. Let $v_{n:n}$ and $v_{n-1:n}$ be the highest and second-highest order statistics, and let $F_{(n)}(v)$

[12] To have our model cover both the case of full market coverage and the case of partial market coverage, we have chosen to normalize the outside option and vary the marginal cost c. The same qualitative insights obtain if instead we normalize the production cost and vary the outside option.

[13] In the duopoly case, our set-up nests Hotelling with a symmetric location distribution if v_1 and v_2 are large enough (to cover the market) and we treat $v_1 - v_2$ as a consumer's location. For some location distributions (e.g., the uniform one), we need a particular correlation structure over (v_1, v_2).

and $F_{(n-1)}(v)$ be their respective CDFs. Then

$$F_{(n)}(v) = \tilde{F}(v, ..., v) = \int_{\underline{v}}^{v} G(v|v_i)dF(v_i) \tag{2}$$

and

$$F_{(n-1)}(v) = F_{(n)}(v) + n\int_{v}^{\overline{v}} G(v|v_i)dF(v_i)\ . \tag{3}$$

To understand $F_{(n-1)}(v)$, notice that for the second-highest valuation to be below v, either all the v_i's must be less than v, or exactly one of them must be above v and the others be below v. In the IID case we have $\tilde{F}(\mathbf{v}) \overset{\text{IID}}{=} \prod_{i=1}^{n} F(v_i)$, $G(v|v_i) \overset{\text{IID}}{=} F(v)^{n-1}$, $F_{(n)}(v) \overset{\text{IID}}{=} F(v)^n$, and

$$F_{(n-1)}(v) \overset{\text{IID}}{=} F(v)^n + n(1 - F(v))F(v)^{n-1}\ .$$

In order to solve the uniform pricing game, it is useful to define the random variable

$$x_z \equiv v_i - \max_{j\neq i}\{z, v_j\}\ , \tag{4}$$

where z is a constant. Let $H_z(x)$ and $h_z(x)$ be respectively the CDF and density function of x_z. More explicitly,

$$1 - H_z(x) = \Pr[v_i - x > \max_{j\neq i}\{z, v_j\}] = \int_{z+x}^{\overline{v}} G(v_i - x|v_i)dF(v_i)\ , \tag{5}$$

and

$$h_z(x) = G(z|z + x)f(z + x) + \int_{z+x}^{\overline{v}} g(v_i - x|v_i)dF(v_i)\ . \tag{6}$$

When z is irrelevant (i.e., when $z \leq \underline{v}$), let $H(x)$ and $h(x)$ be respectively the CDF and density function of $x \equiv v_i - \max_{j\neq i}\{v_j\}$. They are for the case of full market coverage.[14]

3.1 Uniform pricing

We first study the regime of uniform pricing, where firms are unable to price discriminate. We focus on a symmetric pure-strategy pricing equilibrium, and let p denote the equilibrium uniform price.[15] Recalling the definition of $H_z(x)$ in equation (5), when firm i unilaterally deviates to a price p_i its deviation demand is

$$\Pr[v_i - p_i > \max_{j\neq i}\{0, v_j - p\}] = \Pr[v_i - \max_{j\neq i}\{p, v_j\} > p_i - p] = 1 - H_p(p_i - p)\ ,$$

[14] Anderson, Baik, and Larson (2019) also use such notation to simplify demand expressions.

[15] If the joint density $\tilde{f}$ is log-concave, the pricing equilibrium is unique and symmetric in the duopoly case (Caplin and Nalebuff (1991)) and in the IID case (Quint (2014)).

and its deviation profit is $(p_i - c)[1 - H_p(p_i - p)]$. It is clear that a firm will never set a price below marginal cost c or above the maximum valuation $\overline{v}$.

To ensure that the uniform pricing equilibrium is uniquely determined by the first-order condition, we make the following assumption:

Assumption 1. $1 - H_z(x)$ *is log-concave in* x *and* $\frac{1-H_z(0)}{h_z(0)}$ *is non-increasing in* z.

In the Appendix we report some primitive conditions under which this assumption holds. For example, the first condition holds if the joint density $\tilde{f}$ is log-concave (Caplin and Nalebuff (1991)), and both conditions hold in the IID case with a log-concave f. (The second condition must hold if, for $z < z'$, x_z is greater than $x_{z'}$ in the sense of hazard rate dominance.) Assumption 1 also holds in the Hotelling case (see footnote 13) provided that $v_1 - v_2$ has a log-concave density.

Given the first condition in Assumption 1, firm i's deviation profit is log-concave in p_i, and so p must solve the first-order condition

$$p - c = \phi(p) \ , \tag{7}$$

where

$$\phi(p) \equiv \frac{1 - H_p(0)}{h_p(0)} = \frac{\int_p^{\overline{v}} G(v|v)dF(v)}{G(p|p)f(p) + \int_p^{\overline{v}} g(v|v)dF(v)} \ . \tag{8}$$

To interpret this, note that $1 - H_p(0)$ is each firm's equilibrium demand,[16] while $h_p(0)$ is the absolute value of the equilibrium demand slope and it measures how many consumers are marginal for each firm. (The first term in $h_p(0)$ captures the extensive margin, and the second term captures the intensive margin due to competition.)

Given the second condition in Assumption 1, $\phi(p)$ is non-increasing. Notice also that $\phi(p)$ is constant for $p \leq \underline{v}$. Then one can show the following result.

Lemma 1. *Suppose Assumption 1 holds.*
(i) If $c \leq \underline{v} - \phi(\underline{v})$*, the equilibrium uniform price satisfies*

$$p - c = \phi(\underline{v}) = \frac{1/n}{\int_{\underline{v}}^{\overline{v}} g(v|v)dF(v)} \tag{9}$$

and $p \leq \underline{v}$*, such that the market is fully covered in equilibrium.*
(ii) Otherwise, the equilibrium uniform price uniquely solves (7) and $p > \underline{v}$*, such that the market is not fully covered in equilibrium.*

[16] Due to firm symmetry we can also write that $1 - H_p(0) = \frac{1}{n}[1 - F_{(n)}(p)]$.

Intuitively, when cost is relatively low ($c \le \underline{v} - \phi(\underline{v})$), marginal consumers are sufficiently valuable that firms choose to cover the whole market. On the other hand, when cost is relatively high ($c > \underline{v} - \phi(\underline{v})$), firms optimally exclude some low-valuation consumers. Note that a sufficient (but by no means necessary) condition for partial coverage is that $\underline{v} \le c$, i.e., some consumers value a product less than marginal cost.

The literature on random-utility oligopoly models usually studies the IID case, such that $G(v|v) \overset{\text{IID}}{=} F(v)^{n-1}$ and $g(v|v) \overset{\text{IID}}{=} (n-1)F(v)^{n-2}f(v)$, and so

$$\phi(p) \overset{\text{IID}}{=} \frac{[1-F(p)^n]/n}{F(p)^{n-1}f(p) + \int_p^{\overline{v}} f(v)dF(v)^{n-1}} \ . \tag{10}$$

Most papers further assume that the market is covered (e.g., Perloff and Salop (1985), and Gabaix et al. (2016)), in which case $\phi(p)$ simplifies to $1/[n\int_{\underline{v}}^{\overline{v}} f(v)dF(v)^{n-1}]$.[17]

Example: uniform distribution. Suppose the v_i's are independent and uniformly distributed on $[\underline{v}, \underline{v}+1]$. Here, if $p \le \underline{v}$ then $\phi(p) = \frac{1}{n}$, and if $p > \underline{v}$ then $\phi(p) = \frac{1}{n}[1-(p-\underline{v})^n]$. Therefore if $c + \frac{1}{n} \le \underline{v}$ the market is fully covered and the equilibrium price is $p = c + \frac{1}{n}$; otherwise the market is not fully covered and $p > \underline{v}$ uniquely solves

$$p - c = \frac{1-(p-\underline{v})^n}{n} \ . \tag{11}$$

Example: exponential distribution. Suppose the v_i's are independent and exponentially distributed with $F(v) = 1 - e^{-(v-\underline{v})}$ on $[\underline{v}, \infty)$. Here we have $\phi(p) = 1$, and so the equilibrium price is $p = c + 1$ regardless of whether or not the market is covered (and irrespective of the number of firms).[18] Therefore the market is fully covered in equilibrium if and only if $c + 1 \le \underline{v}$.

Industry profit under uniform pricing is

$$\Pi_U \equiv n(p-c)[1-H_p(0)] = (p-c)[1-F_{(n)}(p)] \ , \tag{12}$$

where $1 - F_{(n)}(p)$ is the measure of consumers who value at least one product above price p. Since all consumers buy their favorite product as long as it has a positive

[17] An exception is Section 4.2 of Zhou (2017), which shows that in the IID case $\phi(p)$ is decreasing and the equilibrium price decreases in n if f is log-concave.

[18] Using integration by parts, the denominator in (10) can be rewritten as $f(\overline{v}) - \int_p^{\overline{v}} F(v)^{n-1}f'(v)dv$. For the exponential distribution $f(\overline{v}) = 0$ and $f(v) = -f'(v)$, so this equals the numerator of (10).

surplus, (aggregate) consumer surplus is

$$V_U \equiv \mathbb{E}[\max\{0, v_{n:n} - p\}] = \int_p^{\overline{v}} (v-p)dF_{(n)}(v) = \int_p^{\overline{v}} [1 - F_{(n)}(v)]dv \ , \tag{13}$$

where the last equality is from integration by parts. Notice that these expressions are valid regardless of whether or not the market is fully covered.

3.2 Personalized Pricing

Now consider the regime where firms perfectly observe each consumer's vector of valuations $\mathbf{v} = (v_1, ..., v_n)$ and set personalized prices accordingly. In this case, firms engage in a game of asymmetric Bertrand competition for each consumer. To rule out uninteresting equilibria, we assume that firms do not play dominated strategies, and that when consumers are indifferent between several offers they buy the product with the highest valuation (so that total welfare is maximized). Consider a consumer who values, say, firm 1's product the highest and firm 2's product the second highest. Competition is then essentially between these two firms. Suppose $v_1 \geq c$ (otherwise the consumer takes the outside option). Competition forces firm 2 to price at marginal cost. Firm 1 prices at $c + v_1 - v_2$ if $v_2 \geq c$, and otherwise acts as a monopolist and charges v_1. In both cases the other $n-2$ firms charge weakly more than c, and firm 1 sells to the consumer. (Since the prices of these $n-2$ firms can be anything above c, there are multiple equilibria, but they are all outcome-equivalent.) To ease the exposition, we henceforth focus on the equilibrium in which these other $n-2$ firms charge c. Firm i's equilibrium pricing schedule can then be written as:

$$p(v_i, \mathbf{v}_{-i}) = \begin{cases} v_i - \max_{j \neq i}\{0, v_j - c\} & \text{if } v_i \geq \max_{j \neq i}\{c, v_j\} \\ c & \text{otherwise} \end{cases} \tag{14}$$

where $\mathbf{v}_{-i}$ denotes a consumer's valuations for all products other than i. Intuitively, if a firm's product is a consumer's favorite and has a valuation above cost, the firm charges the consumer a price equal to the difference between her valuation for its product and that of the best alternative (which is either the outside option, or the best rival product sold at marginal cost).[19] Note that if $\underline{v} \geq c$, the "max" constraint from the outside

[19]Note that a consumer-specific pricing schedule includes uniform pricing as a special case. Therefore, in an extended game where firms simultaneously choose whether to adopt discriminatory pricing

option in (14) is irrelevant and so all consumers buy their favorite product in equilibrium (i.e., the market is fully covered under personalized pricing).

To calculate profit, notice that when firm i wins a consumer, its profit margin is $p(v_i, \mathbf{v}_{-i}) - c = v_i - \max_{j \neq i}\{c, v_j\} = x_c$, where x_c has a CDF $H_c(x)$. That is, firm i earns x_c whenever $x_c \geq 0$, so its equilibrium profit is $\int_0^\infty x dH_c(x)$, and industry profit under perfect price discrimination is

$$\Pi_D = n \int_0^\infty x dH_c(x) = n \int_0^\infty [1 - H_c(x)] dx \ . \tag{15}$$

(There are alternative ways to calculate profit as we will show later.)

Consumers always buy their favorite product (as long as it is valued above c). Given the equilibrium pricing schedule in (14), it is clear that a consumer's surplus is equal to either the outside option or the surplus of the second-best product when sold at marginal cost. Therefore, consumer surplus under perfect price discrimination is

$$V_D \equiv \mathbb{E}[\max\{0, v_{n-1:n} - c\}] = \int_c^{\overline{v}} (v - c) dF_{(n-1)}(v) = \int_c^{\overline{v}} [1 - F_{(n-1)}(v)] dv \ . \tag{16}$$

Notice that expressions (15) and (16) are both valid regardless of whether or not $c < \underline{v}$.

