Market Segmentation and Product Steering*

Stefan Terstiege† Adrien Vigier‡

March 15, 2023

Abstract

A monopolistic seller possesses an inventory containing distinct products, each consumer wishes to buy a single product, and the seller can steer consumers' choices. We fully characterize the producer-consumer surplus pairs induced by market segmentation as product variety becomes large. We uncover a trade-off between consumer surplus and social welfare, prove the redundancy of price discrimination, and analyze the welfare implications of privacy. We apply our results to study market segmentation arising from the sale of consumer data by data intermediaries.

Keywords: market segmentation, product variety, steering, consumer privacy, data intermediaries.

JEL-Classification: D42, D83

*We thank Antoine Dubus, Matteo Escudé, Shota Ichihashi, and Kym Pram for giving us very useful feedback, as well as various seminar audiences.

†Maastricht University. Email: s.terstiege@maastrichtuniversity.nl.

‡University of Notthingam. Email: adrien.vigier@nottingham.ac.uk

1 Introduction

The advent of the digital age has transformed marketplaces. Two salient features specifically distinguish online markets from traditional ones: (a) the enormous product variety that they entail, and (b) the availability and use of vast amounts of consumer data. This paper examines how multi-product sellers' access to consumer data affects welfare.

Whereas large brick-and-mortar bookstores hold between 40,000 and 100,000 book titles, Amazon sells several millions of titles (Brynjolfsson, Hu, and Smith, 2003). Shein launches about 6,000 new SKUs every day, and in total stocks about 600,000 distinct products (Economist, 2021). Zalando, an online fashion retailer, has 65,539 items in the category "Men's T-Shirts & Polos" alone.[1] Similarly, online shoe retailers may offer over 50,000 distinct models, whereas traditional retailers usually stock at most a few thousand ones (Quan and Williams, 2018).

While the online shelf space is virtually unlimited, consumers' attention and time are limited. Hence, a central problem for online retailers is to direct consumers towards products which they are likely to value. As an input to advertising and recommender systems, but also to pricing algorithms, online retailers' information about consumers plays a crucial role.

Comprehensive privacy regulation in the European Union has fundamentally modified access to consumer data. The General Data Protection Regulation prohibits the collection, storage, use, and dissemination of personal data unless at least one of several conditions is fulfilled, for example consent (Regulation (EU) 2016/679, Article 6). Moreover, antitrust authorities globally appear to take a more aggressive stance towards market power in data-driven industries. The Digital Markets Act prescribes that platforms do not combine, without additional consent, their personal data with those collected by subsidiaries (Regulation (EU) 2022/1925, Article 5). The US Federal Trade Commission has filed a lawsuit against Meta (*Federal Trade Commission v. Meta Platforms, Inc.*), the parent company of Facebook, alleging that it systematically accumulated market power through its acquisitions of Instagram and WhatsApp.

On this background, we address the following questions: Which producer-consumer sur-

[1] `https://en.zalando.de/men-clothing-shirts/`; accessed February 27, 2023.

plus pairs are attainable when multi-product sellers have access to consumer data? Can regulation enhance the privacy of consumers without sacrificing efficiency or consumer surplus? Why do online retailers often refrain from personalized pricing although they have so much data? What are the welfare consequences of data intermediation, and which data do intermediaries supply? What are the effects of competition in the data market?

In our model, a monopolistic seller possesses an inventory containing distinct products, and each consumer wishes to buy a single product. The seller's information induces a market segmentation, as in Bergemann, Brooks, and Morris (2015). In particular, each market segment is represented by a probability distribution over the set of possible valuation vectors for the different products of the seller. The seller's problem is to choose, for each market, which product to offer and at what price.

The central result of our paper (Theorem 1) characterizes the combinations of producer and consumer surpluses that result from market segmentation. This characterization rests on the observation that when product variety is large, the combinations of producer and consumer surpluses obtained through market segmentation approximately coincide with the producer-consumer surplus pairs obtained in a much simpler auxiliary setting. Specifically, in this auxiliary setting a designer chooses the distribution of consumers' valuations for a *single* product, subject to certain constraints.

We explore the effect of market segmentation on social welfare and the distribution of surplus. Efficiency requires each consumer to purchase the product that he values the most. But if the seller can identify for each consumer the product that he values the most then, in equilibrium, the prices at which products are sold must be high. By this logic, we show that along the Pareto frontier social welfare decreases as consumer surplus goes up (Proposition 1). Hence, when product variety is large, efficiently transferring surplus from the seller to consumers through market segmentation is not possible; in this sense segmentation is less potent with multiple products than with a single product (Bergemann, Brooks, and Morris, 2015). We show on the other hand that transferring surplus from the seller to consumers all the while improving the latter's privacy is possible. This property is important when, as is quite common, privacy is perceived as having intrinsic value.

Interestingly, in sharp contrast to single-product settings, price discrimination has no role to play when the number of products is large. We prove that the surplus pairs that result from market segmentation approximately coincide with the surplus pairs induced by market segmentations under which the seller chooses not to price discriminate (Proposition 3). Our analysis thus sheds light on the use of price discrimination, by showing that product steering effectively makes price discrimination redundant. It is perhaps unsurprising therefore that whereas targeted ads and personalized product recommendations are common, overt price discrimination (personalized pricing) appears not to be widespread (see, e.g., Cavallo, 2017; OECD, 2018; DellaVigna and Gentzkow, 2019).

We then study market segmentation arising from the sale of consumer data by intermediaries. To this end, we augment the model by an initial stage in which data intermediaries propose data policies to a consumer. A data policy specifies which data will be made available to the seller, provided that the consumer gives his consent. The intermediaries that obtain the consumer's consent subsequently sell their data to the seller of the products.

We precisely pin down producer and consumer surplus for a monopolistic and a competitive data market, respectively (Proposition 4). In particular, we show that competition between data intermediaries results in greater consumer surplus, but reduces social welfare because the seller ends up offering less suitable products. We go on to show that competition in the data market may further benefit consumers by giving them more privacy (Proposition 5). Our results thus suggest that efforts made to limit the market power of data-driven businesses (exemplified by the Digital Markets Act and the lawsuit against Meta) can significantly improve consumer welfare.

The rest of the paper is organized as follows. The related literature is discussed below. The model is presented in Section 2. Section 3 states and proves the central theorem of the paper. In Section 4, we examine the effect of market segmentation on social welfare, and the relation between consumer privacy and welfare. We show in Section 5 that price discrimination is irrelevant for the set of producer-consumer surplus pairs feasible through market segmentation. Section 6 studies online markets with data intermediaries. Finally, Section 7 concludes. Omitted proofs are in the Appendix; the Online Appendix contains further proofs and extensions.

1.1 Related Literature

We contribute to the study of market segmentation initiated by Bergemann, Brooks, and Morris (2015). The study of market segmentation in multi-product monopolies starts with Ichihashi (2020); other major contributions include Hidir and Vellodi (2021), Haghpanah and Siegel (2022a,b), and Pram (2021). In a single-product setting, Bergemann, Brooks, and Morris (2015) characterize the entire set of producer-consumer surplus pairs attainable by market segmentation. Obtaining a similar characterization in a multi-product setting is notoriously hard. We show that letting the number of products tend to infinity renders this problem solvable.

Ichihashi (2020) and Hidir and Vellodi (2021) study *consumer-optimal* market segmentation. As in our model, consumer data can be used for steering and pricing. Ichihashi (2020) compares the properties of consumer-optimal market segmentation in two pricing regimes: one in which the seller commits to one price, and one in which the seller sets prices after having observed the market segment to which the consumer belongs. The author proves that letting the seller personalize prices: (a) induces inefficient trade whereby the seller occasionally offers a product that is *not* the consumer's most-preferred product; (b) decreases producer surplus; (c) increases consumer surplus.[2] Hidir and Vellodi (2021) explore a setting in which individual consumers exert a form of control over the market segment to which they belong. Specifically, each consumer chooses his preferred market segment through cheap-talk communication. The authors introduce the notion of an incentive-compatible market segmentation, and show that the consumer-optimal incentive-compatible segmentation consists of pooling segments wide enough to keep prices low but narrow enough to guarantee trade. In contrast to these papers, we do not focus on situations in which consumers have perfect control of the seller's information. We characterize, and study the properties of, the entire set of surplus pairs induced by arbitrary market segmentation as product variety becomes large.

The model of Haghpanah and Siegel (2022b) is more general than ours, both in terms of

[2]The fact that personalized pricing results in lower producer surplus provides a possible explanation for why this form of price discrimination seems quite rare.

consumers' preferences, and in terms of selling mechanisms considered. The authors prove that any market in which profit maximization leads to inefficiency can be segmented in a way that increases welfare in the sense of Pareto. Haghpanah and Siegel (2022a) find conditions under which the multi-product counterpart of the "surplus triangle" of Bergemann, Brooks, and Morris (2015) corresponds to the set of feasible producer-consumer surplus pairs. Pram (2021) explores a different setting in which individual consumers exert some control over the market segment to which they belong. Specifically, each consumer communicates hard information, and thus chooses his preferred market segment from within a subset of segments. The author precisely characterizes those situations in which some equilibrium Pareto dominates the equilibrium without evidence.

Our paper also contributes to recent economic research on markets with data intermediaries. In Hidir and Vellodi (2021), a single online platform provides consumer data to sellers, and each consumer faces an opportunity cost of participation. The authors show that a higher opportunity cost leads to lower product prices and lower match quality between consumers and products. This key insight resonates with our results if one interprets the higher opportunity cost as being due to greater competition among online platforms. De Cornière and De Nijs (2016) consider a setting where online platforms auction advertising slots. To the extent that an increase in the number of slots plays the same role as a decrease in the number of bidders, competition between platforms benefits consumers by inducing lower prices in the product market. In Bounie, Dubus, and Waelbroeck (2022), data intermediaries first acquire costly information, and then choose the information sold to downstream firms. Competition between data intermediaries benefits consumers because it induces the former to acquire less information, which in turn reduces extraction by sellers in the product market.

Bergemann and Bonatti (2015) and Ichihashi (2021) offer different perspectives than the aforementioned papers, and put forth that competition need not benefit consumers. In Bergemann and Bonatti (2015), the key aspect is that raising the price at which information about one consumer is sold to downstream firms reduces the demand for information about all other consumers. By contrast, in Ichihashi (2021) data intermediaries can compensate customers through monetary transfers, and so the degree of competitiveness in the market for data leaves trade efficiency in the product market unaffected.

A distinct literature studies steering of consumers by intermediaries seeking out commissions (e.g., Armstrong and Zhou, 2011; Hagiu and Jullien, 2011; Inderst and Ottaviani, 2012; de Cornière and Taylor, 2019; Teh and Wright, 2022). In Nocke and Rey (2022), a multi-product monopolist influences the search of consumers for suitable products via the design and pricing of its product line. Finally, a large literature studies the economics of privacy (see Acquisti, Taylor, and Wagman, 2016, for a survey). An important focus of this literature is on intertemporal price discrimination based on customer recognition (see, e.g., Taylor, 2004; Villas-Boas, 2004; Acquisti and Varian, 2005; Conitzer, Taylor, and Wagman, 2012; Bonatti and Cisternas, 2020). In this paper, we abstract from the collection of data and directly compare the privacy afforded to consumers by different market segmentations.

2 Baseline Model

Throughout the paper,

$$X := \{x_1, \ldots, x_m\}, \quad 0 < x_1 < \cdots < x_m,$$

and f is a distribution in ΔX that has full support.[3]

There is a seller (she) with an inventory containing n variants of a product; the seller possesses an infinite supply of each product. There is also a continuum of unit-demand consumers.[4] Any consumer's valuations for the n products can be represented by some vector $\mathbf{v} \in X^n$, with v_k (the k^{th} component of the vector $\mathbf{v}$) indicating this consumer's valuation for product k. We use the generic notation μ for a probability distribution over X^n, that is, $\mu \in \Delta X^n$; we refer to such a distribution as a *market*. The k-marginal of a market μ is denoted by μ_k.[5]

[3]The notation ΔY indicates the set of all distributions with finite support over the set Y. Our assumptions that f has full support and that $x_1 > 0$ merely simplify the exposition.

[4]That is, the value attached by a consumer to any set of products is equal to the consumer's maximum valuation for a single item in this set.

[5]Thus, $\mu_k \in \Delta X$, with $\mu_k(x) = \sum_{\mathbf{v}:\, v_k = x} \mu(\mathbf{v})$ for all $x \in X$.

The proportion $\bar{\mu}(\mathbf{v})$ of consumers whose valuations are given by the vector $\mathbf{v}$ satisfies

$$\bar{\mu}(\mathbf{v}) = \prod_k f(v_k), \quad \forall \mathbf{v} \in X^n. \tag{1}$$

We refer to the market $\bar{\mu}$ defined through (1) as the *aggregate market*.

A typical element of $\Delta\Delta X^n$ is denoted by τ; if

$$\sum_\mu \tau(\mu)\mu(\mathbf{v}) = \bar{\mu}(\mathbf{v}), \quad \forall \mathbf{v} \in X^n, \tag{2}$$

then τ is called a *market segmentation*. For a fixed market segmentation τ, the problem of the seller is to choose for each market comprised in the support of τ, which product to offer and at what price.

We use the generic notation ρ for a strategy of the seller, with $\rho_\mu(k,p)$ representing the probability that the seller offers product k at price p in market μ. We suppose that if a consumer's valuation for product k equals v_k then, when offered product k at price v_k, the consumer decides to buy. The producer surplus generated by the strategy ρ is[6]

$$\Pi_\tau(\rho) := \sum_\mu \tau(\mu) \sum_{k,p} \rho_\mu(k,p) p \sum_{x \geq p} \mu_k(x);$$

the corresponding consumer surplus is

$$U_\tau(\rho) := \sum_\mu \tau(\mu) \sum_{k,p} \rho_\mu(k,p) \sum_{x \geq p} \mu_k(x)(x-p).$$

We say that a surplus pair (π, u) is *feasible* if there exist a market segmentation τ as well as a strategy $\rho^* \in \operatorname{argmax}_\rho \Pi_\tau(\rho)$ such that $\pi = \Pi_\tau(\rho^*)$ and $u = U_\tau(\rho^*)$. The set of feasible surplus pairs is denoted by S_n.

Additional expository assumptions: we assume that $p \mapsto p \sum_{x \geq p} f(x)$ possesses a unique maximizer, which we denote by p_0. Then, letting

$$\pi_0 := p_0 \sum_{x \geq p_0} f(x),$$

we assume that $\pi_0 \in X$.

[6]The mass of consumers is normalized to one, to save on notation.

2.1 Discussion of the Model

A market segmentation could either depict geographically distinct markets, or summarize information available to an online seller, perhaps due to the seller's access to consumers' browsing histories or the use of cookies. An online seller might be able to determine, say, the age and nationality of each consumer. In this case, a market would represent the distribution of valuations within a given age group of a certain nationality.

The model supposes that the seller offers a single product in each market. We capture thereby situations in which a firm has a large inventory consisting of many different variants of a given good or service, and where the number of variants is far greater than the constraints imposed by consumers' limited attention or cognitive costs, thus forcing sellers to make strategic choices with regard to the products they offer in any given market. This feature is central to online retailing, among other things (Brynjolfsson, Hu, and Smith, 2003; Anderson, 2006). Note that, what we refer to as a product in the model might in practice represent a sub-category of products, such as "Italian movies from the 1960's", for example.

Our model allows the seller to engage in third-degree price discrimination. In online markets, this assumption seems realistic. At any rate, we show in Section 5 that, when product variety is large, whether or not the seller can price discriminate is inconsequential for our results.

Finally, the model supposes that a consumer's valuation for one product is statistically independent of his valuation for other products. Naturally, in certain applications a consumer's valuations for different products may be correlated; for instance, books by the same author, or from the same genre, might be valued similarly. We show in Section OA.2 of the Online Appendix that our main results carry through in a version of the model that allows for correlation between valuations. Similarly, we show in Section OA.3 that our main results carry through when the set of possible valuations is a continuum.

3 The Welfare Bounds of Market Segmentation

In this section, we characterize the set of feasible surplus pairs when product variety is large.

We begin with a couple of key definitions. Firstly, for $i \in \{1, \dots, m\}$, define $g_i \in \Delta X$ by

$$g_i(x_j) := \begin{cases} 0 & \text{if } j < i, \\ x_i/x_j - x_i/x_{j+1} & \text{if } i \le j < m, \\ x_i/x_m & \text{if } j = m. \end{cases}$$

It is readily checked that

$$p \sum_{x \ge p} g_i(x) = \begin{cases} p & \text{for all } p \in \{x_1, \dots, x_i\}, \\ x_i & \text{for all } p \in \{x_{i+1}, \dots, x_m\}. \end{cases} \tag{3}$$

Thus, if the seller had a single product and the "demand" for this product were given by g_i, then the seller would be indifferent between all prices in $\{x_i, \cdots, x_m\}$; furthermore, if she were to choose the price x_i, the surplus of the consumers would be equal to $\sum_{x \ge x_i} g_i(x)(x - x_i)$. The latter remark motivates the definition of a mapping $\overline{u} : [x_1, x_m] \to \mathbb{R}$ such that

$$\overline{u}(x_i) := \sum_{x \ge x_i} g_i(x)(x - x_i), \tag{4}$$

and

$$\overline{u}\big((1-\lambda)x_i + \lambda x_{i+1}\big) := (1-\lambda)\overline{u}(x_i) + \lambda \overline{u}(x_{i+1}), \quad \forall \lambda \in [0,1]. \tag{5}$$

Finally, let

$$S := \big\{(\pi, u) \in \mathbb{R}^2 \mid \pi \in [\pi_0, x_m], u \in [0, \overline{u}(\pi)]\big\}.$$

Figure 1 illustrates S for $X = \{1,2,3,4,5,6,7\}$ and $\pi_0 = 2$. As product variety becomes large, the set S approximately coincides with the feasible surplus pairs.

Theorem 1. *For every $n \in \mathbb{N}$, the set S_n of feasible surplus pairs is contained in S. Moreover, for every $(\pi, u) \in S$, there exists a sequence $\big((\pi_n, u_n)\big)_{n \in \mathbb{N}}$ such that $(\pi_n, u_n) \in S_n$ and $(\pi_n, u_n) \underset{n \to \infty}{\longrightarrow} (\pi, u)$.*

To shed light on Theorem 1, we will first argue that S may be viewed as the surplus pairs attainable in a single-product setting without market segmentation, but where the valuation distribution is an object of design.

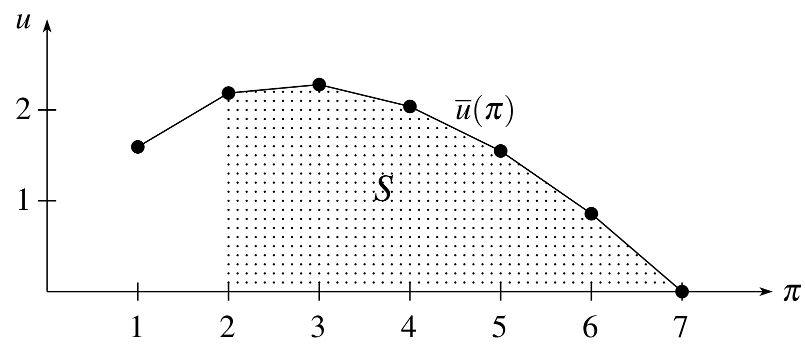


Figure 1: The set of feasible surplus pairs as product variety becomes large

Formally, consider the following auxilliary setting, with a single product and a single market.[7] The valuations of the consumers are distributed according to $\hat{f} \in \Delta X$; $\Pi_{\hat{f}}(\hat{\rho})$ denotes the producer surplus generated by the pricing strategy $\hat{\rho}$, and $U_{\hat{f}}(\hat{\rho})$ the corresponding consumer surplus.[8] Say that a surplus pair (π, u) is attainable if there exist a valuation distribution $\hat{f}$ and a pricing strategy $\hat{\rho}^* \in \arg\max_{\hat{\rho}} \Pi_{\hat{f}}(\hat{\rho})$ such that $\pi = \Pi_{\hat{f}}(\hat{\rho}^*)$ and $u = U_{\hat{f}}(\hat{\rho}^*)$.

One shows that the surplus pairs attainable in this auxiliary setting are[9]

$$\left\{(\pi, u) \in \mathbb{R}^2 \mid \pi \in [x_1, x_m], u \in [0, \overline{u}(\pi)]\right\}.$$

We conclude by Theorem 1 that, as $n \to \infty$, the feasible surplus pairs approximately coincide with the surplus pairs attainable in a single-product setting without market segmentation, but where the valuation distribution is an object of design subject to the constraints that (a) the support of this distribution is contained in X, (b) the resulting surplus of the seller is not smaller than π_0.

Specifically, the first part of Theorem 1 tells us that, regardless of the market segementation, the seller obtains a surplus $\pi \in [\pi_0, x_m]$, while the surplus of the consumers is bounded from above by $\overline{u}(\pi)$.

The second part of the theorem tells us that, as product variety becomes large, any element of S may be approximately attained through market segmentation. The basic idea is

[7]In particular, in this setting a strategy $\hat{\rho}$ of the seller consists simply of a price distribution.

[8]As in the baseline model, suppose that any consumer who is indifferent between buying and not buying always decides to buy.

[9]See Lemma 1 in Condorelli and Szentes (2020).

as follows. Pick some $x_i \in X$ such that $x_i \geq \pi_0$. For each product $k \in \{1, \cdots, n\}$, and independently across products, divide the consumers in two groups such that in one group, the valuation for product k is distributed according to g_i. Mark every consumer in this group with the label $a_k = \texttt{yes}$, and every other consumer with the label $a_k = \texttt{no}$. This creates 2^n markets; each market consists of the consumers with the same sequence of labels $\mathbf{a} \in \{\texttt{yes}, \texttt{no}\}^n$. We show that if $a_k = \texttt{yes}$ for some product, the seller can do no better than to offer such a product at a price of x_i. When n is large, the only market with $a_k = \texttt{no}$ for all products is vanishingly small. By segmenting the aggregate market in this way, we thus generate surplus approaching x_i for the seller and $\overline{u}(x_i)$ for consumers.

3.1 Proof of Theorem 1

Readers uninterested in the technical details of the analysis can jump to Section 4 without loss. Our proof of Theorem 1 builds on three lemmas.