4 The Impact of Personalized Pricing

We now examine how a shift from uniform to personalized pricing affects market performance. We first study the short-run impact, when the number of firms n is taken as given. In the case of full market coverage, we significantly generalize the insight from Thisse and Vives (1988), showing that competitive personalized pricing harms firms and benefits consumers (on average). However, we also show that when the market is only partially covered, their insight can be completely overturned. We then study the long-run impact, when n is determined by firms' free-entry decisions.

4.1 The short-run impact with a fixed market structure

Suppose the number of firms is fixed. Our first result shows that, under a mild regularity condition, the highest personalized price exceeds the uniform price. As a result, some

and what prices to offer, the only equilibrium is that all firms adopt personalized pricing. Thisse and Vives (1988) also consider such a one-stage game and make the same point in their duopoly model.

consumers benefit from personalized pricing while others suffer. Recall that $h(x)$ is the density of $v_i - \max_{j \neq i}\{v_j\}$.

Lemma 2. *Suppose Assumption 1 holds, and that $h(x) < h(0)$ for $x > 0$. Then the highest personalized price exceeds the uniform price.*

Proof. Using equation (14) the highest personalized price is $p_{\max} = \overline{v} - \max\{0, \underline{v} - c\}$. If $\underline{v} \leq c$, $p_{\max} = \overline{v}$ and so it must exceed the uniform price. If $\underline{v} > c$, $p_{\max} = c + \overline{v} - \underline{v}$ and so $p < p_{\max}$ if and only if $p - c < \overline{v} - \underline{v}$. Under Assumption 1, $\phi(p)$ is decreasing and so the uniform price must satisfy $p - c = \phi(p) \leq \phi(\underline{v}) = \frac{1}{nh(0)}$. At the same time,

$$\frac{1}{n} = \int_0^{\overline{v}-\underline{v}} h(x)dx < h(0)(\overline{v} - \underline{v}) \ ,$$

where the equality is from the fact that the probability of $v_i \geq \max_{j \neq i}\{v_j\}$ is $\frac{1}{n}$, and the inequality is from the assumption that $h(x) < h(0)$ for $x > 0$. Therefore we have $p - c < \overline{v} - \underline{v}$. □

Note that the hypotheses of the lemma hold, for example, in the IID case with a log-concave f. However they fail in the linear Hotelling model studied earlier in Section 2, because in that case $h(x)$ is a constant in $x \geq 0$ (which explains why, in that case, the highest personalized price exactly equals the uniform price).

The remainder of this subsection addresses the subtler question of how price discrimination affects profit and aggregate consumer surplus.

4.1.1 The case of full market coverage

We first study the case where the market is fully covered under uniform pricing, i.e., where $p \leq \underline{v}$. From Lemma 1, this happens when $c \leq \underline{v} - \phi(\underline{v})$. This condition in turn implies $c < \underline{v}$, which means that the market is also fully covered under personalized pricing. Total welfare is therefore the same under uniform and personalized pricing, because in both cases all consumers buy their preferred product. The following result reports the impact of personalized pricing on profit and consumer surplus.

Proposition 1. *Suppose Assumption 1 holds and $c \leq \underline{v} - \phi(\underline{v})$ (in which case the market is fully covered under both pricing regimes). Then for any $n \geq 2$, relative to uniform pricing, personalized pricing harms firms and benefits consumers.*

Proof. Under the stated full-coverage condition, $x_z = v_i - \max_{j \neq i}\{z, v_j\}$ simplifies to $x = v_i - \max_{j \neq i}\{v_j\}$ for both $z = p$ and $z = c$ as $c < p \leq \underline{v}$. Recall that H and h are respectively the CDF and density function of x. Then industry profit under uniform pricing is

$$\Pi_U = p - c = \frac{1}{nh(0)} ,$$

while under personalized pricing it is

$$\Pi_D = n \int_0^{\infty} [1 - H(x)] dx = n \int_0^{\infty} \frac{1 - H(x)}{h(x)} dH(x) \leq n \frac{[1 - H(0)]^2}{h(0)} = \frac{1}{nh(0)} .$$

The inequality follows because, under Assumption 1, $1 - H$ is log-concave and therefore $\frac{1-H}{h}$ is decreasing. The final equality follows because firm symmetry implies $1 - H(0) = \frac{1}{n}$. Therefore, firms suffer from personalized pricing. Since total welfare is unchanged, consumers benefit from personalized pricing. □

The intuition for this result is as follows. Notice that consumers with a relatively small gap between their top two valuations pay less under personalized pricing, while the reverse is true for consumers with a relatively large gap between their top two valuations. Under log-concavity (in Assumption 1) there are relatively more of the former consumers, and so personalized pricing harms firms but benefits consumers in aggregate. Note that since our set-up includes Hotelling as a special case (see footnote 13), Proposition 1 significantly generalizes the result in Thisse and Vives (1988).[20]

From the proof, we can also see that if $1 - H(x)$ is log-linear, price discrimination has *no* impact on profit and consumer surplus. This edge case arises with IID exponentially distributed valuations, where $\Pi_U = \Pi_D = 1$.[21]

[20] Although not highlighted in Anderson, Baik, and Larson (2019), this generalization of Thisse-Vives is also implied by their Proposition 6 which does comparative statics with respect to the advertising cost in their model. Our proof is similar to that of Proposition 7 in ABL which shows the opposite result when $1 - H$ is log-convex. If we were to instead work directly with the primitive valuation distribution $\tilde{F}$, this result would be considerably harder to prove, even in the IID case.

[21] On page 13 we showed that, in the IID exponential case with $F(v) = 1 - e^{-(v - \underline{v})}$, the uniform price is $c + 1$ and so with a covered market $\Pi_U = 1$. To compute Π_D, use equation (15) and

$$1 - H(x) = \int_{\underline{v}+x}^{\infty} F(v - x)^{n-1} dF(v) = \int_{\underline{v}}^{\infty} F(t)^{n-1} f(t + x) dt = e^{-x} \int_0^1 (1 - u)^{n-1} du = \frac{e^{-x}}{n} ,$$

where the third equality is from changing the integral variable from t to $u = e^{-(t - \underline{v})}$.

Finally, it is interesting to consider how the impact of personalized pricing changes as valuations become more or less correlated. A general investigation into this issue appears hard, but it is clear that in the limit with perfectly (positively) correlated valuations the impact should disappear. The following example suggests that more correlation in valuations weakens the impact of personalized pricing.

Example: bivariate normal distribution. Suppose v_1 and v_2 are bivariate normal with means μ, variances σ^2, and correlation coefficient ρ.[22] (This distribution has a log-concave joint density provided $\rho \in (-1, 1)$.) Note that $v_1 - v_2$ is normally distributed with mean 0, variance $2\sigma^2 (1-\rho) \equiv \tau^2$, and density function $h(x) = \frac{1}{\sqrt{2\pi}\tau} e^{-\frac{x^2}{2\tau^2}}$. Under uniform pricing, industry profit is

$$\Pi_U = \frac{1}{2h(0)} = \tau\sqrt{\frac{\pi}{2}} \ ,$$

while under personalized pricing, industry profit is

$$\Pi_D = 2\int_0^\infty xh(x)dx = -2\tau^2\int_0^\infty h'(x)dx = 2\tau^2 h(0) = \tau\sqrt{\frac{2}{\pi}} \ ,$$

where the first equality follows because $xh(x) = -\tau^2 h'(x)$. (Notice that Π_U and Π_D both decrease in ρ. Intuitively, as correlation increases, products become less differentiated and so price competition is fiercer.) The reduction in profit—or equivalently, the increase in consumer surplus—due to personalized pricing is

$$\tau\left(\sqrt{\frac{\pi}{2}} - \sqrt{\frac{2}{\pi}}\right) \ ,$$

which decreases in ρ and, as expected, goes to zero as $\rho \to 1$.

4.1.2 The case of partial market coverage

We now turn to the (perhaps more realistic) case where the market is not fully covered under uniform pricing. From Lemma 1, we know this happens when $c > \underline{v} - \phi(\underline{v})$. One simple impact of personalized pricing is that it now expands total demand: under uniform pricing, a consumer buys if the best match is above the uniform price $p > c$;

[22] Strictly speaking it is impossible to have full market coverage in this example, but it is almost the case if μ is large. Given full market coverage, c does not enter the expressions for Π_U and Π_D.

under personalized pricing, a consumer buys if the best match is above c. Personalized pricing therefore strictly improves total welfare.

Before investigating the impact on firms and consumers, we offer an alternative formula to calculate industry profit under personalized pricing, which is more convenient to use in some of the subsequent analysis. Conditional on firm i winning a consumer and its product being valued at v_i, its expected profit margin is

$$\bar{p}(v_i) - c = v_i - \int_{\underline{v}}^{v_i} \max\{0, x - c\} d\frac{G(x|v_i)}{G(v_i|v_i)} - c = \frac{\int_c^{v_i} G(x|v_i)dx}{G(v_i|v_i)} \ , \tag{17}$$

where the first equality used (14) and the second one is from integration by parts. Then industry profit under perfect price discrimination is

$$\Pi_D = n \int_c^{\overline{v}} [\bar{p}(v_i) - c] G(v_i|v_i) dF(v_i) = n \int_c^{\overline{v}} [G(v|\overline{v}) - G(v|v)F(v)] dv \ , \tag{18}$$

where the second equality is from using (17) and integration by parts. In the IID case, $G(v|\overline{v}) = G(v|v) = F(v)^{n-1}$, so it simplifies to

$$\Pi_D \overset{\text{IID}}{=} \int_c^{\overline{v}} \frac{1 - F(v)}{f(v)} dF(v)^n \ . \tag{19}$$

We will now show that when the market is only partially covered, competitive personalized pricing can raise profit and lower aggregate consumer surplus. To understand why, it is useful to first investigate why the simple proof in Proposition 1 breaks down with partial coverage. Under Assumption 1, we still have that

$$\Pi_D = n \int_0^{\infty} [1 - H_c(x)] dx \le n \frac{[1 - H_c(0)]^2}{h_c(0)} \ , \tag{20}$$

but now the last term is greater than

$$\Pi_U = n \frac{[1 - H_p(0)]^2}{h_p(0)} \ ,$$

because $p > c$ and both $1 - H_z(0)$ and $\frac{1 - H_z(0)}{h_z(0)}$ decrease in z. (In the full-coverage case, $c < p \le \underline{v}$ and so $H_c = H_p = H$.) This observation also suggests that if $1 - H_z(x)$ is log-linear in x, then the inequality in (20) binds and so we have $\Pi_D > \Pi_U$ whenever the market is not fully covered. That is indeed what we show in the following example.

An exponential distribution example. Before pursuing some general analytic results, it is illuminating to first consider an IID example with an exponential distribution $F(v) = 1 - e^{-(v-\underline{v})}$. We showed earlier that with full market coverage under uniform pricing (which requires $c \leq \underline{v} - 1$), personalized pricing has *no* impact on profit or consumer surplus. We now show that with partial coverage (meaning that $c > \underline{v} - 1$), personalized pricing always *benefits* firms but *harms* consumers. This example suggests that, at least for some distributions close to the exponential one, moving away from full market coverage can qualitatively change the impact of personalized pricing.

A convenient feature of this exponential example is that the equilibrium uniform price equals $1 + c$ irrespective of whether or not the market is fully covered. Under uniform pricing, a fraction $F(1 + c)^n$ of consumers are excluded from the market, so industry profit is $\Pi_U = 1 - F(1 + c)^n$. Under personalized pricing, using (19) and the fact that $1 - F(v) = f(v)$ in this exponential example, we immediately have $\Pi_D = 1 - F(c)^n$. Therefore, under the condition for partial coverage ($1 + c > \underline{v}$), $\Pi_D > \Pi_U$, i.e., personalized pricing *boosts* profit.

The impact on consumer surplus is still unclear, because as noted earlier personalized pricing also increases total welfare. Consumer surplus under uniform pricing and personalized pricing are respectively

$$V_U = \int_{1+c}^{\infty} (v - c) dF(v)^n - \Pi_U \text{ and } V_D = \int_{c}^{\infty} (v - c) dF(v)^n - \Pi_D \ ,$$

where the integral term in each expression is the total welfare in each regime. The former is greater than the latter if and only if

$$F(1 + c)^n - F(c)^n > \int_{c}^{1+c} (v - c) dF(v)^n \ ,$$

which is true as $v - c < 1$ for $v \in (c, 1 + c)$. Therefore, personalized pricing boosts profit so much that consumers always *suffer* from it.