Lemma 1. *If $(\pi, u) \in S_n$ then $u \leq \overline{u}(\pi)$.*

Proof. We treat below the case $\pi \in X$; the proof for the remaining case is similar, and therefore omitted. As $(\pi, u) \in S_n$, there exist a market segmentation τ and a strategy ρ of the seller that is optimal given τ, such that $\pi = \Pi_\tau(\rho)$ and $u = U_\tau(\rho)$. Then, define the distribution $h \in \Delta X$ by

$$h(x) := \sum_{\mu} \tau(\mu) \sum_{k,p} \rho_\mu(k,p) \mu_k(x), \ \ \forall x \in X.$$

Letting x_i be the element of X such that $\pi = x_i$, we have for any $q \in \{x_i, \cdots, x_m\}$:

$$\begin{aligned} \sum_{\mu} \tau(\mu) \sum_{k,p} \rho_\mu(k,p) q \sum_{x \geq q} \mu_k(x) \quad &\leq \quad \sum_{\mu} \tau(\mu) \sum_{k,p} \rho_\mu(k,p) p \sum_{x \geq p} \mu_k(x) \\ &= \quad \Pi_\tau(\rho) = \pi = q \sum_{x \geq q} g_i(x). \end{aligned}$$

The inequality in the previous sequence follows from ρ being optimal given τ; the last equality follows from (3). Dividing through by q, we see that g_i first-order stochastically dominates h. Hence,

$$\begin{aligned}
u &= \sum_{\mu}\tau(\mu)\sum_{k,p}\rho_{\mu}(k,p)\sum_{x\geq p}\mu_k(x)(x-p) \\
&= \sum_{\mu}\tau(\mu)\sum_{k,p}\rho_{\mu}(k,p)\sum_{x\geq p}\mu_k(x)x-\pi \\
&\leq \sum_{\mu}\tau(\mu)\sum_{k,p}\rho_{\mu}(k,p)\sum_{x}\mu_k(x)x-\pi \\
&= \sum_{x}h(x)x-\pi \\
&\leq \sum_{x}g_i(x)x-\pi \\
&= \overline{u}(\pi). \qquad \square
\end{aligned}$$

In what follows, we say that a distribution $\tau\in\Delta\Delta X^n$ is the product of distributions $(\tau^k)_{k=1}^n$ in $\Delta\Delta X$, if $\tau(\mu)>0$ implies

$$\mu(\mathbf{v})=\prod_k\mu_k(v_k), \quad \forall\mathbf{v}\in X^n, \tag{6}$$

and

$$\tau(\mu)=\prod_k\tau^k(\mu_k). \tag{7}$$

Lemma 2. *Let τ be the product of $(\tau^k)_{k=1}^n$. If*

$$\sum_{\mu_k}\tau^k(\mu_k)\mu_k(x)=f(x) \quad \text{for all } k\in\{1,\cdots,n\} \text{ and all } x\in X, \tag{8}$$

then τ is a market segmentation.

The proof of this elementary result is relegated to the appendix. Our last lemma identifies points of S which may be approached by feasible surplus pairs as the number of products becomes large.

Lemma 3. *For every $x_i\in\{\pi_0,\ldots,x_m\}$ and every $p\in\{x_i,\ldots,x_m\}$, there exists a sequence $\big((\pi_n,u_n)\big)_{n\in\mathbb{N}}$ such that $(\pi_n,u_n)\in S_n$ for every n, and*

$$(\pi_n,u_n)\underset{n\to\infty}{\longrightarrow}\left(x_i,\sum_{x\geq p}g_i(x)(x-p)\right). \tag{9}$$

Proof. Let $x_i \in \{\pi_0, \ldots, x_m\}$, and $p \in \{x_i, \ldots, x_m\}$. Choose $\lambda \in (0,1)$ such that $\lambda g_i(x) \leq f(x)$ for all $x \in X$, and define

$$h(x) := \frac{f(x) - \lambda g_i(x)}{1-\lambda}. \tag{10}$$

Note that $h(x) \geq 0$ for all $x \in X$, and $\sum_x h(x) = 1$, whence $h \in \Delta X$. Moreover,

$$\lambda g_i(x) + (1-\lambda)h(x) = f(x), \quad \forall x \in X. \tag{11}$$

We claim that

$$\max_q q \sum_{x \geq q} g_i(x) = p \sum_{x \geq p} g_i(x) = x_i \geq \max_q q \sum_{x \geq q} h(x). \tag{12}$$

The equalities in (12) follow from (3). The fact that

$$x_i \geq q \sum_{x \geq q} h(x), \quad \forall q \leq x_i,$$

is immediate, as h is a distribution. Lastly, for all $q > x_i$:

$$x_i \ \geq \pi_0 \geq q \sum_{x \geq q} f(x) = \lambda q \sum_{x \geq q} g_i(x) + (1-\lambda) q \sum_{x \geq q} h(x) = \lambda x_i + (1-\lambda) q \sum_{x \geq q} h(x).$$

So

$$x_i \geq q \sum_{x \geq q} h(x), \quad \forall q > x_i,$$

which finishes the proof of (12).

Next, define $\tau^k \in \Delta\Delta X$ by

$$\tau^k(g_i) = \lambda = 1 - \tau^k(h),$$

and let τ be the product of $(\tau^k)_{k=1}^n$. By coupling (11) with Lemma 2, notice that τ is a market segmentation.

Now let ρ be a strategy of the seller with the following properties. For every market μ in the support of τ such that $\mu_k = g_i$ for some product k, offer any such product at price p. If $\mu_k = h$ for all products $k \in \{1, \cdots, n\}$, on the other hand, offer any product at some fixed price

$$q' \in \operatorname*{argmax}_q q \sum_{x \geq q} h(x).$$

By (12), the strategy ρ is optimal given τ. The resulting surplus of the seller is

$$\pi_n := \Pi_\tau(\rho) = (1-(1-\lambda)^n)x_i + (1-\lambda)^n q' \sum_{x \geq q'} h(x);$$

the consumer surplus is

$$u_n := U_\tau(\rho) = (1-(1-\lambda)^n) \sum_{x \geq p} g_i(x)(x-p) + (1-\lambda)^n \sum_{x \geq q'} h(x)(x-q').$$

Then $(\pi_n, u_n) \in S_n$, and since $\lambda > 0$, the limit in (9) is established. □

We are now ready to prove the theorem.

Proof of Theorem 1. The strategy ρ given by $\rho_\mu(1, p_0) = 1$ for every μ yields surplus π_0 to the seller, so the seller can guarantee herself a surplus of π_0 regardless of the market segmentation. The first part of the theorem then follows from Lemma 1.

We now prove the second part of the theorem. Let (π', u') and (π'', u'') be arbitrary points in the set S. Suppose $(\pi'_n, u'_n) \in S_n$ for every n, with

$$(\pi'_n, u'_n) \underset{n\to\infty}{\longrightarrow} (\pi', u').$$

Similarly, suppose $(\pi''_n, u''_n) \in S_n$ for every n, with

$$(\pi''_n, u''_n) \underset{n\to\infty}{\longrightarrow} (\pi'', u'').$$

Let τ'_n and τ''_n be market segmentations inducing the surplus pairs (π'_n, u'_n) and (π''_n, u''_n), respectively. The set of market segmentations is evidently convex. Furthermore, note that for all $\zeta \in [0,1]$, some optimal strategy of the seller given $(1-\zeta)\tau'_n + \zeta\tau''_n$ yields a surplus of $(1-\zeta)\pi'_n + \zeta\pi''_n$ for the seller and $(1-\zeta)u'_n + \zeta u''_n$ for the consumers. We conclude that there exists a sequence $\big((\pi_n, u_n)\big)_{n\in\mathbb{N}}$ such that $(\pi_n, u_n) \in S_n$ for every n, and

$$(\pi_n, u_n) \underset{n\to\infty}{\longrightarrow} (1-\zeta)(\pi', u') + \zeta(\pi'', u'').$$

Now, for all $x_i \in \{\pi_0, \ldots, x_m\}$, Lemma 3 gives us sequences $\big((\pi'_n, u'_n)\big)_{n\in\mathbb{N}}$ and $\big((\pi''_n, u''_n)\big)_{n\in\mathbb{N}}$ such that, firstly, (π'_n, u'_n) and (π''_n, u''_n) belong to S_n for every n, and secondly,

$$(\pi'_n, u'_n) \underset{n\to\infty}{\longrightarrow} \big(x_i, \overline{u}(\pi)\big) \quad \text{and} \quad (\pi''_n, u''_n) \underset{n\to\infty}{\longrightarrow} (x_i, 0).$$

We conclude using the previous observation that, for every $(\pi, u) \in S$, there exists a sequence $\big((\pi_n, u_n)\big)_{n\in\mathbb{N}}$ such that $(\pi_n, u_n) \in S_n$ for every n and $(\pi_n, u_n) \underset{n\to\infty}{\longrightarrow} (\pi, u)$. □

4 Social Welfare and Consumer Privacy

In view of Theorem 1, we refer to the set of maximal elements of S as the *Pareto frontier*; that is, a surplus pair (π, u) belongs to the Pareto frontier if (i) $u = \overline{u}(\pi)$ and (ii) $\overline{u}(\pi) > \overline{u}(\pi')$ for all $\pi' \in (\pi, x_m]$. The *social welfare* at a surplus pair (π, u) is defined as $\pi + u$.

We start this section by showing that, along the Pareto frontier, increasing the surplus of consumers, or decreasing the producer surplus, implies lowering social welfare. To see this, note that, for $1 \leq i \leq m-1$ and $p \in \{x_{i+1}, \cdots, x_m\}$, (3) gives

$$x_i = p \sum_{x \geq p} g_i(x) < x_{i+1} = p \sum_{x \geq p} g_{i+1}(x).$$

Thus

$$\sum_{x \geq p} g_i(x) < \sum_{x \geq p} g_{i+1}(x), \quad \forall p \in \{x_{i+1}, \cdots, x_m\},$$

whence g_{i+1} first-order stochastically dominates g_i. By Shaked and Shanthikumar (2007, Thm. 1.A.8), we conclude that

$$x_i + \overline{u}(x_i) = \sum_{x \geq x_i} x g_i(x) < \sum_{x \geq x_{i+1}} x g_{i+1}(x) = x_{i+1} + \overline{u}(x_{i+1}).$$

In other words, $x_i + \overline{u}(x_i)$ is strictly increasing in i. The previous remark establishes:

Proposition 1. *Along the Pareto frontier, increasing consumer surplus implies lowering social welfare.*

A central insight of Bergemann, Brooks, and Morris (2015) is that, in a single-product setting, market segmentation can be used as a tool to efficiently redistribute the gains from trade. Proposition 1 shows that, contrastingly, when the number of products is large, efficiently transferring surplus from the seller to the consumers through segmentation is infeasible. The broad idea is simple. Whereas in a single-product setting efficiency obtains as long as trade occurs with probability 1, with product variety efficiency also requires each consumer to buy one of the products that he values the most. When product variety is large, the goal of achieving efficiency thus collides with that of inducing low prices. Along the Pareto frontier, the transfer of surplus from seller to consumers is achieved by segmenting the aggregate market in a way that leads the seller to occasionally offer products which do not accurately fit consumers' tastes.

We next examine the link between consumer privacy and welfare. We formalize the notion of privacy by building on Blackwell (1953). Specifically, we say that a market segmentation τ' is *finer* than τ if there exists a function $\xi : \text{supp } \tau \to \Delta\Delta X^n$ such that, for every $\mu \in \text{supp } \tau$,

$$\mu(\mathbf{v}) = \sum_{\tilde{\mu}} \xi(\tilde{\mu} \mid \mu)\tilde{\mu}(\mathbf{v}), \quad \forall \mathbf{v} \in X^n,$$

and

$$\tau'(\tilde{\mu}) = \sum_{\mu} \tau(\mu)\xi(\tilde{\mu} \mid \mu).$$

Intuitively, τ' is obtained by splitting every market μ comprised in the support of τ.[10] The market segmentation τ can thus be viewed as giving greater privacy to consumers than τ'.

Giving consumers greater privacy evidently harms the seller. The effect of consumer privacy on the welfare of consumers is a lot more complex. On the one hand, privacy prevents the seller from extracting surplus through personalized prices. On the other hand, making detailed information available to the seller enables the latter to improve the match quality between consumers and products.

We now argue that if $P' = (\pi', u')$ and $P = (\pi, u)$ are two points on the Pareto frontier such that $u' > u$, then we can "move" from point P to P' by giving consumers greater privacy. Slightly more generally:

Proposition 2. *Let (π, u) and (π', u') be two points in S, with $\pi' < \pi$. For every $n \in \mathbb{N}$, there exist market segmentations τ_n, τ'_n, where τ_n is finer than τ'_n, and strategies $\rho_n \in \text{argmax}_\rho \Pi_{\tau_n}(\rho)$ as well as $\rho'_n \in \text{argmax}_\rho \Pi_{\tau'_n}(\rho)$, such that*

$$\big(\Pi_{\tau_n}(\rho_n), U_{\tau_n}(\rho_n)\big) \underset{n\to\infty}{\longrightarrow} (\pi, u) \quad \text{and} \quad \big(\Pi_{\tau'_n}(\rho'_n), U_{\tau'_n}(\rho'_n)\big) \underset{n\to\infty}{\longrightarrow} (\pi', u').$$

The proof of Proposition 2 rests on two basic ideas. Firstly, different market segmentations typically lead the seller to offer different products. Secondly, for a given market

[10] Interpreting τ and τ' as distributions of posterior beliefs induced by Blackwell–experiments α and α', respectively, our notion corresponds to α' being "sufficient" for α, one of several equivalent definitions of "more informative" in Blackwell (1953).

segmentation, the seller tends to offer the subset of products reagarding which the segmentation is "most informative" (that is, with regard to which the market segmentation best distinguishes consumers). Now suppose $\pi' < \pi$, and we want to find market segmentations τ and τ' respectively generating surplus π and π' for the seller. Proceed as follows. Firstly, partition the products in two subsets of equal size, say K_1 and K_2. Then construct τ' by separating consumers *exclusively* with respect to their valuations for the products in K_1. Under the market segmentation τ', the seller offers products in the subset K_1 and obtains surplus π'. Finally, construct τ by splitting every market in the support of τ' according to consumers' valuations for the products in K_2. Under this finer market segmentation τ, the seller offers products in the subset K_2 and obtains surplus $\pi > \pi'$.

5 Irrelevance of Price Discrimination

We say that a strategy ρ involves price discrimination if some product k is sold at different prices depending on the market in which this product is offered.[11]

It is easy to see that price discrimination may strictly benefit the seller. For example, suppose $X = \{x_1, x_2\}$, let τ be the market segmentation comprising 2^n markets separating consumers with different valuation vectors, and ρ^* some strategy of the seller that is optimal given τ. Now let μ^- denote the market in the support of τ in which every consumer's valuation vector equals $(x_1, \cdots, x_1)$, and μ^k the market in which every consumer values product k at x_2 and all other products at x_1. Then any product offered by the seller in market μ^- must be sold at a price of x_1, whence $\rho^*_{\mu^-}(k, x_1) > 0$ for some product $k \in \{1, \cdots, n\}$. On the other hand, the definition of the market μ^k implies $\rho^*_{\mu^k}(k, x_2) = 1$. So any strategy of the seller that is optimal given τ^* involves price discrimination.

We now argue that price discrimination is irrelevant for the characterization of feasible surplus pairs in Theorem 1: in the limit, when the number of products grows without bound, any surplus pair that is feasible at all is also feasible without price discrimination.

Formally, we say that a surplus pair (π, u) is *feasible without price discrimination* if there exist a market segmentation τ, as well as a strategy $\rho^* \in \operatorname{argmax}_\rho \Pi_\tau(\rho)$, such that

[11] That is, formally, if there exist k, $\mu \neq \mu'$, and $p \neq p'$ such that $\rho_\mu(k, p) > 0$ and $\rho_{\mu'}(k, p') > 0$.

$\pi = \Pi_\tau(\rho^*)$, $u = U_\tau(\rho^*)$, and ρ^* does not price discriminate. The set of surplus pairs that are feasible without price discrimination is denoted by $\tilde{S}_n$.

Proposition 3. *For every $n \in \mathbb{N}$, the set $\tilde{S}_n$ of surplus pairs that are feasible without price discrimination is contained in the set S. Moreover, for every $(\pi, u) \in S$, there exists a sequence $\big((\pi_n, u_n)\big)_{n\in\mathbb{N}}$ such that $(\pi_n, u_n) \in \tilde{S}_n$ and $(\pi_n, u_n) \underset{n\to\infty}{\longrightarrow} (\pi, u)$.*

Consider again the simple binary valuation setting examined above. We illustrate the underlying idea of the proposition by showing that, as product variety becomes large, the surplus pair $(x_2, 0)$ maximizing the surplus of the seller can almost be attained without price discrimination. To this end, let τ be the market segmentation with 2^{n-1} markets separating consumers whose vectors of valuations differ in some other component than the first one, and let μ^- denote the market in the support of τ in which every consumer values all products $k \neq 1$ at x_1. Now let ρ^* be some strategy of the seller such that:

- in market μ^-, the seller offers product 1 at a price of x_1;
- in any other market μ contained in the support of τ, the seller offers at a price of x_2 one of the products $k \neq 1$ to which all consumers in the market μ attach value x_2.

Notice that the strategy ρ^* is optimal given τ,[12] and does not price discriminate. Furthermore, the proportion of consumers who belong to some market $\mu \neq \mu^-$ approaches 1 as n tends to infinity, whence $\big(\Pi_\tau(\rho^*), U_\tau(\rho^*)\big)$ approaches the surplus pair $(x_2, 0)$.

More generally, Proposition 3 shows that product steering makes price discrimination redundant when product variety is large.

6 Online Markets with Data Intermediaries

We have until now examined the implications of market segmentation in very general terms. In particular, we have remained agnostic regarding the source of market segmentation. In this section, we study market segmentation arising from the sale of consumer data by data

[12] In particular, the additional expository assumptions of the baseline model imply for $X = \{x_1, x_2\}$ that $x_1 > x_2 f(x_2)$, whence price x_1 is optimal for product 1 in market μ^-.

intermediaries. By building on the analysis of the previous sections, we determine the implications of data markets for welfare and consumer privacy.

The setting is as follows. There are a seller with an inventory comprising n products, and a unit-demand consumer. The consumer's valuation for product k is denoted by V_k, and the valuation vector by $\mathbf{V} = (V_1, \ldots, V_n)$. These valuations are initially unknown to all parties; the common prior probability assigned to $\mathbf{V} = \mathbf{v}$ is given by $\bar{\mu}(\mathbf{v})$, defined by (1).

The setting also comprises $l \geq 1$ data intermediaries, each of whom chooses a *data policy*, that is, a tuple (D, ϕ) where D is a set of signals and ϕ a mapping

$$\phi : X^n \to \Delta D.$$

Under data policy (D, ϕ), the signal $d \in D$ is drawn with probability $\phi(d \mid \mathbf{v})$ if the consumer has valuation vector $\mathbf{V} = \mathbf{v}$. The signals of different data intermediaries are drawn independently conditional on $\mathbf{V}$.

There are two stages: Stage 1 describes the data market, and Stage 2 describes the product market. The timeline is depicted in Figure 2.

Stage 1 (data market). First, every data intermediary $j = 1, \ldots, l$ chooses a data policy (D_j, ϕ_j), as well as a fee t_j at which it intends to sell the data d_j generated by this policy. Afterwards, the consumer selects a subset of data intermediaries, say $J \subseteq \{1, \ldots, l\}$, comprising all data intermediaries receiving his consent. The seller then purchases data from a subset of data intermediaries J^* selected from the set J.

Stage 2 (product market). First, the valuation vector $\mathbf{V}$ is drawn from the distribution $\bar{\mu}$. Then, the signals of the data intermediaries $j \in J^*$ are drawn according to their data policies. The seller observes these signals, and chooses which product to offer and at what price. Lastly, the consumer decides whether to buy.

The payoff of a data intermediary is its revenue from selling data. If the consumer buys the product offered by the seller, his payoff equals his valuation minus the price; otherwise his payoff is zero. The payoff of the seller equals her revenue minus the cost $\sum_{j \in J^*} t_j$ paid to acquire data.

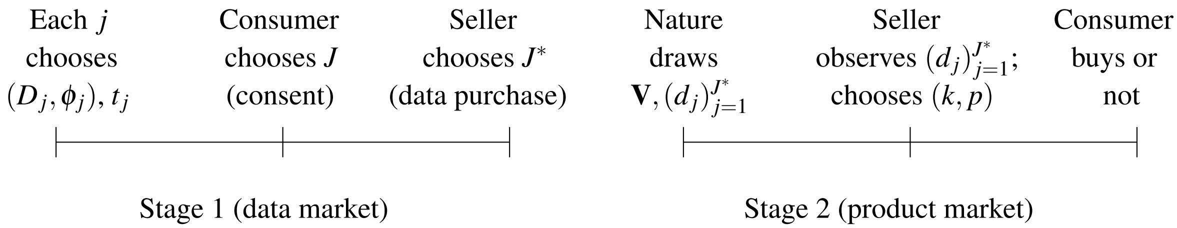


Figure 2: Timeline

The solution concept is perfect Bayesian equilibrium, with two refinements: firstly, the data intermediaries use pure strategies; secondly, the seller breaks ties in favor of the consumer, both when purchasing data and when choosing a product-price combination.

6.1 Discussion of the Model

A key assumption of the model is that the data intermediaries must obtain the consent of the consumer before selling information to the seller. This assumption is consistent with the EU General Data Protection Regulation (Regulation (EU) 2016/679, Article 6), among other things.

Our assumption that the data intermediaries sell information to the seller *directly* accords with the business model of firms such as Acxiom, Nielsen, and Oracle, for example. On the other hand, online platforms acting as data intermediaries, such as Google and Facebook, do not sell information per se, but sell instead access to targeted consumer segments. To keep the analysis focused, we disregard in this paper the distinction between direct and indirect sale of information.[13]

The assumption that the data intermediaries know precisely the consumer's valuation vector evidently lacks realism, and merely ensures tractability. In particular, in practice one of the gains from having multiple data intermediaries may be that different intermediaries possess complementary information about consumers' preferences. Such considerations are beyond the scope of our analysis.

Finally, the model makes a number of technical assumptions. The assumption that a data

[13]See Federal Trade Commission (2014) and Bergemann and Bonatti (2019) for details about the various business models of data intermediaries.

intermediary simultaneously chooses its data policy and the fee at which it intends to sell its data simplifies the structure of the game, but is irrelevant for our results. Our assumption that the seller breaks ties in favor of the consumer ensures that each data market outcome at the end of Stage 1 induces both a unique expected revenue for the seller and a unique expected payoff for the consumer. Our focus on pure strategies circumvents possible miscoordination among data intermediaries.