One way to understand the above results is as follows. Notice that under personalized pricing total demand is $1 - F(c)^n$, and so the *average* price that consumers pay is $1 + c$, which is exactly equal to the uniform price. Personalized pricing therefore raises profit, because it expands the size of the market. At the same time, this market expansion is from consumers whose highest valuation is between c and $1 + c$—and since this is below the average price, personalized pricing lowers aggregate consumer surplus. This insight can be generalized. If personalized pricing weakly raises or only slightly decreases the average price paid by consumers, it should harm consumers overall even if it expands the demand.

The production cost and market coverage. Given the full-coverage result in Proposition 1, it is clear that for a more general (regular) distribution, the impact of personalized pricing can only be reversed when the market is sufficiently far away from being fully covered. As we saw earlier, by changing the marginal cost c we can change the degree of market coverage. In particular, when c is sufficiently close to the valuation upper bound $\overline{v}$, most consumers are excluded from the market. In that case we can show that the impact of personalized pricing is completely different from the full coverage case.[23]

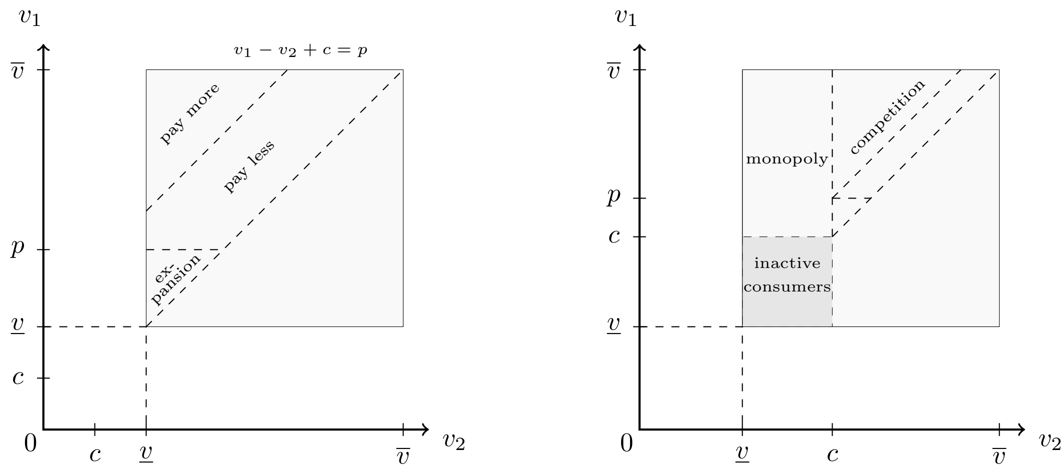
Proposition 2. *Suppose $\overline{v} < \infty$ and $f(\overline{v}) > 0$. For any given $n \geq 2$, there exists $\hat{c} < \overline{v}$ such that when $c > \hat{c}$ personalized pricing benefits firms and harms consumers compared to uniform pricing.*

(One may wonder whether this result is immediate from (13) and (16). From those expressions, it is clear that if c and p are sufficiently close to each other but both are bounded away from $\overline{v}$, it must be the case that $V_U > V_D$ given $v_{n:n}$ is greater than $v_{n-1:n}$ in the sense of first-order stochastic dominance. This ranking, however, is no longer obvious when both c and p approach $\overline{v}$ but c is always smaller than p.)

To understand this result—and more generally how varying c affects the impact of personalized pricing—we refer to the following graphs which illustrate the duopoly case. (The graphs also work for the $n > 2$ case if we interpret v_2 as $\max_{j \geq 2}\{v_j\}$.)

Consider first the case depicted in Figure 1a, where $c < \underline{v}$ (so the market is fully covered under personalized pricing) but $p > \underline{v}$ (so the market is only partially covered under uniform pricing). Let us focus on the consumers in the upper triangle area who prefer firm 1's product. Under personalized pricing, these consumers buy from firm 1 and pay $v_1 - v_2 + c$. Compared to the regime of uniform pricing with price p, those consumers in the northwest corner with $v_1 - v_2 + c > p$ pay more, those with $v_1 - v_2 + c < p$ and $v_1 > p$ pay less, and those in the "expansion" region, who were excluded from the

[23] As noted earlier, we could also normalize marginal cost and vary consumers' outside option. When the outside option is sufficiently close to $\overline{v}$ personalized pricing benefits firms and harms consumers. As indicated by some examples below, the assumption $\overline{v} < \infty$ in Proposition 2 does not appear to be crucial, but we have not been able to extend our proof to the case with an unbounded support.



(a) The case of $c < \underline{v}$. As c increases, consumers in the expansion region benefit, but others suffer.

(b) The case of $c > \underline{v}$. As c increases, the monopoly region becomes more important relative to the competition region.

Figure 1: The impact of personalized pricing with partial market coverage

market under uniform pricing, now buy. If c is small and p is close to $\underline{v}$, the situation is close to full market coverage. As c increases, the "expansion" triangle grows, which is a positive effect for consumers; at the same time the line of $v_1 - v_2 + c = p$ moves downward, which is a negative effect for consumers. (The latter effect is because the profit margin $p - c$ decreases in c, which can be seen from (7). This reflects the fact that under personalized pricing firms fully pass cost increases through to consumers, whereas under uniform pricing firms share some of the burden.) Since the consumers in the expansion region have low valuations, the positive effect is relatively small, and so it becomes possible for personalized pricing to hurt consumers.

When c exceeds $\underline{v}$, as depicted in Figure 1b, a new effect emerges. The "monopoly" region contains consumers who value product 2 below, but product 1 above, the cost. Firm 1 acts as a true monopolist over these consumers and extracts all their surplus under personalized pricing. For those consumers in the "competition" region who value both products above the cost, the situation is the same as in Figure 1a. As c increases, both the monopoly and competition regions shrink, but the monopoly region becomes proportionally more important. When c is close to $\overline{v}$, the monopoly region dominates, so the impact of personalized pricing becomes qualitatively the same as in the monopoly case, as proved in Proposition 2.[24] (Notice also that since both regions shrink as c increases, the size of the impact goes to zero in the limit.)

Following this discussion when c is large, we have an even stronger result when c is sufficiently large. Using the results in the proof of Proposition 2, we can show the following result:

Corollary 1. *Suppose $\overline{v} < \infty$ and $f(\overline{v}) > 0$. For any given $n \geq 2$, there exists $\hat{c} < \overline{v}$ such that for $c > \hat{c}$ each firm is better off and consumers are worse off under competitive personalized pricing than under monopoly uniform pricing.*

This is again because when c is large, each firm acts almost as a monopolist.

Proposition 2 and the above discussion suggests a possible cutoff result, whereby personalized pricing benefits firms if and only if c exceeds a threshold c', and harms consumers if and only if c exceeds another threshold c''. (Where $c'' > c'$ because personalized pricing raises total welfare.) Although it appears hard to formally prove such a cutoff result, numerical simulations suggest it is true. In particular, Figure 2 plots the impacts of personalized pricing on industry profit $(\Pi_D - \Pi_U)$ and on consumer surplus $(V_D - V_U)$ for four common distributions (all in the IID case) and for different values of c.

Figure 2a considers the exponential case with $F(v) = 1 - e^{-(v-1)}$ on $[1, \infty)$. At $c = 0$ the market is (just) covered and so the impact is zero, but for higher values of c the market is only partly covered, so as explained before personalized pricing benefits firms and harms consumers. Figures 2b and Figure 2c consider, respectively, the extreme value distribution with $F(v) = e^{-e^{-(v-2)}}$ (which leads to the logit model), and the normal distribution with mean 2 and variance 1 (which leads to the probit model). In both cases, for low values of c (when coverage is high) personalized pricing benefits consumers and harms firms as in the full-coverage case, for high values of c (when coverage is low) personalized pricing has the opposite impact, while for intermediate c both consumers and firms benefit from personalized pricing. Finally, the same pattern is also observed in Figure 2d, which considers the case where valuations are uniformly distributed on $[1, 2]$. In this example, when $c \leq 1/2$ the market is covered, so we

[24] Alternatively, when c is close to $\overline{v}$, conditional on a consumer valuing one firm's product more than c, it is very unlikely that the consumer values any other product more than c. Therefore when c is high, each firm is essentially a monopolist competing only against the outside option.

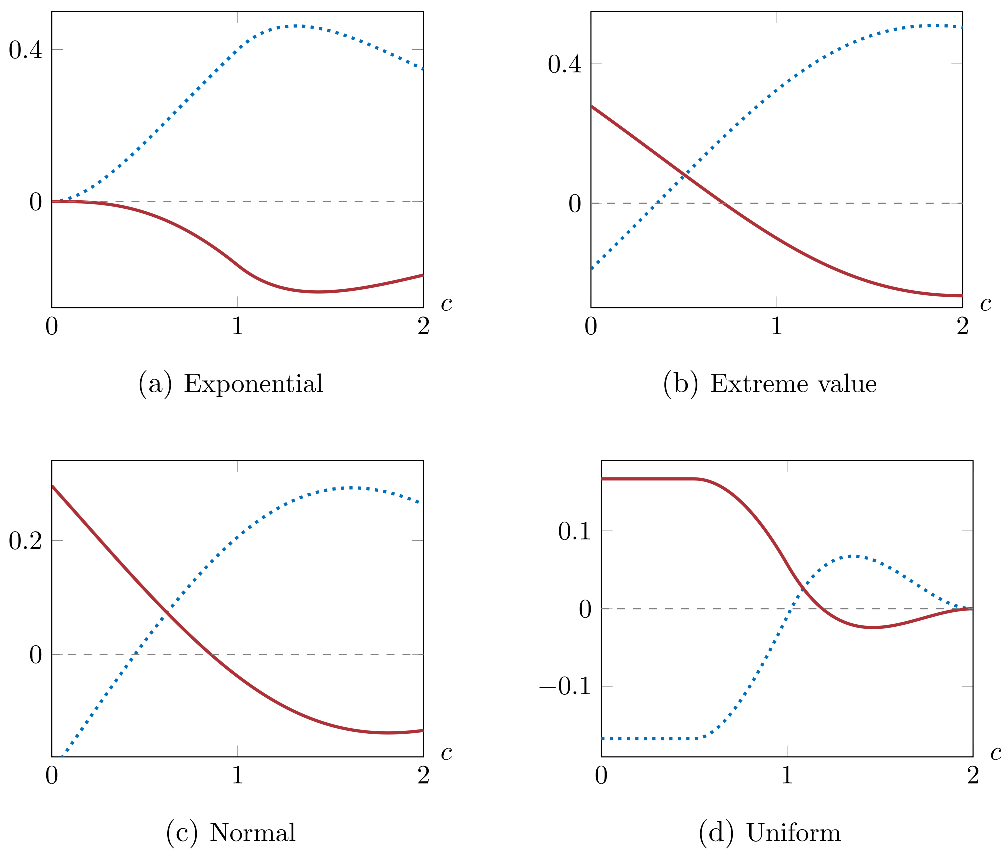


Figure 2: The impact of personalized pricing when $n = 2$, for different values of c.
(The dotted and the solid lines represent, respectively, industry profit and consumer surplus.)

know from earlier that personalized pricing harms firms and benefits consumers. When $c > 1/2$ the market is only partially covered, and personalized pricing benefits firms whenever c is above about 1.02, and harms consumers whenever c is above about 1.19.

The number of firms and market coverage. Another parameter which influences market coverage is the number of firms. When n is small, the impact of personalized pricing should be similar to under monopoly; when n is large, the best match should be relatively high, and so the impact of personalized pricing should be similar to the full-coverage case.

We first report an analytical result for the case where n is large, using an approximation argument similar to the one for the case of large c. However, the approximation when n is large is technically more difficult. We rely on the approximation results developed in Gabaix et al. (2016) which use extreme value theory, but this technique works only in the IID case. Let

$$\gamma = \lim_{v \to \overline{v}} \frac{d}{dv}\left(\frac{1 - F(v)}{f(v)}\right) \tag{21}$$

denote the tail index of the valuation distribution of each product. When f is log-concave, we must have $\gamma \in [-1, 0]$.[25]

Proposition 3. *Consider the IID case without full market coverage for any n. Suppose f is strictly positive and finite everywhere on its support and has a tail index $\gamma \in (0, 1)$. Then there exists $\hat{n}$ such that when $n > \hat{n}$ personalized pricing harms firms and benefits consumers.*

Given the opposite is true in the monopoly case, this again suggests the possibility of a cutoff result in terms of n. Since an analytic result seems hard to obtain, we instead report some numerical examples in Figure 3 below (the IID case with $c = 0$). Figure 3a is for the exponential distribution, and confirms our earlier analytic result that personalized pricing always benefits firms and harms consumers. Figure 3b shows that for the standard extreme value distribution, personalized pricing benefits firms if and only if $n < 10$, and harms consumers if and only if $n < 7$. So the predictions from Thisse and Vives (1988) fail for a relatively large range of n. A qualitatively similar pattern emerges in Figure 3c for the standard normal distribution. Figure 3d considers the uniform distribution with support $[0, 1]$, where the impact of personalized pricing is reversed once we move beyond monopoly. (However, simulations for higher values of c show that the impact of personalized pricing can be similar to in monopoly for some values of $n > 1$.)

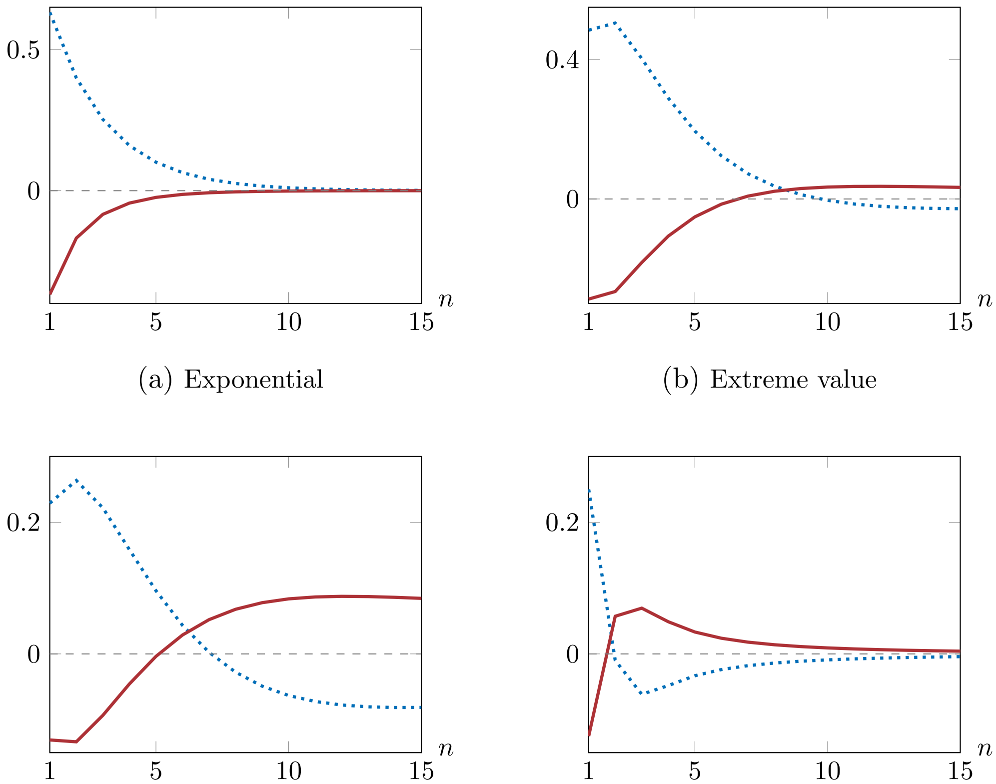
4.2 The long-run impact in a free-entry market

In the long-run, personalized pricing may also influence the market structure. To investigate this we now consider a free-entry game, where firms first decide whether or

[25] Given f is log-concave, $1 - F$ is log-concave, so $\gamma \leq 0$. To see $\gamma \geq -1$, notice that

$$\frac{d}{dv}\left(\frac{1 - F(v)}{f(v)}\right) = -1 - \frac{1 - F(v)}{f(v)} \frac{f'(v)}{f(v)} .$$

If $\lim_{v \to \overline{v}} f'(v) \leq 0$, the claim is obvious. If $\lim_{v \to \overline{v}} f'(v) > 0$, then we must have $\overline{v} < \infty$ and $f(\overline{v}) > 0$, in which case $\frac{1 - F(\overline{v})}{f(\overline{v})} = 0$ and given the log-concavity of f we also have $\frac{f'(\overline{v})}{f(\overline{v})} < \infty$. Then $\gamma = -1$. (Without the log-concavity of f, $\frac{f'(\overline{v})}{f(\overline{v})}$ can be ∞, in which case γ can be less than -1.)



(a) Exponential

(b) Extreme value

(c) Normal

(d) Uniform

Figure 3: The impact of personalized pricing when $c = 0$, for different values of n.
(The dotted and the solid lines represent, respectively, industry profit and consumer surplus.)

not to pay a fixed entry cost, and then after entering they compete in prices. The free-entry equilibrium is determined by the usual zero-profit condition. (As in the literature, we implicitly assume a sequential entry game to avoid coordination problems, and we ignore integer constraints on n.)

Let us first study the case with personalized pricing. Due to Bertrand competition, the profit on each consumer is simply the difference between her best and second-best product valuations, adjusted for the marginal cost. Hence, with n firms in the market, each firm's profit can be expressed as

$$\frac{1}{n}\Pi_D = \frac{1}{n}\mathbb{E}[\max\{c, v_{n:n}\} - \max\{c, v_{n-1:n}\}] \ . \tag{22}$$

On the other hand, the increase in match efficiency when the number of firms goes from $n-1$ to n is

$$\mathbb{E}[\max\{c, v_{n:n}\}] - \mathbb{E}[\max\{c, \hat{v}_{n-1:n-1}\}] \ , \tag{23}$$

where $\hat{v}_{n-1:n-1}$ denotes the best match among the original $n-1$ products. (We use $\hat{v}_i$ to denote the valuation for product $i \leq n-1$ when there are only $n-1$ firms in the market.) To determine whether there is too much or too little entry relative to the social optimum, it then suffices to compare (22) and (23).

Assumption 2. *Entry of a new firm does not affect consumers' valuations for existing products. That is, the distribution of* $(\hat{v}_1, ..., \hat{v}_{n-1})$ *when there are* $n-1$ *firms in the market is the marginal distribution of* $(v_1, ..., v_{n-1})$ *when there are* n *firms in the market.*

Under the above assumption (which is clearly true, e.g., in the IID case),[26] it turns out that (22) and (23) are actually equal to each other. Hence one can show that:

Lemma 3. *Under Assumption 2, the free-entry equilibrium under personalized pricing is unique and it is also socially optimal.*

The intuition for this result is straightforward. Suppose $n-1$ firms are already in the market, and consider the entry of an n^{th} firm. Amongst consumers for whom $v_n \leq \max_{j \leq n-1}\{c, v_j\}$, this additional firm creates no social surplus and earns zero profit. However, amongst consumers with $v_n > \max_{j \leq n-1}\{c, v_j\}$, this new firm raises total surplus by $v_n - \max_{j \leq n-1}\{c, v_j\}$, and fully extracts it via Bertrand competition. As a result, the incentives of the social planner and this new firm are perfectly aligned.[27]

Now consider the case with uniform pricing. Let n^* denote the socially optimal number of firms. A simple corollary of Lemma 3 is the following:

Corollary 2. *Suppose Assumption 2 holds and each firm's profit under uniform pricing decreases in* n*.*
(i) Entry under uniform pricing is excessive if $\Pi_U > \Pi_D$ *at* $n = n^*$*, but insufficient if* $\Pi_U < \Pi_D$ *at* $n = n^*$*.*
(ii) Uniform pricing therefore leads to excessive entry if the market is fully covered and Assumption 1 holds.

[26] Assumption 2 will fail if the entry of a new product induces the existing firms to reposition their products or if consumers have consideration-set dependent preferences.

[27] Note that Assumption 2 fails in the Salop circle model, because entry of a new firm causes existing ones to relocate, and this changes consumer valuations for their products. (Contrary to Lemma 3, Section 3.5 of Stole (2007) shows that entry is socially excessive in the Salop circle model under perfect price discrimination.) However Assumption 2 can hold in other spatial models, such as in Chen and Riordan (2007) where entry of a new firm does not lead to repositioning by existing firms.

To understand part (i) of the corollary, note that we have just shown that the increase in match efficiency due to an extra firm entering the market is exactly equal to that firm's profit under personalized pricing. Hence, with uniform pricing, entry is either excessive or insufficient depending upon whether Π_U is respectively above or below Π_D.[28] To understand part (ii), recall that Proposition 1 showed that with full coverage and Assumption 1 it is always the case that $\Pi_U > \Pi_D$. Anderson, de Palma, and Nesterov (1995) and Tan and Zhou (2021) also prove this excessive entry result in the IID case; our proof is much simpler than theirs, and our result is more general since it also potentially allows for correlated valuations.

Finally, another simple but important consequence of Lemma 3 is the following:

Proposition 4. *Compared to uniform pricing, personalized pricing (weakly) benefits consumers in the long-run.*

In the long-run firms earn zero profit (after accounting for the fixed entry cost) irrespective of the pricing regime. Therefore since total welfare is maximized under personalized pricing, so is aggregate consumer surplus.

5 Consumer Privacy Choice

We now consider an extended game in which consumers decide whether or not to share their data. When a consumer shares her data, she incurs a privacy cost τ; this is drawn independently of her product valuations, according to a CDF $T(\tau)$ with support $[\underline{\tau}, \overline{\tau}]$, where $\underline{\tau} \geq 0$ and $\overline{\tau}$ can be infinity.[29] We assume that if a consumer chooses to share her data, all firms learn her vector of valuations $\mathbf{v} = (v_1, \ldots, v_n)$ and offer personalized prices. If a consumer chooses not to share her data, she remains anonymous and firms

[28]The assumption that each firm's profit under uniform pricing decreases in n ensures the uniqueness of the free-entry equilibrium. It must hold if the uniform price decreases in n (which, as shown in Zhou (2017), holds, e.g., in the IID case with a log-concave f). To see this, let $\pi_n(p_i, p)$ be firm i's profit when it offers a uniform price p_i and the other $n-1$ firms offer a price p. Let p_n be the equilibrium uniform price. Then $\pi_n(p_n, p_n) < \pi_{n-1}(p_n, p_{n-1}) \leq \pi_{n-1}(p_{n-1}, p_{n-1})$, where the first inequality is because a firm's profit must increase when the number of competitors drops and they further set higher prices $p_{n-1} > p_n$, and the second inequality is simply from the no-deviation equilibrium condition when there are $n-1$ firms in the market.

[29]This privacy cost reflects, e.g., a consumer's intrinsic concern about data security.

have to offer her a common “list” price. We assume that a firm’s list price is publicly available and therefore caps its personalized prices.

The timing is as follows. At the first stage, each consumer learns her privacy cost τ. Then, *before* learning her product valuations $\mathbf{v}$, each consumer simultaneously decides whether or not to share her data. At the second stage, each consumer’s valuations are realized. Firms observe a consumer’s $\mathbf{v}$ if and only if she shared her data. Firms simultaneously choose a uniform price for every consumer who hid her data, and a (weakly lower) personalized price for each consumer who shared her data. Consumers then decide which product (if any) to buy.

Discussion. We are assuming that each firm’s list price is public and caps its personalized prices. As an example, the list price could be freely available on a firm’s website, which consumers can consult before choosing whether or not to accept their personalized price. One can therefore interpret personalized prices as targeted discounts off the list price. We are also assuming that consumers choose whether or not to share their data before knowing their product valuations. This captures the idea that consumers often need to accept or reject cookies on non-merchant websites, and have limited information about how and when in the future those cookies will be used.[30]

5.1 Price competition in a mixed market

Let α denote the fraction of consumers who choose to hide their data and thus remain anonymous. (Recalling the previous section, $\alpha = 0$ corresponds to personalized pricing and $\alpha = 1$ corresponds to uniform pricing.) In this section we consider a general $\alpha \in [0, 1]$ and solve for a symmetric pricing equilibrium where each firm uses the same list price $p(\alpha)$; when there is no confusion we simply denote this list price by p. (In the next section we then endogenize α via consumers’ privacy choices.)

Competition for anonymous consumers is the same as in the uniform pricing regime studied in Section 3.1. In particular, if firm i unilaterally deviates to a list price $p_i > c$, its expected profit from an anonymous consumer is

$$\pi_a(p_i, p) \equiv (p_i - c)[1 - H_p(p_i - p)] \ . \tag{24}$$

[30] As discussed in the introduction, our approach contrasts with many papers on consumer privacy choice, which often assume that consumers know their product valuations when deciding whether or not to accept cookies.