6.2 Data Intermediation and Welfare

We now characterize the equilibrium payoffs when the number of products is large.[14]

Define

$$u_0 := \sum_{x \geq p_0} f(x)(x-p);$$

intuitively, u_0 captures the consumer surplus resulting from the market segmentation attaching mass 1 to the aggregate market. Let also

$$\pi_A := \max\left\{\pi \in [\pi_0, x_m] \mid \overline{u}(\pi) = u_0\right\}.$$

To simplify the statement of the next proposition, we assume that the function $\overline{u}$ possesses a unique maximizer in $[\pi_0, x_m]$,[15] which we denote by π_B. Figure 3 illustrates the points $A = \big(\pi_A, \overline{u}(\pi_A)\big)$ and $B = \big(\pi_B, \overline{u}(\pi_B)\big)$ in an example where $X = \{1,2,3,4,5,6,7\}$ and $\pi_0 = 2$.

Proposition 4. *For every $n \in \mathbb{N}$, fix some equilibrium. Let $(\pi_n, u_n)_{n\in\mathbb{N}}$ be the corresponding combinations of expected revenue of the seller and expected payoff of the consumer. If $l = 1$, then $(\pi_n, u_n) \underset{n\to\infty}{\longrightarrow} \big(\pi_A, \overline{u}(\pi_A)\big)$; if $l > 1$, then $(\pi_n, u_n) \underset{n\to\infty}{\longrightarrow} \big(\pi_B, \overline{u}(\pi_B)\big)$.*

The proposition can be understood as follows. A monopolistic data intermediary ($l = 1$) fully extracts the seller's gain from purchasing data. This results in a data policy that maximizes the seller's expected revenue, subject to the constraint that the consumer gives his consent. By contrast, when the data market is competitive ($l > 1$), the implemented

[14] We omit a proof of the existence of an equilibrium. Lemma B1 in Section B of the Appendix implies that both consumer-optimal and seller-optimal data policies exist. Based on this, it is straightforward to deduce existence of an equilibrium.

[15] This is the case if $\sum_{x\geq x_i} g_i(x)(x-x_i) \neq \sum_{x\geq x_{i+1}} g_{i+1}(x)(x-x_{i+1})$ for all $i = 1,\ldots,m-1$.

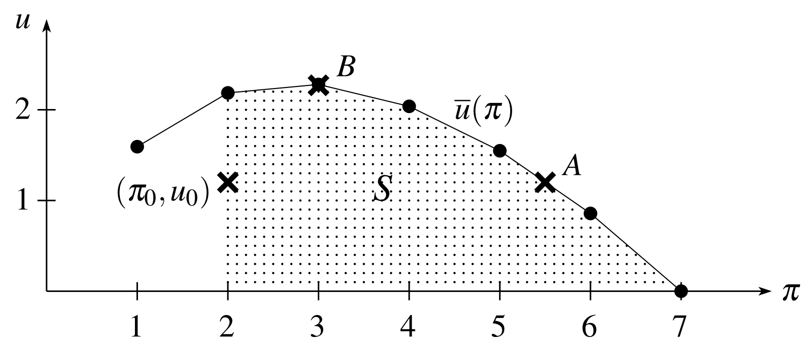


Figure 3: The welfare consequences of competition in the data market

payoff pair maximizes the consumer's expected payoff, subject to the constraint that the seller purchases the data.

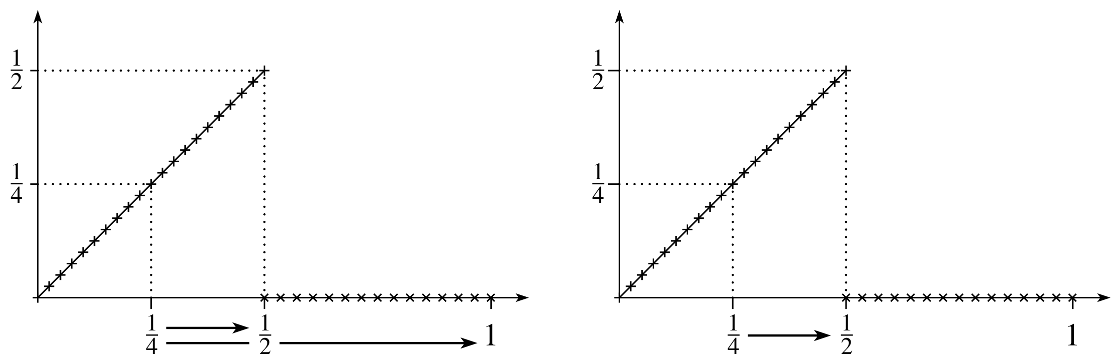
We sketch here the main ideas of the proof. Every data market outcome at the end of Stage 1 induces a market segmentation τ. In the current setting, $\tau(\mu)$ represents the probability that the seller's posterior belief concerning the consumer's valuations (after observing the signals) is equal to μ. Thus, every data market outcome induces an expected revenue π_n for the seller, and an expected payoff u_n for the consumer, such that (π_n, u_n) belongs to the set S_n of feasible surplus pairs defined in Section 2.[16] We then prove that S_n satisfies a number of properties which enable us to pin down both the expected revenue of the seller and the expected payoff of the consumer in any equilibrium. Finally, an application of Theorem 1 yields the convergences stated in Proposition 4.

Note that combining Propositions 1 and 4 shows that competition in the data market benefits consumers at the cost of a loss in social welfare. Specifically, while competition in the data market helps to reduce prices in the product market, it also results in lower match quality between consumers and products.

6.3 Data Intermediation and Consumer Privacy

We saw in the previous subsection that competition in the data market increased consumer surplus along the Pareto frontier (Proposition 4). On the other hand, we saw in Section 4 that, along the Pareto frontier, greater consumer surplus was compatible with more privacy

[16] Specifically, (π_n, u_n) belongs to the subset of S_n that consists of the surplus pairs which are consistent with the seller breaking ties in favor of the consumer.



(a) Monopoly ($l = 1$)

(b) Competition ($l > 1$)

Figure 4: Illustration of equilibrium data policies, assuming $x_1 = 1$, $x_2 = 2$, $f(x_2) = 1/4$. Horizontal axis: probability $\Pr[v_{k^*} = x_2]$; vertical axis: consumer's expected payoff.

(Proposition 2). We now combine these insights to show that competition in the data market ultimately benefits consumers by giving them more privacy. We assume here that valuations are binary ($X = \{x_1, x_2\}$), which will allow us to characterize equilibrium data policies.

We first illustrate the general idea by way of an example in which $x_1 = 1$, $x_2 = 2$, and $f(x_2) = 1/4$. As $x_1 > x_2 f(x_2)$, in the absence of any data the seller offers an arbitrary product at a price of x_1. The probability that the consumer's valuation for the offered product is equal to x_2 is then $f(x_2)$. The seller thus obtains a payoff of 1, while the consumer obtains an expected payoff equal to

$$f(x_2)(x_2 - x_1) = \frac{1}{4}.$$

Figure 4, Panel (a), illustrates the equilibrium in the case of a monopolistic data intermediary. In order to enable the seller to increase her revenue, the data intermediary identifies, with probability $1/2$, a product for which the consumer's valuation is x_2. With the remaining probability, the data intermediary identifies a product to which the consumer is equally likely to attach value x_1 or x_2. In the latter case, the seller offers said product at a price of x_1, leaving the consumer with a surplus of $(x_2 - x_1)$ every time his valuation is x_2. The consumer thus obtains an expected payoff of

$$\frac{1}{2} \cdot \frac{1}{2}(x_2 - x_1) = \frac{1}{4},$$

making him indifferent between consenting to the policy or not. The seller, on the other hand, obtains an expected revenue equal to

$$\frac{1}{2}x_1 + \frac{1}{2}x_2 = \frac{3}{2},$$

allowing the monopolistic data intermediary to extract $1/2$ through the ex ante fee.

Figure 4, Panel (b), illustrates the equilibrium in the case of a competitive data market. Competition among data intermediaries results in the consumer's most-preferred data policy. This data policy identifies with probability 1 a product to which the consumer is equally likely to attach value x_1 or x_2 (and, thus, transmits less information than the data policy of a monopolistic data intermediary). The seller offers said product at a price of x_1, giving the consumer an expected payoff of

$$\frac{1}{2}(x_2 - x_1) = \frac{1}{2}.$$

In order to generalize the previous example, we rank the privacy afforded by different data policies according to Blackwell–informativeness. A data policy (D, ϕ) is *more informative* than another data policy (D', ϕ') if there exists a function $\sigma : D \to \Delta D'$ such that

$$\phi'(d' \mid \mathbf{v}) = \sum_{d \in D} \phi(d \mid \mathbf{v})\sigma(d' \mid d), \quad \forall d' \in D', \forall \mathbf{v} \in X^n.$$

Thus, (D', ϕ') is a garbled version of (D, ϕ). Below, say that the seller purchases data given by (D, ϕ) if in Stage 1 the seller purchases data from a single data intermediary, say j, and $(D_j, \phi_j) = (D, \phi)$.

Proposition 5. *Let $X = \{x_1, x_2\}$. Fix $l' \in \mathbb{N}$ with $l' > 1$, and $n \in \mathbb{N}$ with $n \geq (\ln f(x_2) + \ln(x_2 - x_1) - \ln x_1)/\ln f(x_1)$. There exist two data policies, (D, ϕ) and (D', ϕ'), as well as an equilibrium for $l = 1$ and an equilibrium for $l = l'$, such that:*

- *(D, ϕ) is more informative than (D', ϕ');*
- *in the equilibrium for $l = 1$, the seller purchases data given by (D, ϕ);*
- *in the equilibrium for $l = l'$, the seller purchases data given by (D', ϕ');*
- *the expected payoff of the consumer is greater in the equilibrium for $l = l'$ than in the equilibrium for $l = 1$.*

Proposition 5 thus formalizes the idea that competition in the data market benefits consumers by giving them more privacy.

7 Conclusion

We have studied the welfare consequences of market segmentation in a multi-product monopoly. Our central result is an approximation of the feasible producer-consumer surplus pairs when product variety is large. This characterization revealed a trade-off between consumer surplus and social welfare. We also showed that, when product variety is large, price discrimination has no role to play. Specifically, product steering makes price discrimination redundant. Finally, we have shown that greater consumer surplus is compatible with more privacy. In particular, along the Pareto frontier, increasing consumer surplus can be achieved by giving consumers greater privacy. We applied our results to study market segmentation arising from the sale of consumer data by data intermediaries, and showed that competition in the data market can give consumers both greater surplus and greater privacy.

Appendix

A Omitted Proofs for Sections 3–5

Proof of Lemma 2. Since τ is the product of $(\tau^k)_{k=1}^n$, notice that $\mu \in \text{supp } \tau$ if and only if (6) holds and $\mu_k \in \text{supp } \tau^k$ for every k. Then, using (6), (7), and (8) gives

$$\begin{aligned}\sum_{\mu} \tau(\mu)\mu(\mathbf{v}) &= \sum_{\substack{\mu_1 \in \text{supp } \tau^1, \\ \cdots, \\ \mu_n \in \text{supp } \tau^n}} \prod_k \tau^k(\mu_k)\mu_k(v_k) \\ &= \prod_k \left(\sum_{\mu_k \in \text{supp } \tau^k} \tau^k(\mu_k)\mu_k(v_k) \right) \\ &= \prod_k f(v_k) = \bar{\mu}(\mathbf{v}),\end{aligned}$$

for all $\mathbf{v} \in X^n$. □

Proof of Proposition 2. We will use the following lemma. Its proof is analogous to the proof of Lemma 2, and therefore omitted.