Competition for consumers who share their data is also the same as in Section 3.2, except that now each firm's personalized prices are capped by their list price. As a result, firm i's profit margin when it wins a consumer is

$$p(v_i, \mathbf{v}_{-i}) - c = \min\{v_i - \max_{j \neq i}\{c, v_j\}, p_i - c\} \ .$$

Recalling the definition $x_c = v_i - \max_{j \neq i}\{c, v_j\}$ from earlier, firm i's expected profit from a sharing consumer can then be written as

$$\pi_s(p_i) \equiv \int_0^{p_i - c} x dH_c(x) + (p_i - c)[1 - H_c(p_i - c)] = \int_0^{p_i - c} [1 - H_c(x)]dx \ , \qquad (25)$$

where the second equality follows from integration by parts. Notice that $\pi_s(p_i)$ is increasing in firm i's list price—intuitively, an increase in p_i enables firm i to charge a higher personalized price to each consumer for whom the list price was binding.[31] Hence firm i faces a trade-off when increasing its list price: it loses demand from anonymous consumers, but earns more profit from consumers who shared their data.

Using the above, firm i's deviation profit from charging a list price p_i is

$$\alpha \pi_a(p_i, p) + (1 - \alpha)\pi_s(p_i) \ . \qquad (26)$$

When $\alpha = 0$—such that all consumers share their data—the fact that $\pi_s(p_i)$ is strictly increasing implies that all firms set their list price (weakly above) $\overline{v}$; the list price never binds, and firms use personalized pricing as in Section 3.2. To ensure the existence and uniqueness of a symmetric equilibrium when $\alpha > 0$, we make the following assumption:

Assumption 3. *The profit expression (26) is quasi-concave in p_i for any $\alpha > 0$ and $p \geq c$, and both $\frac{1 - H_z(0)}{h_z(0)}$ and $\frac{1 - H_c(z - c)}{h_z(0)}$ are non-increasing in z.*

We show in the Appendix that Assumption 3 holds, for example, in the leading case of IID valuations with a log-concave f.[32] We can then state the following result.

Lemma 4. *Under Assumption 3, for any $\alpha > 0$ the symmetric equilibrium list price $p(\alpha)$ uniquely solves*

$$p - c = \frac{1 - H_p(0)}{h_p(0)} + \frac{1 - \alpha}{\alpha} \frac{1 - H_c(p - c)}{h_p(0)} \ , \qquad (27)$$

and it decreases in α.

[31] Moreover $\pi_s(p_i)$ is also concave in p_i. Intuitively, when p_i is larger, the list price binds for fewer consumers, and so an increase in p_i has less impact on firm i's profit.

[32] Note that although $\pi_s(p_i)$ is concave in p_i (see footnote 31 above), and $\pi_a(p_i, p)$ is log-concave in p_i under Assumption 1 from earlier, a linear combination of them need not necessarily be quasi-concave.

An uninteresting case arises when the uniform price in Section 3 exceeds the highest possible personalized price. In this case it is easy to show that the equilibrium list price is exactly equal to the uniform price—meaning that it is independent of α and never caps the personalized prices.[33] This situation, however, is ruled out if Lemma 2 holds. In the following, unless stated otherwise, we do not consider this uninteresting case.

The property that the list price decreases in α plays an important role in the remainder of this section.[34] It captures an interesting *externality* across consumers: when more consumers choose to share their data, firms optimally raise their list prices in order to allow more price personalization, which harms anonymous consumers as well as sharing consumers for whom the list price binds.

5.2 Equilibrium privacy choice

We now turn to consumers' privacy choice. Suppose a consumer expects a fraction α of consumers to be anonymous. If she chooses to hide her data and remain anonymous, she expects to face the list price in (27) and so her expected surplus is

$$V_a(\alpha) \equiv \mathbb{E}[\max\{v_{n:n} - p, 0\}] \ . \tag{28}$$

If instead the consumer chooses to share her data, her expected surplus is

$$V_s(\alpha) \equiv \mathbb{E}[\max\{v_{n:n} - p, v_{n-1:n} - c, 0\}] \ , \tag{29}$$

because when she buys, if the list price binds she obtains surplus $v_{n:n} - p$, whereas if the list price does not bind we know from Section 3.2 that her surplus under price discrimination is $v_{n-1:n} - c$. It is easy to see that both V_a and V_s increase in α because p decreases in α, and $V_s > V_a$ given $p > c$. (Notice that if $\alpha = 0$ then $p = \overline{v}$, and so $V_a(0) = 0$ while $V_s(0)$ degenerates to the consumer surplus V_D in the regime of personalized pricing in Section 3.2.)

[33] To see why, notice that $1 - H_c(p-c) = \int_p^{\overline{v}} G(v-p+c|v)dF(v) = 0$ if $p \geq c + \overline{v} - \underline{v}$. Therefore, if the uniform price in Section 3.1, which solves $p - c = \phi(p)$, exceeds the highest possible personalized price, it also solves (27) and so is the equilibrium list price.

[34] This property holds as long as the list price equilibrium is locally stable (i.e., each firm's best-response function has a slope less than one at the equilibrium list price). Assumption 3 is sufficient but not necessary for the equilibrium to be locally stable.

When we deal with some specific examples, it will be useful to write V_a and V_s more explicitly as follows:

$$V_a(\alpha) = \int_p^{\overline{v}} (v - p) dF_{(n)}(v) \ , \tag{30}$$

and

$$V_s(\alpha) = \int_c^{\overline{v}} F_{n|n-1}(x + p - c|x)(x - c) dF_{(n-1)}(x) + \int_p^{\overline{v}} F_{n-1|n}(v - p + c|v)(v - p) dF_{(n)}(v) \ , \tag{31}$$

where $F_{n|n-1}(v|x)$ denotes the CDF of the highest valuation conditional on the second highest being equal to x, and $F_{n-1|n}(x|v)$ denotes the CDF of the second highest valuation conditional on the highest being equal to v.[35] Hence:

$$F_{n|n-1}(v|x) = \frac{F(v|x) - F(x|x)}{1 - F(x|x)}; \ F_{n-1|n}(x|v) = \frac{G(x|v)}{G(v|v)} \ ,$$

where $F(v|x)$ is the CDF of v_j conditional on $v_i = x$.

Comparing (28) and (29), we immediately have the following observations:

Lemma 5. *Define $\Delta(\alpha) \equiv V_s(\alpha) - V_a(\alpha)$ as the consumption benefit of sharing data. Then (i) $\Delta(\alpha) = 0$ in the monopoly case, and (ii) $\Delta(\alpha) > 0$ and decreases in α for any $n \geq 2$ under Assumption 3.*

To understand the monopoly case, note that if a consumer's valuation exceeds p then she gets the same surplus $v - p$ regardless of whether or not she shares—because when she shares her data the price she pays is capped by p. Similarly, if the consumer's valuation is less than p, she also gets zero surplus regardless of whether or not she shared her data—when she is anonymous she does not buy, whereas when she shares she buys but her surplus is fully extracted by the monopolist.

In the competition case, on the other hand, there is always a chance that the consumer will benefit from personalized pricing (i.e., when $v_{n-1:n}$ is sufficiently close to $v_{n:n}$). Moreover, when p is smaller (which is the case when α is higher), it is less likely that $v_{n-1:n} - c$ exceeds $v_{n:n} - p$, and so $\Delta(\alpha)$ is smaller.

[35] The first term in V^s is the expected surplus when the list price does not bind (conditional on $x \geq c$, $x - c$ is the relevant surplus if $v - p < x - c$); the second term is the expected surplus when the list price binds (conditional on $v \geq p$, $v - p$ is the relevant surplus if $x - c < v - p$).

Any equilibrium in the privacy choice game must follow a cut-off rule: if consumers expect a fraction α of them to be anonymous, then all consumers with privacy cost $\tau > \Delta(\alpha)$ choose to be anonymous, while all the others share. As a result, α forms a rational expectations equilibrium if and only if

$$\alpha = 1 - T(\Delta(\alpha)) \ . \tag{32}$$

We then obtain the following results:

Proposition 5. *(i) The privacy choice game must have an equilibrium.*
(ii) $\alpha = 0$ *is an equilibrium if and only if* $\Delta(0) \geq \overline{\tau}$, $\alpha = 1$ *is an equilibrium if and only if* $\Delta(1) \leq \underline{\tau}$, *and any equilibrium of the privacy choice game must be interior with* $\alpha \in (0,1)$ *if and only if* $\underline{\tau} < \Delta(1) < \Delta(0) < \overline{\tau}$.

Existence of equilibrium simply follows from Tarski's fixed-point theorem since the right-hand side is an increasing function. The remainder of Proposition 5 then gives conditions for existence of either "corner" (i.e., $\alpha = 0, 1$) or interior equilibria, which we now discuss further.

Consider first the case of monopoly. As discussed earlier, in this case $\Delta(\alpha) = 0$, and given $\underline{\tau} \geq 0$ and the privacy cost distribution is continuous, the unique equilibrium has all consumers hiding their data and so remaining anonymous.

Consider next the case of oligopoly. One difference with monopoly is that, provided $\underline{\tau}$ is not too large, some consumers will share their data in equilibrium. Another difference with monopoly is that the privacy game can have multiple equilibria. To see why, recall from earlier that $\Delta(\alpha)$ is strictly positive and decreasing in α. If privacy costs are sufficiently homogeneous, such that $\Delta(1) \leq \underline{\tau} < \overline{\tau} \leq \Delta(0)$, then Proposition 5 says there are at least two corner equilibria. Even if privacy costs are very heterogeneous, it is easy to find examples with multiple interior equilibria.[36] Intuitively, multiplicity arises due to "complementarity" in consumers' privacy decisions: when more consumers are expected to remain anonymous, consumers anticipate a lower list price and hence a smaller benefit from sharing their data and getting personalized prices, which induces more of them to remain anonymous.

When the privacy choice game has multiple equilibria, they can be ranked according to consumer surplus. This is immediate from the following simple result:

[36] One can do this even when $\underline{\tau} = 0$ and $\overline{\tau} = \infty$, such that corner equilibria do not exist. Examples are available on request.

Lemma 6. *Suppose the privacy choice game has an equilibrium with $\alpha > 0$. Then each individual consumer prefers this equilibrium over any market situation with fewer anonymous consumers, regardless of whether that situation is an equilibrium or not.*

To understand the lemma, note that given an *equilibrium* α, the expected surplus of a consumer with privacy cost τ is $\max\{V_a(\alpha), V_s(\alpha) - \tau\}$. At the same time, for any situation with $\alpha' < \alpha$, a consumer with privacy cost τ obtains a surplus of at most $\max\{V_a(\alpha'), V_s(\alpha') - \tau\}$ (because if α' is not an equilibrium, some consumers make a suboptimal privacy choice). Therefore, since both V_a and V_s increase in α, it is clear that every consumer is better off in an equilibrium with α compared to any market situation with $\alpha' < \alpha$.[37] We then obtain the following corollary:

Corollary 3. *Consider two equilibria with α and $\alpha' < \alpha$. Each consumer strictly prefers the α equilibrium.*

Lemma 6 also has the following simple policy implication:

Corollary 4. *A privacy policy like GDPR, which enables consumers to costlessly hide their data, strictly benefits consumers if it induces some of them to hide their data.*

Absent a privacy policy firms have full access to consumers' data, which in terms of the above discussion is equivalent to $\alpha' = 0$. If all consumers respond to the policy by sharing their data then the policy has no effect. However, provided some consumers react by hiding their data, the policy benefits all consumers—because they all pay less, and those who hide their data avoid incurring the privacy cost.

5.3 Comparison with consumer-optimal privacy choice

We now compare the equilibrium privacy choice with the consumer-optimal choice. Suppose a social planner can choose α of the consumers with the highest τ to be anonymous, whilst forcing the remaining consumers to share their data. Then aggregate consumer welfare is

$$V(\alpha) = \alpha V_a(\alpha) + (1 - \alpha) V_s(\alpha) - P(\alpha) \ , \tag{33}$$

[37] This argument depends on α being an equilibrium: if neither α nor $\alpha' < \alpha$ constitute an equilibrium, some consumers may be worse off in the latter situation even though the list price is lower. This is because those who are forced to hide their data may suffer as $V_a(\alpha) < V_s(\alpha)$ for any given α.

where $P(\alpha) = \int_{\underline{\tau}}^{T^{-1}(1-\alpha)} \tau dT(\tau)$ is the total privacy cost paid by consumers and it is decreasing in α.

Proposition 6. *In any interior privacy choice equilibrium, there are too many sharing consumers relative to the consumer optimum.*

Proof. Let $\hat{\alpha}$ be an interior equilibrium of the privacy choice game. According to Lemma 6 each consumer gets strictly higher payoff in the $\hat{\alpha}$ equilibrium compared to any market situation with $\alpha' < \hat{\alpha}$. We now show that aggregate consumer surplus is strictly increasing in α around $\hat{\alpha}$, which is enough to prove the result. Note that using $P'(\alpha) = -T^{-1}(1-\alpha)$, we have

$$V'(\alpha) = \alpha V_a'(\alpha) + (1-\alpha)V_s'(\alpha) - \Delta(\alpha) + T^{-1}(1-\alpha) \ .$$

Since $\hat{\alpha}$ is an interior equilibrium we must have $\Delta(\hat{\alpha}) = T^{-1}(1-\hat{\alpha})$, and so

$$V'(\hat{\alpha}) = \hat{\alpha} V_a'(\hat{\alpha}) + (1-\hat{\alpha})V_s'(\hat{\alpha}) > 0$$

given both V_a and V_s increase in α. Hence the consumer-optimal privacy choice must have $\alpha > \hat{\alpha}$. □

The intuition for this result is straightforward: as pointed out before, when more consumers share their data they impose a negative externality on other consumers by inducing a higher list price. (More precisely, both the anonymous consumers and the sharing consumers who end up paying the list price are adversely affected.) However individual consumers ignore this externality when making their privacy choice.[38]

5.4 The impact of competition

We have already seen that under monopoly all consumers choose to remain anonymous, and hence equilibrium consumer surplus is

$$V_m = \int_{p_m}^{\overline{v}} (v - p_m) dF(v) = \int_{p_m}^{\overline{v}} [1 - F(v)] dv \ , \tag{34}$$

[38] Proposition 6 focuses on interior equilibria. Over-sharing may not arise with corner equilibria. For example, if the privacy choice game has an equilibrium with $\alpha = 1$, Lemma 6 implies that the consumer-optimal outcome is also $\alpha = 1$. Similarly, if the *unique* equilibrium of the privacy game has $\alpha = 0$, it is possible that the consumer-optimal outcome is also $\alpha = 0$. This is because when some consumers are forced to hide their data, the reduction in their surplus (from not being able to get personalized prices) might outweigh the gain in other consumers' surplus due to a lower list price.

where p_m is the standard monopoly price. We now show that competition can—at least over some range of $n \geq 2$—lower consumer surplus. This is because with competition $\Delta(\alpha) > 0$, so if $\underline{\tau}$ is not too large some consumers start sharing their data, which due to negative externalities can harm consumers overall. The following result reports a sufficient condition for this to happen:

Proposition 7. *There exist privacy-cost distributions such that competition harms consumers relative to monopoly, provided that*

$$V_s(0) - V_m < V_s(1) - V_a(1) \ . \tag{35}$$

This condition must hold if $V_D \leq V_m$ for $n \geq 2$, which is true if c is sufficiently high and if $\overline{v} < \infty$ and $f(\overline{v}) > 0$.

Proof. With competition, if $\overline{\tau} < \Delta(1)$, the unique privacy choice equilibrium is $\alpha = 0$, i.e., all consumers choose to share their data, and equilibrium consumer surplus is $V_s(0) - \mathbb{E}[\tau]$. This is worse than under monopoly if $V_s(0) < V_m + \mathbb{E}[\tau]$. Therefore, a sufficient condition for competition to harm consumers is

$$\overline{\tau} < \Delta(1) < V_s(0) < V_m + \mathbb{E}[\tau] \ ,$$

where $\Delta(1) < V_s(0)$ is always true since $V_a(0) = 0$ and so $V_s(0) = \Delta(0)$. (Recall that $\Delta(\alpha)$ is decreasing.) Other things equal, this condition is most likely to hold when $T(\cdot)$ is almost degenerate at, say, τ. In that case the above condition simplifies to

$$V_s(0) - V_m < \tau < \Delta(1) \ ,$$

from which (35) follows.

A sufficient condition for (35) to hold is that $V_D \leq V_m$, i.e., consumer surplus under competitive price discrimination is no greater than under uniform pricing in monopoly. (This is because the right-hand side of (35) is positive, while on the left-hand side $V_s(0)$ equals V_D.) From Corollary 1, we know this must be true if c is sufficiently large. $\square$

As we discussed before, for any given n, if c is sufficiently close to the valuation upper bound, each firm essentially acts as a monopolist over the consumers who value its product above the cost. That is why competitive personalized pricing results in lower consumer surplus than under monopoly uniform pricing.

Uniform example. To illustrate the above result, consider the IID uniform example with support $[0,1]$ and cost $0 \leq c < 1$. The monopoly price is $p_m = \frac{1+c}{2}$ and so $V_m = \frac{(1-c)^2}{8}$. Using $F_{(n-1)}(v) = v^n + n(1-v)v^{n-1}$ and $V_D = \int_c^1 [1 - F_{(n-1)}(v)]dv$, one can check that

$$V_D = c^n - c + \frac{n-1}{n+1}(1 - c^{n+1}) \ .$$

For example, $V_D = \frac{(1-c)^3}{3}$ when $n = 2$, which is less than V_m if and only if $c \geq \frac{5}{8}$; $V_D = \frac{(1-c)^3}{2}(1+c)$ when $n = 3$, which is less than V_m if and only if $c \geq \frac{\sqrt{3}}{2}$. For a higher c, $V_D \leq V_m$ (so competition harms consumers relative to monopoly) for a wider range of n.

Of course in the limit case with $n \to \infty$, if competition drives all (uniform or personalized) prices to marginal cost c (which is the case, e.g., when the valuation upper bound is finite), then consumers have no incentive to share their data, the same as in the monopoly case. In that case, consumers must be better off compared to the monopoly case. This is clear from the above uniform example.

6 Conclusion

This paper has investigated the impact of personalized pricing, a form of price discrimination which is becoming increasingly relevant in the digital economy. The paper delivers three main insights: (i) In a general oligopoly model, competitive personalized pricing intensifies competition, harms firms and benefits consumers under a log-concavity condition if the market is fully covered; however, the impact can be completely reversed in the (arguably more realistic) case without full market coverage. (ii) In a long-run free-entry equilibrium, personalized pricing leads to the socially optimal market structure and so favors consumers. (iii) In the presence of personalized pricing, when some consumers choose to share their data, this imposes a negative externality on other consumers who end up paying a higher list price. Due to this externality, too many consumers share their data relative to the consumer optimum, and more competition can harm consumers by inducing more data sharing.

There are many issues that deserve further exploration. For example, we have focused on the case where firms know a consumer's preferences for all products (if she shares her data). An alternative information structure, which might be more reasonable in some circumstances, is that each firm knows only a consumer's valuation for their own product. In this case, price competition between firms resembles a first-price auction. Therefore, in the IID case, the revenue equivalence theorem implies that the outcome (for both firms and consumers) is exactly the same as in our current setting. However more investigation is needed for the non-IID case. We have also assumed that firms have symmetric information on consumer preferences. However, in reality some firms may have more information than others about consumer preferences (consider, e.g., competition between Amazon and other smaller retailers); it would be interesting to investigate how such asymmetries affect competition and market performance.

Appendix: Omitted Proofs and Details

Primitive conditions for Assumption 1. We report some primitive conditions for Assumption 1. Define a piece of notation

$$G_2(x|y) \equiv \frac{\partial G(x|y)}{\partial y} \ . \tag{36}$$

Lemma 7. *(i) If the joint density $\tilde{f}$ is log-concave, then $1 - H_z(x)$ is log-concave in x. (ii) $\phi(z) = \frac{1-H_z(0)}{h_z(0)}$ is non-increasing in z if (a) $G_2(v|v) \geq 0$ and $f'(v) \geq 0$, or (b) if $\tilde{f}$ is log-concave and $\frac{G_2(v|v)}{G(v|v)}$ is non-increasing in v. (In particular, condition (b) holds in the IID case with a log-concave f.)*

Proof. (i) Note that

$$1 - H_z(x) = \int_{A_x} \tilde{f}(\mathbf{v}) d\mathbf{v} \ ,$$

where $A_x = \{\mathbf{v} : v_i - \max_{j \neq i}\{z, v_j\} > x\}$. To prove $1 - H_z(x)$ is log-concave in x, according to the Prékopa-Borell Theorem (see, e.g., Caplin and Nalebuff (1991)), it suffices to show that, for any $\lambda \in [0,1]$, we have

$$\lambda A_{x_0} + (1-\lambda) A_{x_1} \subset A_{\lambda x_0 + (1-\lambda) x_1} \ , \tag{37}$$

where the former is the Minkowski average of A_{x_0} and A_{x_1}. Let $\mathbf{v}^0 \in A_{x_0}$ and $\mathbf{v}^1 \in A_{x_1}$, i.e.,

$$v_i^0 > z + x_0 \ \text{ and } \ v_i^0 > v_j^0 + x_0 \ \text{ for any } j \neq i \ ,$$

and

$$v_i^1 > z + x_1 \ \text{ and } \ v_i^1 > v_j^1 + x_1 \ \text{ for any } j \neq i \ .$$

These immediately imply that

$$v_i^\lambda > z + \lambda x_0 + (1-\lambda) x_1 \ \text{ and } \ v_i^\lambda > v_j^\lambda + \lambda x_0 + (1-\lambda) x_1 \ \text{ for any } j \neq i \ ,$$

where $v_i^\lambda = \lambda v_i^0 + (1-\lambda) v_i^1$. Hence, we have $\mathbf{v}^\lambda \in A_{\lambda x_0 + (1-\lambda) x_1}$, and so (37) holds.

(ii) Recall that

$$\phi(z) = \frac{\int_z^{\overline{v}} G(v|v) dF(v)}{G(z|z) f(z) + \int_z^{\overline{v}} g(v|v) dF(v)} \ .$$

For $z \leq \underline{v}$, $\phi(z)$ is a constant and so is non-increasing. In the following, we focus on $z > \underline{v}$. Using $\frac{dG(v|v)}{dv} = g(v|v) + G_2(v|v)$, one can check that $\phi'(z) \leq 0$ if and only if

$$G(z|z) f(z) + \int_z^{\overline{v}} g(v|v) dF(v) + \left(\frac{f'(z)}{f(z)} + \frac{G_2(z|z)}{G(z|z)} \right) \int_z^{\overline{v}} G(v|v) dF(v) \geq 0 \ . \tag{38}$$

This must be true if condition (a) holds. To see condition (b), notice that the log-concavity of the joint density $\tilde{f}$ implies the log-concavity of the marginal density f and so $\frac{f'(z)}{f(z)} \geq \frac{f'(v)}{f(v)}$ for $v \geq z$. Therefore, a sufficient condition for (38) is

$$G(z|z)f(z) + \int_z^{\overline{v}} g(v|v)dF(v) + \int_z^{\overline{v}} G(v|v)f'(v)dv + \frac{G_2(z|z)}{G(z|z)} \int_z^{\overline{v}} G(v|v)dF(v) \geq 0 \ .$$

Applying integration by parts to the third term, we can rewrite the above condition as

$$f(\overline{v}) - \int_z^{\overline{v}} G_2(v|v)f(v)dv + \frac{G_2(z|z)}{G(z|z)} \int_z^{\overline{v}} G(v|v)dF(v) \geq 0 \ .$$

This holds if

$$\frac{G_2(z|z)}{G(z|z)} \geq \frac{G_2(v|v)}{G(v|v)}$$

for any $v \in [z, \overline{v}]$. This is true if $\frac{G_2(v|v)}{G(v|v)}$ is non-increasing in v. □

Proof of Lemma 1. Here we establish existence and uniqueness of the equilibrium price. (The rest of the lemma follows from arguments in the text.)

Clearly $p - c < \phi(p)$ when $p = c$. Since $\phi(p)$ is non-increasing due to Assumption 1, it suffices to show that $p - c > \phi(p)$ when $p = \overline{v}$. The latter must be true if $\overline{v} = \infty$, because $\phi(p)$ is non-increasing and thus finite as $p \to \infty$. It also holds if $\overline{v} < \infty$ and $f(\overline{v}) > 0$, because in that case $\phi(\overline{v}) = 0$. Finally, then, consider $\overline{v} < \infty$ and $f(\overline{v}) = 0$, in which case $f(v)$ must be decreasing for v sufficiently close to $\overline{v}$. Notice that $\phi(p) \leq \frac{\int_p^{\overline{v}} f(v)dv}{G(p|p)f(p)}$, which for p close to $\overline{v}$ is itself weakly less than $\frac{(\overline{v}-p)f(p)}{G(p|p)f(p)} = \frac{\overline{v}-p}{G(p|p)}$. This is clearly less than $\overline{v} - c$ when p is close to $\overline{v}$. □

Remark. To ease the exposition we assumed that the joint distribution of valuations has full support on $[\underline{v}, \overline{v}]^n$. However Lemma 1 also holds under the weaker condition that if $G(\hat{v}|\hat{v}) = 0$ for some $\hat{v}$, then $G(v|v) = 0$ for any $v \leq \hat{v}$. (This is true, e.g., when the joint distribution has a convex support.) Define $\underline{v}^* \equiv \max\{v : G(v|v) = 0\}$. Then $\underline{v}$ in the lemma can be replaced by $\underline{v}^*$.