Lemma A1. *For every $k \in \{1, \dots, n\}$, let $\tau^k \in \Delta X$ and $\xi^k : \Delta X \to \Delta\Delta X$ satisfy*

$$\sum_{\mu_k} \tau^k(\mu_k)\mu_k(x) = f(x), \quad \forall x \in X,$$

$$\sum_{\tilde{\mu}_k} \xi^k(\tilde{\mu}_k \mid \mu_k)\tilde{\mu}_k(x) = \mu_k(x), \quad \forall x \in X, \forall \mu_k \in \Delta X.$$

Define $\langle \tau^k, \xi^k \rangle \in \Delta X$ by

$$\langle \tau^k, \xi^k \rangle(\tilde{\mu}_k) = \sum_{\mu_k} \tau^k(\mu_k)\xi^k(\tilde{\mu}_k \mid \mu_k), \quad \forall \tilde{\mu}_k \in \Delta X. \tag{A.1}$$

Let τ be the product of $(\tau^k)_{k=1}^n$, and $\hat{\tau}$ be the product of $(\langle \tau^k, \xi^k \rangle)_{k=1}^n$. Then both τ and $\hat{\tau}$ are market segmentations, and $\hat{\tau}$ is finer than τ.

We can now prove the proposition. We treat below the case $(\pi, u) = (x_i, \overline{u}(x_i))$ and $(\pi', u') = (x_j, \overline{u}(x_j))$ with $x_i, x_j \in X$ and $x_i < x_j$; the proof for the remaining cases is similar, and therefore omitted.

Let $\lambda_i \in (0,1)$, and define $h_i \in \Delta X$ by

$$h_i(x) := \frac{f(x) - \lambda_i g_i(x)}{1 - \lambda_i},$$

as in the proof of Lemma 3. Moreover, let $\lambda_j \in (0,1)$, and define $h_j \in \Delta X$ analogously.

Next, for $n > 1$, define $\tau_n^k \in \Delta\Delta X$ by

$$\tau_n^k(g_i) = \lambda_i = 1 - \tau_n^k(h_i), \quad \forall k \in \{1, \ldots, n \div 2\}$$

(where $\div$ denotes division with remainder), and

$$\tau_n^k(f) = 1, \quad \forall k \in \{(n \div 2) + 1, \ldots, n\}.$$

Moreover, define $\xi_n^k : \Delta X \to \Delta\Delta X$ by

$$\xi_n^k(g_j \mid f) = \lambda_j = 1 - \xi_n^k(h_j \mid f), \quad \forall k \in \{(n \div 2) + 1, \ldots, n\},$$

and $\xi_n^k(\mu_k \mid \mu_k) = 1$ if $k \in \{1, \ldots, n \div 2\}$ or $\mu_k \neq f$. Lastly, let τ_n be the product of $(\tau_n^k)_{k=1}^n$, and $\hat{\tau}_n$ the product of $(\langle \tau_n^k, \xi_n^k \rangle)_{k=1}^n$, where $\langle \tau_n^k, \xi_n^k \rangle$ was defined in (A.1). By Lemma A1, both τ_n and $\hat{\tau}_n$ are market segmentations, and $\hat{\tau}_n$ is finer than τ_n.

Every market μ in the support of τ_n satisfies $\mu_k \in \{g_i, h_i\}$ for all $k \in \{1, \ldots, n \div 2\}$, and $\mu_k = f$ for all $k \in \{(n \div 2) + 1, \ldots, n\}$. By (12), there exists a strategy $\rho_n \in \operatorname{argmax}_\rho \Pi_{\tau_n}(\rho)$ with the following property:

> For every market $\mu \in \operatorname{supp} \tau_n$ satisfying $\mu_k = g_i$ for some $k \in \{1, \ldots, n \div 2\}$, offer a product k for which $\mu_k = g_i$ at a price of x_i.

Because the probability that $\mu_k = g_i$ for some $k \in \{1, \ldots, n \div 2\}$ is $1 - (1 - \lambda_i)^{n \div 2}$, which tends to 1 as n grows without bound, it holds that

$$\lim_{n \to \infty} \Pi_{\tau_n}(\rho_n) = \lim_{n \to \infty} (1 - (1 - \lambda_i)^{n \div 2}) x_i,$$
$$\lim_{n \to \infty} U_{\tau_n}(\rho_n) = \lim_{n \to \infty} (1 - (1 - \lambda_i)^{n \div 2}) \sum_{x \geq x_i} g_i(x)(x - x_i).$$

Consequently,

$$\lim_{n \to \infty} \big(\Pi_{\tau_n}(\rho_n), U_{\tau_n}(\rho_n)\big) = \big(x_i, \bar{u}(x_i)\big).$$

Every market μ in the support of $\hat{\tau}_n$ satisfies $\mu_k \in \{g_i, h_i\}$ for all $k \in \{1, \ldots, n \div 2\}$, and $\mu_k \in \{g_j, h_j\}$ for all $k \in \{(n \div 2) + 1, \ldots, n\}$. By (12), and since $x_j > x_i$, there exists a strategy $\hat{\rho}_n \in \operatorname{argmax}_\rho \Pi_{\hat{\tau}_n}(\rho)$ with the following property:

For every market $\mu \in \operatorname{supp} \hat{\tau}_n$ satisfying $\mu_k = g_j$ for some $k \in \{(n \div 2)+1, \ldots, n\}$, offer a product k for which $\mu_k = g_j$ at a price of x_j.

Then, as above,

$$\lim_{n\to\infty} \big(\Pi_{\hat{\tau}_n}(\hat{\rho}_n), U_{\hat{\tau}_n}(\hat{\rho}_n)\big) = \big(x_j, \overline{u}(x_j)\big). \qquad \square$$

Proof of Proposition 3. The first part of the proposition follows from Theorem 1 because $\tilde{S}_n \subseteq S_n$.

We next prove the second part. We treat below the case $\pi = x_i \in X$ and $u = \zeta \overline{u}(x_i)$, where $\zeta \in [0,1]$; the proof for the remaining case is similar, and therefore omitted.

Let $\lambda \in (0,1)$, and let $h \in \Delta X$ be given by (10). We claim that

$$x_i \geq \pi_0 \geq \max_p \; p \sum_{x \geq p} h(x). \tag{A.2}$$

The first inequality holds because $(\pi, u) \in S$. We next show the second inequality. For all $p \leq x_i$,

$$\pi_0 \geq p \sum_{x\geq p} f(x) = \lambda p \sum_{x \geq p} g_i(x) + (1-\lambda) p \sum_{x\geq p} h(x) = \lambda p + (1-\lambda) p \sum_{x \geq p} h(x).$$

Hence, $(1-\lambda) p \sum_{x\geq p} f(x) \geq p \sum_{x \geq p} f(x) - \lambda p = (1-\lambda) p \sum_{x\geq p} h(x)$. For all $p > x_i$,

$$\pi_0 \geq p \sum_{x\geq p} f(x) = \lambda p \sum_{x \geq p} g_i(x) + (1-\lambda) p \sum_{x\geq p} h(x) = \lambda x_i + (1-\lambda) p \sum_{x \geq p} h(x).$$

In either case, we obtain $\pi_0 \geq p \sum_{x \geq p} h(x)$, which finishes to prove (A.2).

Next, let K^{x_i} and K^{x_m} be two disjoint subsets of $\{1, \cdots, n\}$, each containing $(n-1) \div 2$ elements (where $\div$ denotes division with remainder). Let $k_0 \in \{1, \cdots, n\} \setminus (K^{x_i} \cup K^{x_m})$. For each $p \in \{x_i, x_m\}$, define $\tau_k^p \in \Delta\Delta X$ by

$$\tau_k^p(g_i) = \lambda = 1 - \tau_k^p(h), \quad \forall k \in K^p,$$

and

$$\tau_k^p(f) = 1, \quad \forall k \in \{1, \ldots, n\} \setminus K^p.$$

Lastly, let τ^p be the product of $(\tau_k^p)_{k=1}^n$. By coupling (11) with Lemma 2, notice that τ^p is a market segmentation. Consequently, the mixture

$$\tau := \zeta \tau^{x_i} + (1-\zeta) \tau^{x_m}$$

is a market segmentation too.

Each market μ in the support of τ satisfies exactly one of the following conditions:

(a) $\mu_k = g_i$ for some $k \in K^{x_i}$;

(b) $\mu_k = g_i$ for some $k \in K^{x_m}$;

(c) $\mu_k \in \{h, f\}$ for all $k \in K^{x_i} \cup K^{x_m}$.

Now let ρ be a strategy of the seller with the following properties:

- for every market $\mu \in \operatorname{supp} \tau$ satisfying (a), offer one of the products $k \in K^{x_i}$ for which $\mu_k = g_i$ at a price of x_i;
- for every market $\mu \in \operatorname{supp} \tau$ satisfying (b), offer one of the products $k \in K^{x_m}$ for which $\mu_k = g_i$ at a price of x_m;
- for every market $\mu \in \operatorname{supp} \tau$ satisfying (c), offer product k_0 at price p_0.

By (A.2), the strategy ρ is optimal given τ. Furthermore, note that ρ does not price discriminate: the products in K^{x_i} (respectively, K^{x_m}) are always offered at a price of x_i (respectively, x^m), and product k_0 is always offered at price p_0.

The resulting surplus of the seller is

$$\pi_n := \Pi_\tau(\rho) = \big(1 - (1-\lambda)^{(n-1)\div 2}\big)x_i + (1-\lambda)^{(n-1)\div 2}\pi_0;$$

the consumer surplus is

$$u_n := U_\tau(\rho) = \zeta\big(1 - (1-\lambda)^{(n-1)\div 2}\big)\sum_{x \geq x_i} g_i(x)(x - x_i) + (1-\lambda)^{(n-1)\div 2}\sum_{x \geq p_0} f(x)(x - p_0).$$

So $(\pi_n, u_n) \in \tilde{S}_n$, and as $\lambda > 0$,

$$(\pi_n, u_n) \underset{n\to\infty}{\longrightarrow} \big(x_i, \zeta\bar{u}(x_i)\big). \qquad \square$$

B Proofs for Section 6

The proofs of Propositions 4 and 5 build on a characterization of the expected revenues of the seller and the expected payoffs of the consumer that can result from arbitrary data policies. We state this characterization in the following subsection.

B.1 Preliminaries

Every data policy induces a distribution $\tau \in \Delta\Delta X^n$ of posterior beliefs $\mu \in \Delta X^n$; furthermore, this τ is a market segmentation. Conversely: every market segmentation τ is the distribution of posterior beliefs induced by some data policy (see Kamenica and Gentzkow, 2011).

Next, the profile of data policies $\big((D_j, \phi_j)\big)_{j \in J^*}$ has the same informational content as the "aggregate" data policy (D', ϕ') with $D' = \prod_{j \in J} D_j$ and ϕ' given by

$$\phi'\big((d_j)_{j \in J^*} \mid \mathbf{v}\big) = \prod_{j \in J} \phi_j(d_j \mid \mathbf{v}).$$

Hence, we can represent any profile of data policies by a single data policy.

Combining the previous observations shows that every data market outcome at the end of Stage 1 induces a subgame in Stage 2 in which the seller obtains an expected revenue π_n and the consumer obtains an expected payoff u_n such that (π_n, u_n) belongs to the set S_n of feasible surplus pairs defined in Section 2. In fact, since here the seller breaks ties in favor of the consumer, the previous payoffs are uniquely pinned down by the aggregate data policy (D', ϕ'). We thus say that (D', ϕ') *implements* (π_n, u_n). The set of pairs (π, u) which can be implemented by some data policy (D', ϕ') will be denoted by $\hat{S}_n$.

The following two lemmas provide a characterization of the set $\hat{S}_n$. Their proofs are relegated to Section OA.1 of the Online Appendix.

Lemma B1. *Define* $\pi_n^{\max} := \sum_{\mathbf{v}} \bar{\mu}(\mathbf{v}) \max_k v_k$. *For every* $n \in \mathbb{N}$, *it holds that:*

a) $\{\pi \in \mathbb{R} \mid$ *there exists* $u \in \mathbb{R}$ *s.t.* $(\pi, u) \in \hat{S}_n\} = [\pi_0, \pi_n^{\max}]$*;*

b) $\{u \in \mathbb{R} \mid (\pi, u) \in \hat{S}_n\}$ *has a greatest element* $\bar{u}_n(\pi)$ *for every* $\pi \in [\pi_0, \pi_n^{\max}]$*;*

c) $\bar{u}_n : [\pi_0, \pi_n^{\max}] \to \mathbb{R}$ *is concave and continuous, and* $\bar{u}_n(\pi_n^{\max}) = 0$.

Lemma B2. *For every* $n \in \mathbb{N}$, $\hat{S}_n \subseteq S$. *Moreover, for every* $(\pi, \bar{u}(\pi)) \in S$, *there exists a sequence* $\big((\pi_n, u_n)\big)_{n \in \mathbb{N}}$ *such that* $(\pi_n, u_n) \in \hat{S}_n$ *and* $(\pi_n, u_n) \underset{n \to \infty}{\longrightarrow} (\pi, \bar{u}(\pi))$.

B.2 Proofs of Propositions 4 and 5

Proof of Proposition 4. We abbreviate "data intermediary" to "DI".

(i) Suppose there is a single DI. Fix some $n \in \mathbb{N}$. In every perfect Bayesian equilibrium, the consumer's expected payoff is at least u_0, which the consumer obtains if he does not give consent to the proposed data policy. Hence, the seller's expected gross payoff is at most

$$\max\left\{\pi \in \mathbb{R} \mid \text{there exists } u \geq u_0 \text{ s.t. } (\pi, u) \in \hat{S}_n\right\}.$$

By Lemma B1, the maximum exists and is equal to

$$\pi_n^* := \max\left\{\pi \in [\pi_0, \pi_n^{\max}] \mid \overline{u}_n(\pi) = u_0\right\}.$$

In every perfect Bayesian equilibrium, the DI chooses fee $\pi_n^* - \pi_0$ and a data policy that implements (π_n^*, u_0), the consumer consents, and the seller purchases the data.

It remains to show that $\lim_{n\to\infty} \pi_n^* = \pi_A$. There are two cases. Case (i): π_A coincides with π_B, the unique maximizer of $\overline{u}$. Then $\overline{u}(\pi_B) = \overline{u}(\pi_A) = u_0$, which implies $\pi_n^* = \pi_B = \pi_A = \pi_0$ for all n and hence $\lim_{n\to\infty} \pi_n^* = \pi_A$. Case (ii): $\pi_A \neq \pi_B$. Then $\overline{u}$ is strictly decreasing at π_A. By the second part of Lemma B2, we can find $n \in \mathbb{N}$ such that $(\pi, u) \in \hat{S}_n$ with π in any neighborhood of π_A and $u > \overline{u}(\pi_A) = u_0$. Consequently, $\liminf_{n\to\infty} \pi_n^* \geq \pi_A$. On the other hand, $\limsup_{n\to\infty} \pi_n^* \leq \pi_A$ because $\hat{S}_n \subseteq H$ by the first part of Lemma B2. Thus, $\lim_{n\to\infty} \pi_n^* = \pi_A$

(ii) Suppose there is more than one DI. Fix some $n \in \mathbb{N}$. First, we show that the consumer's expected payoff in every perfect Bayesian equilibrium is equal to

$$\max\left\{u \in \mathbb{R} \mid \text{there exists } \pi \in \mathbb{R} \text{ s.t. } (\pi, u) \in \hat{S}_n\right\}.$$

By Lemma B1, the maximum exists and is

$$u_n^* := \max\left\{\overline{u}_n(\pi) \mid \pi \in [\pi_0, \pi_n^{\max}]\right\}.$$

By contradiction, suppose the consumer's expected payoff in some perfect Bayesian equilibrium is $u < u_n^*$. Let π be the seller's expected revenue in this equilibrium. Then, the sum of the expected payoffs of the DIs is at most $\pi - \pi_0$. Consequently, there exists a DI j whose expected payoff is strictly smaller than $(\pi - \pi_0)/2$. By Lemma B1, part c), we can find $\left(\pi', \overline{u}_n(\pi')\right) \in \hat{S}_n$ with $\pi' > (\pi + \pi_0)/2$ and $\overline{u}_n(\pi') > u$. Suppose DI j chooses fee $\pi' - \pi_0$ and a data policy that implements $\left(\pi', \overline{u}_n(\pi')\right)$. We show that the consumer would give consent to j, and the seller would purchase j's data.

Indeed, if the consumer gives consent to DI j alone, then the seller must purchase j's data, given our equilibrium restriction that the seller breaks ties in favor of the consumer when purchasing data. The consumer's expected payoff is then $\overline{u}_n(\pi') > u$. Suppose the consumer gives consent to a subset of DIs not including j. This choice was also possible when j did not deviate, and it would have resulted in the same expected payoff for the consumer. Hence, the consumer's expected payoff is at most $u < \overline{u}_n(\pi')$ in this case. Suppose finally the consumer gives consent to a subset J of DIs that includes j. If the seller then does not purchase j's data, the consumer's expected payoff is the same as if he gives consent just to the DIs $J \setminus \{j\}$.

Thus, if DI j chooses chooses fee $\pi' - \pi_0$ and a data policy that implements $\left(\pi', \overline{u}_n(\pi')\right)$, the consumer consents and the seller purchases the data. With this deviation, j's payoff is $\pi' - \pi_0 > (\pi - \pi_0)/2$, contradicting the hypothesis that j earns $(\pi - \pi_0)/2$ in equilibrium. Hence, the consumer obtains $u = u_n^*$ in every perfect Bayesian equilibrium.

Choose $\hat{\pi}_n \in [\pi_0, \pi_n^{\max}]$ such that $\overline{u}_n(\hat{\pi}_n) = u_n^*$. We show $\lim_{n\to\infty}(\hat{\pi}_n, \overline{u}(\hat{\pi}_n)) = (\pi_B, \overline{u}(\pi_B))$. Because π_B is the unique maximizer of $\overline{u}$, $\overline{u}_n(\hat{\pi}_n) \leq \overline{u}(\pi_B)$ by the first part of Lemma B2. Hence, $\limsup_{n\to\infty} \overline{u}_n(\hat{\pi}_n) \leq \overline{u}(\pi_B)$. By the second part of Lemma B2, $\liminf_{n\to\infty} \overline{u}_n(\hat{\pi}_n) \geq \overline{u}(\pi_B)$. Thus, $\lim_{n\to\infty} \overline{u}_n(\hat{\pi}_n) = \overline{u}(\pi_B)$. By contradiction, suppose $\liminf_{n\to\infty} \hat{\pi}_n = \pi' < \pi_B$ or $\limsup_{n\to\infty} \hat{\pi}_n = \pi' > \pi_A$. Then, $\overline{u}_n(\pi') \leq \overline{u}(\pi') < \overline{u}(\pi_B)$, contradicting $\lim_{n\to\infty} \overline{u}_n(\hat{\pi}_n) = \overline{u}(\pi_B)$. Thus, $\lim_{n\to\infty} \hat{\pi}_n = \pi_B$. □

Proof of Proposition 5. Note that for $X = \{x_1, x_2\}$, the assumption $p_0 \in X$ implies $\pi_0 = p_0 = x_1$. Thus, $x_1 > f(x_2)x_2$. Moreover, $u_0 = f(x_2)(x_2 - x_1)$, and

$$\overline{u}(\pi) = \frac{x_2 - \pi}{x_2 - x_1} \frac{x_1}{x_2}(x_2 - x_1), \quad \forall \pi \in [x_1, x_2]. \tag{B.1}$$

Lastly, the hypothesis

$$n \geq \frac{\ln f(x_2) + \ln(x_2 - x_1) - \ln x_1}{\ln f(x_1)}$$

implies

$$(f(x_1))^n \leq f(x_2) \frac{x_2 - x_1}{x_1}. \tag{B.2}$$

The proof consists of three steps.

Step 1: an equilibrium for $l = 1$. Suppose $l = 1$. If the data intermediary proposes a data policy that implements (π, u), it is optimal for the consumer to give his consent if $u \geq u_0$, and it is optimal for the seller to purchase the data if the fee is at most $\pi - \pi_0$.

By Lemma B2, the set of (π, u) that can be implemented by some data policy is comprised in the set S. Furthermore, $\left(\pi_A, \overline{u}(\pi_A)\right)$ maximizes the seller's expected revenue among all points in S giving the consumer an expected payoff at least as large as u_0. Therefore, if some data policy (D, ϕ) implements $\left(\pi_A, \overline{u}(\pi_A)\right)$, then an equilibrium exists in which the data intermediary proposes (D, ϕ) at fee $\pi_A - \pi_0$, the consumer gives his consent, and the seller purchases the data.

We now show that some data policy implements $\left(\pi_A, \overline{u}(\pi_A)\right)$. For future reference, note that (B.1) combined with $\overline{u}(\pi_A) = u_0$ yield

$$\pi_A = x_2 - f(x_2)\frac{x_2 - x_1}{x_1}x_2.$$

Consider the data policy (D, ϕ), where $D = \{1, \ldots, n\} \times X$ and ϕ is defined as follows. With probability

$$\frac{\pi_A/x_2 - f(x_2)}{1 - (f(x_1))^n - f(x_2)},$$

the first component of the signal is drawn uniformly at random from the set $\operatorname{argmax}_{k'} v_{k'}$; by (B.2), this number is between zero and one. With the remaining probability, the first component of the signal is drawn uniformly at random from $\{1, \ldots, n\}$. Thus, the first component displays with some noise a product for which the consumer has the highest valuation across all products. Based on this information, the posterior probability that $v_k = x_2$ when k is displayed is

$$\frac{\pi_A/x_2 - f(x_2)}{1 - (f(x_1))^n - f(x_2)}(1 - (f(x_1))^n) + \left(1 - \frac{\pi_A/x_2 - f(x_2)}{1 - (f(x_1))^n - f(x_2)}\right) f(x_2) = \frac{\pi_A}{x_2}.$$

If the first component of the signal is k and $v_k = x_2$, then the second component is $x = x_2$ with probability

$$\frac{\pi_A - x_1}{x_2 - x_1}\frac{x_2}{\pi_A},$$

and $x = x_1$ with the remaining probability. If $v_k = x_1$, on the other hand, then $x = x_1$ with probability one. Thus, the second component potentially reveals the consumer's valuation for the displayed product k if the valuation is equal to x_2. The posterior probability that $v_k = x_2$ based on the two components of the signal is one if $x = x_2$, and

$$\frac{\pi_A}{x_2}\left(1 - \frac{\pi_A - x_1}{x_2 - x_1}\frac{x_2}{\pi_A}\right) \Big/ \left(1 - \frac{\pi_A}{x_2}\frac{\pi_A - x_1}{x_2 - x_1}\frac{x_2}{\pi_A}\right) = \frac{x_1}{x_2}$$

if $x = x_1$.