Proof of Proposition 2. Suppose $c = \overline{v} - \varepsilon$ with $\varepsilon > 0$ close to zero. We first approximate the uniform price in equation (7), which in this case must be close to $\overline{v}$. One can check that $\phi(\overline{v}) = 0$ and $\phi'(\overline{v}) = -1$,[39] and so using a Taylor expansion we have $\phi(p) \approx \overline{v} - p$. The uniform price is therefore $p \approx \frac{1}{2}(\overline{v} + c) = \overline{v} - \frac{\varepsilon}{2}$, and the profit margin is $p - c \approx \frac{\varepsilon}{2}$. (Note that the approximated price is actually independent of the number of firms.) The equilibrium demand is $1 - F_{(n)}(p) \approx f_{(n)}(\overline{v})(\overline{v} - p) \approx f_{(n)}(\overline{v})\frac{\varepsilon}{2}$, where $f_{(n)}(\cdot)$ is the density of the highest order statistic. Hence industry profit under uniform pricing is

$$\Pi_U \approx f_{(n)}(\overline{v})\frac{\varepsilon^2}{4} \ .$$

On the other hand, under perfect price discrimination recall from (18) that industry profit is $\Pi_D(c) = n\int_c^{\overline{v}}[G(v|\overline{v}) - G(v|v)F(v)]dv$. One can check that $\Pi_D'(\overline{v}) = 0$ and

$$\Pi_D''(\overline{v}) = n\left[f(\overline{v}) + G_2(\overline{v}|\overline{v})\right] = nf(\overline{v}) \ ,$$

where $G_2(\cdot|\cdot)$ is defined in (36) and we have used the fact $G_2(\overline{v}|\overline{v}) = 0$ given $G(\overline{v}|y) = 1$ for any y. We can therefore approximate industry profit as

$$\Pi_D \approx nf(\overline{v})\frac{(\overline{v} - c)^2}{2} = nf(\overline{v})\frac{\varepsilon^2}{2} \ .$$

Notice that from

$$F_{(n)}(v) = \int_{[\underline{v},v]^n} \tilde{f}(\mathbf{v})d\mathbf{v} \ ,$$

we have

$$f_{(n)}(v) = n\int_{[\underline{v},v]^{n-1}} \tilde{f}(v, \mathbf{v}_{-i})d\mathbf{v}_{-i} \ ,$$

where we have used the symmetry of $\tilde{f}$. Therefore, $f_{(n)}(\overline{v}) = nf(\overline{v})$.[40] With this result, we can claim that the approximated Π_D is greater than the approximated Π_U, i.e., price discrimination improves profit when c is sufficiently large.

---

[39] Notice that

$$\phi'(p) = \frac{-h_p(0)G(p|p)f(p) - \int_p^{\overline{v}} G(v|v)dF(v)\frac{\partial h_p(0)}{\partial p}}{h_p(0)^2} \ ,$$

and so $\phi'(\overline{v}) = -1$ given $h_{\overline{v}}(0) = f(\overline{v})$.

[40] The same result holds if the support of $\tilde{f}$ is not full on $[\underline{v}, \overline{v}]^n$ but is of full dimension. Due to the symmetry, we can write $F_{(n)}(v) = n\int_{\underline{v}}^{v}\int_{S_{n-1}(v_i)} \tilde{f}(v_i, \mathbf{v}_{-i})d\mathbf{v}_{-i}dv_i$, where $S_{n-1}(v_i) \subset [\underline{v}, v_i]^{n-1}$ is the support of $\mathbf{v}_{-i}$ conditional on v_i and when all $v_{j\neq i} \leq v_i$. Then $f_{(n)}(v) = n\int_{S_{n-1}(v)} \tilde{f}(v, \mathbf{v}_{-i})d\mathbf{v}_{-i}$. When $v = \overline{v}$, $S_{n-1}(\overline{v})$ becomes the whole support of $\mathbf{v}_{-i}$. Therefore, $f_{(n)}(\overline{v}) = nf(\overline{v})$. Intuitively, when $\tilde{f}$ has a support of full dimension, when one product has the highest valuation, the conditional chance that another product also has the highest valuation is zero. Since it is equally likely for each product to have the highest valuation, we have $f_{(n)}(\overline{v}) = nf(\overline{v})$.

Now consider consumer surplus. Recall from (13) that consumer surplus under uniform pricing is $V_U(p) = \int_p^{\overline{v}}[1 - F_{(n)}(v)]dv$. Using $V_U'(\overline{v}) = 0$ and $V_U''(\overline{v}) = f_{(n)}(\overline{v})$, we can approximate it as

$$V_U \approx f_{(n)}(\overline{v})\frac{(\overline{v}-p)^2}{2} \approx f_{(n)}(\overline{v})\frac{\varepsilon^2}{8} \ .$$

On the other hand, recall from (16) that consumer surplus under price discrimination is $V_D(c) = \int_c^{\overline{v}}[1 - F_{(n-1)}(v)]dv$. Using $V_D'(\overline{v}) = 0$ and $V_D''(\overline{v}) = f_{(n-1)}(\overline{v})$, we can approximate it as

$$V_D \approx f_{(n-1)}(\overline{v})\frac{(\overline{v}-c)^2}{2} = f_{(n-1)}(\overline{v})\frac{\varepsilon^2}{2} \ ,$$

where $f_{(n-1)}(\cdot)$ is the density of the second highest order statistic. From (3) we have $f_{(n-1)}(v) = f_{(n)}(v) - nG(v|v)f(v) + n\int_v^{\overline{v}} g(v|v_i)dF(v_i)$, and so $f_{(n-1)}(\overline{v}) = f_{(n)}(\overline{v}) - nf(\overline{v}) = 0$, where we have used $f_{(n)}(\overline{v}) = nf(\overline{v})$. In other words, V_D is at most of the magnitude of ε^3 when $c = \overline{v} - \varepsilon$. Therefore, when $f(\overline{v}) > 0$, price discrimination harms consumers when c is sufficiently large. □

Proof of Corollary 1. From the proof of Proposition 2, we derive that each firm's profit under competitive personalized pricing is $\frac{1}{n}\Pi_D \approx f(\overline{v})\frac{\varepsilon^2}{2}$ when $c = \overline{v} - \varepsilon$, and the monopoly profit under uniform pricing is $\Pi_U^m \approx f(\overline{v})\frac{\varepsilon^2}{4}$. Therefore, we have $\frac{1}{n}\Pi_D > \Pi_U^m$. For consumer surplus, when $c = \overline{v} - \varepsilon$, $V_U^m \approx f(\overline{v})\frac{\varepsilon^2}{8}$ while V_D is at most of the magnitude of ε^3. □

Proof of Proposition 3. In the IID case industry profit under uniform pricing can be written as

$$\Pi_U = \frac{[1 - F(p)^n]^2/n}{F(p)^{n-1}f(p) + \int_p^{\overline{v}} f(v)dF(v)^{n-1}} \ .$$

Under the log-concavity condition, the uniform price p is decreasing in n, and so $F(p)^n$ must be of order $o(\frac{1}{n})$, i.e., $\lim_{n\to\infty} \frac{F(p)^n}{1/n} = 0$. Meanwhile, Theorem 1 in Gabaix et al. (2016), which approximates the Perloff-Salop price, has shown that as $n \to \infty$,

$$\int_{\underline{v}}^{\overline{v}} f(v)dF(v)^{n-1} \sim f(F^{-1}(1 - \frac{1}{n})) \cdot \Gamma(2 + \gamma) \ ,$$

where $\Gamma(\cdot)$ is the Gamma function. (Notice that $\Gamma(x)$ is non-monotonic in $x \in [1, 2]$ (decreases first and then increases) and is strictly positive but no greater than 1 in that range (with $\Gamma(1)=\Gamma(2)=1$).) At the same time, notice that $\int_{\underline{v}}^{p} f(v)dF(v)^{n-1} < F(p)^{n-1}\times\max_{v\in[\underline{v},p]} f(v)$, so it must be of order $o(\frac{1}{n})$ given f is finite. Therefore, as $n\to\infty$, we have

$$\Pi_U \sim \frac{[1-o(\frac{1}{n})]^2/n}{o(\frac{1}{n})+f(F^{-1}(1-\frac{1}{n}))\cdot\Gamma(2+\gamma)} \ .$$

Since the price is decreasing in n, π_U must be finite for any n. This implies that $\lim_{n\to\infty} nf(F^{-1}(1-\frac{1}{n}))>0$. Therefore, when n is large, those $o(\frac{1}{n})$ terms can be safely ignored. This yields

$$\Pi_U \sim \frac{1}{nf(F^{-1}(1-\frac{1}{n}))\cdot\Gamma(2+\gamma)} \ . \tag{39}$$

The case of perfect price discrimination is simpler. Industry profit in this case is

$$\Pi_D = \int_c^{\overline{v}} \frac{1-F(v)}{f(v)}dF(v)^n = \int_{F(c)}^{1} \frac{1-t}{f(F^{-1}(t))}dt^n \ .$$

Proposition 2 in Gabaix et al. (2016) has shown that, as $n\to\infty$,[41]

$$\mathbb{E}[v_{n:n}-v_{n-1:n}] = \int_0^1 \frac{1-t}{f(F^{-1}(t))}dt^n \sim \frac{\Gamma(1-\gamma)}{nf(F^{-1}(1-\frac{1}{n}))} \ .$$

Notice that

$$\Pi_D = \mathbb{E}[v_{n:n}-v_{n-1:n}] - \int_0^{F(c)} \frac{1-t}{f(F^{-1}(t))}dt^n \ ,$$

and the integrand in the second term is decreasing and so the second term is less than $\frac{F(c)^n}{f(\underline{v})}$ which is of order $o(\frac{1}{n})$. Therefore, the second term can be safely ignored when n is large, and so

$$\Pi_D \sim \frac{\Gamma(1-\gamma)}{nf(F^{-1}(1-\frac{1}{n}))} \ . \tag{40}$$

Comparing (39) and (40), we can claim that when n is sufficiently large, price discrimination reduces profit (and so improves consumer surplus) if

$$\Gamma(1-\gamma)\Gamma(2+\gamma) < 1 \ ,$$

which is true when $\gamma\in(-1,0)$.[42] □

[41] Note that $\int_0^1 t dt^n = 1-\frac{1}{n}$, so the approximation is intuitive up to the adjustment $\Gamma(1-\gamma)$.

[42] The equality holds when $\gamma=-1$ or 0. Unfortunately, in these cases the approximations are not precise enough to help us compare profit meaningfully in the limit.

Proof of Lemma 3. Notice that

$$\mathbb{E}[\max\{c, v_{n:n}\}] = \tfrac{1}{n}\mathbb{E}[\max\{c, v_n\}|v_n > \max\{v_1, ..., v_{n-1}\}] \\ + (1 - \tfrac{1}{n})\mathbb{E}[\max\{c, v_1, ..., v_{n-1}\}|v_n < \max\{v_1, ..., v_{n-1}\}] \ ,$$

and with Assumption 2 we also have

$$\mathbb{E}[\max\{c, \hat{v}_{n-1:n-1}\}] = \tfrac{1}{n}\mathbb{E}[\max\{c, v_1, ..., v_{n-1}\}|v_n > \max\{v_1, ..., v_{n-1}\}] \\ + (1 - \tfrac{1}{n})\mathbb{E}[\max\{c, v_1, ..., v_{n-1}\}|v_n < \max\{v_1, ..., v_{n-1}\}] \ .$$

Therefore, the match efficiency improvement in (23) is equal to

$$\tfrac{1}{n}\mathbb{E}[\max\{c, v_n\} - \max\{c, v_1, ..., v_{n-1}\}|v_n > \max\{v_1, ..., v_{n-1}\}] \ ,$$

which is just equal to (22).

We need to further show that both the free-entry equilibrium and the socially optimal solution are unique. (Otherwise, a free-entry equilibrium could differ from a socially optimal solution due to a selection issue.) It suffices to show that (22) is decreasing in n. To see that, it is more convenient to use the expression for Π_D in (15). Under Assumption 2, $x_c = v_i - \max_{j \neq i}\{c, v_j\}$ must become smaller in the sense of first-order stochastic dominance as one more firm is added, and so $1 - H_c(x)$ decreases in n for any x. This implies that $\frac{1}{n}\Pi_D$ decreases in n. □

Primitive conditions for Assumption 3. We show that Assumption 3 holds in the IID case with a log-concave f.

Claim 1. *In the IID case with a log-concave f, the profit function (26) is quasi-concave in p_i for any $p \geq c$.*

Proof. When $\alpha = 1$, we have the uniform pricing regime in Section 3.1; we showed earlier that profit here is quasi-concave. When $\alpha = 0$, profits equals $\pi_s(p_i)$ which is concave (see footnote 31) and hence also quasi-concave.

The remainder of the proof deals with the case $\alpha \in (0, 1)$ and $n \geq 2$ (the monopoly case is easy to deal with and so omitted). The profit function (26) looks similar to profit in models where part of the demand is price-elastic and part of it is captive; in such cases the profit function is not usually quasi-concave. The important difference here is that the "captive" segment is also price-elastic, and it turns out that at least in the IID case with log-concavity it "shrinks" in price just quickly enough.

To prove the result, we can focus on $p_i \geq c$ for which $1 - H_p(p_i - p) > 0$. Notice that the derivative of (26) with respect to p_i has the same sign as

$$1 - (p_i - c)\frac{h_p(p_i - p)}{1 - H_p(p_i - p)} + \frac{1-\alpha}{\alpha}\frac{1 - H_c(p_i - c)}{1 - H_p(p_i - p)}. \tag{41}$$

Therefore, if (41) is decreasing in p_i, (26) must be quasi-concave in p_i.[43] From Section 3, we already know that $1 - H_p(p_i - p)$ is log-concave in p_i in the IID case with a log-concave f, and so the first two terms in (41) are decreasing in p_i. Therefore, it suffices to show that the final term is decreasing in p_i.[44] A sufficient condition for this is that $1 - H_y(x - y)$ be *total positive of order 2* (TP_2) in (x, y). TP_2 implies that for x' and $x'' < x'$ and also y' and $y'' < y'$, we have

$$[1 - H_{y'}(x' - y')][1 - H_{y''}(x'' - y'')] \geq [1 - H_{y'}(x'' - y')][1 - H_{y''}(x' - y'')] \ ,$$

and so

$$\frac{1 - H_{y''}(x'' - y'')}{1 - H_{y'}(x'' - y')} \geq \frac{1 - H_{y''}(x' - y'')}{1 - H_{y'}(x' - y')} \ .$$

Then the desired result follows by setting $y'' = c$ and $y' = p$.

To prove the TP_2 property, we invoke the following theorem from Karlin (1968):

Theorem 1 (Theorem 5.2 in Karlin (1968))**.** *Let $f(\lambda, x)$ and $g(\lambda, x)$ be defined for $\Lambda \times X$, where Λ is linearly ordered and X is $(-\infty, \infty)$ (or the set of all integers). Suppose both f and g are TP_2 in (λ, x), and are PF_2 in x (i.e., $f(\lambda, x-\zeta)$ and $g(\lambda, x-\zeta)$ are TP_2 in (x, ζ) for fixed λ). Assume that*

$$h(\lambda, x) = \int_{-\infty}^{\infty} f(\lambda, x - \zeta) g(\lambda, \zeta) d\zeta$$

is well defined. Then h is TP_2 in (λ, x) and PF_2 in x.

Specialize to the IID case. Then

$$1 - H_y(x - y) = \int_x^{\infty} G(v - x + y) dF(v) = \int_{-\infty}^{\infty} G(y - (x - v))\mathbf{1}_{x-v<0} f(v) dv \ , \tag{42}$$

[43] Note that (41) is not equal to the derivative of (26), so we are not proving that (26) is concave.

[44] Note that if $c < p \leq \underline{v}$, then this last term is actually equal to 1. Otherwise, it is strictly less than 1, and as p_i increases the numerator becomes 0 before the denominator.

where $G(v) = F(v)^{n-1}$. First, notice that $f(v)$ is PF_2 in v (because f is log-concave) and TP_2 in (v, y) (which is trivially true). Second, we prove that $G(y-(x-v))\mathbf{1}_{x-v<0}$ is TP_2 in (y, x) for fixed v, and also TP_2 in (x, v) for fixed y. Since $F(v)$ is log-concave, $G(v)$ is log-concave and so PF_2 in v. Using the fact that if $k(x)$ is PF_2 then $k(x-y)$ is TP_2 in (x, y), we can claim that $G(y-(x-v))$ is TP_2 in (y, x) and also in (x, v). Indicator functions are TP_2.[45] Products of TP_2 functions are also TP_2. Hence the result follows. □

Claim 2. *In the IID case with a log-concave f, both $\frac{1-H_p(0)}{h_p(0)}$ and $\frac{1-H_c(p-c)}{h_p(0)}$ are non-increasing in p.*

Proof. The first result was already shown in Section 3. We now prove the second. Notice that

$$\frac{1-H_c(p-c)}{h_p(0)} = \frac{\int_p^{\overline{v}} G(v-p+c|v)dF(v)}{G(p|p)f(p) + \int_p^{\overline{v}} g(v|v)dF(v)} \ . \tag{43}$$

Following what we did in deriving primitive conditions for Assumption 1 in Section 3, one can show that (43) is decreasing in p if and only if

$$[G(p|p)f(p) + \int_p^{\overline{v}} g(v|v)dF(v)][G(c|p)f(p) + \int_p^{\overline{v}} g(v-p+c|v)dF(v)] \\ + (G_2(p|p)f(p) + G(p|p)f'(p)) \int_p^{\overline{v}} G(v-p+c|v)dF(v) \geq 0 \ ,$$

where G_2 was defined in (36). Dividing each side by $G(p|p)f(p)$ yields

$$\frac{G(p|p)f(p) + \int_p^{\overline{v}} g(v|v)dF(v)}{G(p|p)f(p)}[G(c|p)f(p) + \int_p^{\overline{v}} g(v-p+c|v)dF(v)] \\ + \left(\frac{G_2(p|p)}{G(p|p)} + \frac{f'(p)}{f(p)}\right) \int_p^{\overline{v}} G(v-p+c|v)dF(v) \geq 0 \ .$$

When f is log-concave, f'/f is decreasing. Using this property, we derive a sufficient

[45] First, note that $\mathbf{1}_{x-v<0}$ is trivially TP_2 in (y, x). Second, note that $\mathbf{1}_{x-v<0}$ is TP_2 in (x, v) if and only if for any $x' > x''$ and $v' > v''$ we have $\mathbf{1}_{x'-v'<0}\mathbf{1}_{x''-v''<0} \geq \mathbf{1}_{x'-v''<0}\mathbf{1}_{x''-v'<0}$. Note that if $x'-v'' < 0$, then all indicator functions are 1 and the inequality holds, whereas if $x'-v'' \geq 0$, then the right-hand side of the inequality is zero and so it also holds.

condition

$$\frac{G(p|p)f(p) + \int_p^{\overline{v}} g(v|v)dF(v)}{G(p|p)f(p)}[G(c|p)f(p) + \int_p^{\overline{v}} g(v-p+c|v)dF(v)]$$
$$+ \frac{G_2(p|p)}{G(p|p)} \int_p^{\overline{v}} G(v-p+c|v)dF(v) + \int_p^{\overline{v}} G(v-p+c|v)f'(v)dv \geq 0 \ .$$

Applying integration by parts to the last term and noticing the first fraction term is greater than 1 yields a sufficient condition

$$\frac{G_2(p|p)}{G(p|p)} \int_p^{\overline{v}} G(v-p+c|v)dF(v) + G(\overline{v}-p+c|\overline{v})f(\overline{v}) - \int_p^{\overline{v}} G_2(v-p+c|v)dF(v) \geq 0 \ .$$

In the IID case we have $G_2(v|v) = 0$ and so the above condition holds. $\square$

References

Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2020): "Too Much Data: Prices and Inefficiencies in Data Markets," *American Economic Journal: Microeconomics*, forthcoming.

Acquisti, A., C. Taylor, and L. Wagman (2016): "The Economics of Privacy," *Journal of Economic Literature*, 54(2), 442–92.

Ali, N., G. Lewis, and S. Vasserman (2020): "Voluntary Disclosure and Personalized Pricing," mimeo.

Anderson, S., A. Baik, and N. Larson (2019): "Price Discrimination in the Information Age: Prices, Poaching, and Privacy with Personalized Targeted Discounts," mimeo.

Anderson, S., A. de Palma, and Y. Nesterov (1995): "Oligopolistic Competition and the Optimal Provision of Products," *Econometrica*, 63(6), 1281–1301.

Anderson, S., A. de Palma, and J.-F. Thisse (1989): "Spatial Price Policies Reconsidered," *Journal of Industrial Economics*, 38(1), 1–18.

Aparicio, D., Z. Metzman, and R. Rigobon (2021): "The Pricing Strategies of Online Grocery Retailers," mimeo.

Armstrong, M. (2007): "Recent Developments in the Economics of Price Discrimination," in *Advances in Economics and Econometrics (Theory and Applications, Ninth World Congress)*, ed. by R. Blundell, W. Newey, and T. Persson. Cambridge University Press.

Bergemann, D., and A. Bonatti (2019): "Markets for Information: An Introduction," *Annual Review of Economics*, 11(1), 85–107.

Bergemann, D., A. Bonatti, and T. Gan (2020): "The Economics of Social Data," mimeo.

Caplin, A., and B. Nalebuff (1991): "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 59(1), 25–59.

Chen, Y., and G. Iyer (2002): "Consumer Addressability and Customized Pricing," *Marketing Science*, 21(2), 197–208.

Chen, Y., and M. Riordan (2007): "Price and Variety in the Spokes Model," *The Economic Journal*, 117(522), 897–921.

Chen, Z., C. Choe, and N. Matsushima (2020): "Competitive Personalized Pricing," *Management Science*, 66(9), 4003–4023.

Choi, J. P., D.-S. Jeon, and B.-C. Kim (2019): "Privacy and personal data collection with information externalities," *Journal of Public Economics*, 173, 113–124.

Council of Economic Advisers (2015): "Big Data and Differential Pricing," Discussion paper, The White House.

Dube, J.-P., and S. Misra (2019): "Personalized Pricing and Customer Welfare," mimeo.

Fudenberg, D., and J. M. Villas-Boas (2007): "Behavior-Based Price Discrimination and Customer Recognition," in *Handbook on Economics and Information Systems (Vol. 1)*, ed. by T. Hendershott. Elsevier.

Gabaix, X., D. Laibson, D. Li, H. Li, S. Resnick, and C. G. de Vries (2016): "The Impact of Competition on Prices with Numerous Firms," *Journal of Economic Theory*, 165, 1–24.

Hannak, A., G. Soeller, D. Lazer, A. Mislove, and C. Wilson (2014): "Measuring Price Discrimination and Steering on E-Commerce Web Sites," in *Proceedings of the 2014 Conference on Internet Measurement Conference*, p. 305–318.

Ichihashi, S. (2020): "Online Privacy and Information Disclosure by Consumers," *American Economic Review*, 110(2), 569–595.

——— (2021): "The Economics of Data Externalities," *Journal of Economic Theory*, 196, 105316.

Kehoe, P., B. Larsen, and E. Pastorino (2020): "Dynamic Competition in the Era of Big Data," mimeo.

Montes, R., W. Sand-Zantman, and T. Valletti (2019): "The Value of Personal Information in Online Markets with Endogenous Privacy," *Management Science*, 65(3), 1342–1362.

OECD (2018): "Personalised Pricing in the Digital Era," Discussion paper, Organisation for Economic Co-operation and Development.

OFT (2013): "Personalised Pricing - Increasing Transparency to Improve Trust," Discussion paper, Office of Fair Trading.

Perloff, J., and S. Salop (1985): "Equilibrium with Product Differentiation," *Review of Economic Studies*, 52(1), 107–120.

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

Quint, D. (2014): "Imperfect Competition with Complements and Substitutes," *Journal of Economic Theory*, 152, 266–290.

Shaffer, G., and Z. J. Zhang (2002): "Competitive One-to-One Promotions," *Management Science*, 48(9), 1143–1160.

Shiller, B. R. (2020): "Approximating Purchase Propensities and Reservation Prices from Broad Consumer Tracking," *International Economic Review*, 61(2).

——— (2021): "Optimized Sticky Targeted Pricing," mimeo.

Stole, L. (2007): "Price Discrimination and Competition," in *Handbook of Industrial Organization (Vol. 3)*, ed. by M. Armstrong, and R. Porter. North Holland.

Tan, G., and J. Zhou (2021): "Consumer Heterogeneity and Inefficiency in Oligopoly Markets," mimeo.

Thisse, J.-F. T., and X. Vives (1988): "On The Strategic Choice of Spatial Price Policy," *American Economic Review*, 78(1), 122–137.

Varian, H. (1989): "Price Discrimination," in *Handbook of Industrial Organization (Vol. 1)*, ed. by R. Schmalensee, and R. Willig. North-Holland.

Waldfogel, J. (2015): "First Degree Price Discrimination Goes to School," *Journal of Industrial Economics*, 63(4), 569–597.

Zhou, J. (2017): "Competitive Bundling," *Econometrica*, 85(1), 145–172.