After signal $d = (k, x_2)$, the seller knows for sure that she can sell product k at price x_2; doing this is optimal. After signal $d = (k, x_1)$, the posterior probability that $v_k = x_2$ is x_1/x_2. For any product $k' \neq k$, by contrast, the posterior probability that $v_{k'} = x_2$ is bounded by the prior probability $f(x_2) < x_1/x_2$ because k' may not belong to the products for which the consumer's valuation is the highest among all products. Hence, it is again optimal to offer product k. Furthermore, x_1 is an optimal price because $x_1 = x_1/x_2 \cdot x_2$.

The signals $d \in \{(1, x_1), \ldots, (n, x_1)\}$ have total probability

$$1 - \frac{\pi_A}{x_2} \frac{\pi_A - x_1}{x_2 - x_1} \frac{x_2}{\pi_A} = \frac{x_2 - \pi_A}{x_2 - x_1}.$$

Consequently, the expected revenue of the seller is

$$\frac{x_2 - \pi_A}{x_2 - x_1} x_1 + \left(1 - \frac{x_2 - \pi_A}{x_2 - x_1}\right) x_2 = \pi_A,$$

and the expected payoff of the consumer is

$$\frac{x_2 - \pi_A}{x_2 - x_1} \frac{x_1}{x_2} (x_2 - x_1) = \overline{u}(\pi_A).$$

Thus, (D, ϕ) implements $\big(\pi_A, \overline{u}(\pi_A)\big)$.

Step 2: an equilibrium for $l = l'$. Suppose $l = l'$. In S, the consumer's expected payoff is maximized at $\big(\pi_B, \overline{u}(\pi_B)\big) = \big(x_1, \overline{u}(x_1)\big)$. Suppose there exists a data policy (D', ϕ') that implements $\big(x_1, \overline{u}(x_1)\big)$. Then, there exists an equilibrium in which every data intermediary proposes (D', ϕ'), along with a fee of zero, the consumer gives his consent to exactly one intermediary, selected uniformly at random, and the seller purchases its data. To see this, note that the seller is indifferent whether to purchase the data, and the consumer cannot benefit by giving his consent to more than one intermediary. If an intermediary deviates to a data policy that does not implement $\big(x_1, \overline{u}(x_1)\big)$, it would not be optimal for the consumer to give his consent to this intermediary, and at any fee strictly greater than zero the seller would not purchase the data.

We present a data policy that implements $\big(x_1, \overline{u}(x_1)\big)$. Consider the data policy (D', ϕ'), where $D' = \{1, \ldots, n\}$. With probability

$$\frac{x_1/x_2 - f(x_2)}{1 - (f(x_1))^n - f(x_2)},$$

the signal is drawn uniformly at random from the set $\text{argmax}_{k'} v_{k'}$; this is a number between zero and one by (B.2) and because $x_1 > f(x_2)x_2$. With the remaining probability, the signal is drawn uniformly at random from $\{1, \cdots, n\}$.

When (D', ϕ') displays signal $d' = k$, the posterior probability that $v_k = x_2$ is

$$\frac{x_1/x_2 - f(x_2)}{1 - (f(x_1))^n - f(x_2)}(1 - (f(x_1))^n) + \left(1 - \frac{x_1/x_2 - f(x_2)}{1 - (f(x_1))^n - f(x_2)}\right) f(x_2) = \frac{x_1}{x_2}.$$

For any product $k' \neq k$, by contrast, the posterior probability that $v_{k'} = x_2$ is bounded by the prior probability $f(x_2) < x_1/x_2$ because k' may not belong to the products for which the consumer's valuation is the highest among all products. Hence, it is optimal to offer product k, and x_1 is an optimal price.

Consequently, the expected revenue of the seller is x_1. The expected payoff of the consumer is

$$\frac{x_1}{x_2}(x_2 - x_1) = \overline{u}(x_1).$$

Thus, (D', ϕ') implements $\big(x_1, \overline{u}(x_1)\big)$.

Step 3: (D, ϕ) is more informative than (D', ϕ'). For every $(k, x) \in D$, define $\sigma(\cdot \mid (k, x)) \in \Delta D'$ as follows. With probability

$$\frac{x_1/x_2 - f(x_2)}{\pi_A/x_2 - f(x_2)},$$

$\sigma(\cdot \mid (k, x))$ draws k. With the remaining probability, $\sigma(\cdot \mid (k, x))$ draws k' uniformly at random from D'. Then,

$$\phi'(d' \mid \mathbf{v}) = \sum_{d \in D} \phi(d \mid \mathbf{v}) \sigma(d' \mid d), \quad \forall d' \in D', \forall \mathbf{v} \in X,$$

so (D, ϕ) is more informative than (D', ϕ'). □

Online Appendix

The Online Appendix contains the proofs of Lemmas B1 and B2 (Section OA.1) and extends Theorem 1 to correlated valuations (Section OA.2) and continuous valuations (Section OA.3).

OA.1 Proofs of Lemmas B1 and B2

To prove the lemmas, we first provide a formal statement of the set $\hat{S}_n$. We will use the generic notation ρ for the restriction of the seller's strategy to the problem of choosing, for each posterior belief μ, which product to offer and at what price, analogous to Section 2. We call ρ a strategy for short. Given a market segmentation τ and a strategy ρ, the expected gross payoff of the seller is $\Pi_\tau(\rho)$ and the expected payoff of the consumer is $U_\tau(\rho)$. We defined these expected payoffs in Section 2.

For convenience, we define ρ on the entire set of posterior beliefs ΔX^n, rather than just for the beliefs that have positive probability under the relevant market segmentation. A strategy ρ is optimal for the seller if

$$\forall \mu \in \Delta X^n : \quad (k^*, p^*) \in \operatorname{supp} \rho_\mu \implies (k^*, p^*) \in \underset{(k,p)}{\operatorname{argmax}}\, p \sum_{x \geq p} \mu_k(x), \tag{OA.1}$$

and it breaks ties in favor of the consumer if furthermore

$$(k', p') \in \underset{(k,p)}{\operatorname{argmax}}\, p \sum_{x \geq p} \mu_k(x) \implies \sum_{x \geq p^*} \mu_{k^*}(x)(x - p^*) \geq \sum_{x \geq p'} \mu_{k'}(x)(x - p'). \tag{OA.2}$$

Thus,

$$\hat{S}_n = \big\{ \big(\Pi_\tau(\rho), U_\tau(\rho)\big) \mid \rho \text{ satisfies (OA.1) and (OA.2)} \big\}.$$

Proof of Lemma B1. a) We first show that $\hat{S}_n$ is convex. Let $(\pi', u') \in \hat{S}_n$ and $(\pi'', u'') \in \hat{S}_n$. Thus, there exist market segmentations $\tau', \tau'' \in \Delta\Delta X^n$, and strategies ρ', ρ'' that satisfy (OA.1) and (OA.2), such that $\big(\Pi_{\tau'}(\rho'), U_{\tau'}(\rho')\big) = (\pi', u')$ as well as $\big(\Pi_{\tau''}(\rho''), U_{\tau''}(\rho'')\big) = (\pi'', u'')$. Then for $\lambda \in (0,1)$, the mixture $\tau = \lambda \tau' + (1-\lambda)\tau''$ is another market segmentation, and $\big(\Pi_\tau(\rho'), U_\tau(\rho')\big) = \big(\Pi_\tau(\rho''), U_\tau(\rho'')\big) = \lambda(\pi', u') + (1-\lambda)(\pi'', u'')$. Thus, $\hat{S}_n$ is convex.

Next, we show that $\{\pi \in \mathbb{R} \mid \text{there exists } u \in \mathbb{R} \text{ s.t. } (\pi,u) \in \hat{S}_n\} \subseteq [\pi_0, \pi_n^{\max}]$. Let τ be any market segmentation and ρ any optimal strategy. Then

$$\Pi_\tau(\rho) = \sum_\mu \tau(\mu) \max_{(k,p)} p \sum_{x \geq p} \mu_k(x)$$

and, letting k' be any product, we obtain

$$\Pi_\tau(\rho) \geq p_0 \sum_{x \geq p_0} \sum_\mu \tau(\mu)\mu_{k'}(x) = p_0 \sum_{x \geq p_0} f(x) = \pi_0,$$

and

$$\Pi_\tau(\rho) \leq \sum_\mu \tau(\mu) \sum_{\mathbf{v}} \mu(\mathbf{v}) \max_k v_k = \sum_{\mathbf{v}} \bar{\mu}(\mathbf{v}) \max_k v_k = \pi_n^{\max}.$$

Now, let τ be such that $\tau(\bar{\mu}) = 1$. Then, $\Pi_\tau(\rho) = \pi_0$ if ρ is optimal. Let τ' be the market segmentation that is supported on the Dirac measures of ΔX^n. That is, supp $\tau' = \{\delta^{\mathbf{v}} \in \Delta X^n \mid \mathbf{v} \in X^n\}$, where $\delta^{\mathbf{v}}$ assigns probability 1 to $\mathbf{v} \in X^n$, and $\tau'(\delta^{\mathbf{v}}) = \bar{\mu}(\mathbf{v})$. Then, $\max_{(k,p)} p \sum_{x \geq p} \delta_k^{\mathbf{v}}(x) = \max_k v_k$, implying $\Pi_{\tau'}(\rho) = \pi_n^{\max}$ at an optimal ρ. Part a) now follows, as $\hat{S}_n$ is convex.

b) We start with preliminaries. Define on $\Delta X^n \times \{1, \cdots, n\} \times X$ the functions $(\mu,k,p) \mapsto p\sum_{x \geq p} \mu_k(x)$ and $(\mu,k,p) \mapsto \sum_{x \geq p} \mu_k(x)(x-p)$. For fixed (k,p), $\mu \mapsto p\sum_{x \geq p} \mu_k(x)$ and $\mu \mapsto \sum_{x \geq p} \mu_k(x)(x-p)$ are continuous. Because $\{1, \cdots, n\} \times X$ is finite, it follows that $(\mu,k,p) \mapsto p\sum_{x \geq p} \mu_k(x)$ and $(\mu,k,p) \mapsto \sum_{x \geq p} \mu_k(x)(x-p)$ are continuous. Define the value function $a : \Delta X^n \to \mathbb{R}$ by

$$a(\mu) := \max_{(k,p) \in \{1,\cdots,n\} \times X} p \sum_{x \geq p} \mu_k(x),$$

and the correspondence $\phi : \Delta X^n \rightrightarrows \{1, \cdots, n\} \times X$ of maximizers by

$$\phi(\mu) := \{(k,p) \in \{1, \cdots, n\} \times X \mid p \sum_{x \geq p} \mu_k(x) = a(\mu)\}.$$

By the continuity of $(\mu,k,p) \mapsto p\sum_{x \geq p} \mu_k(x)$, the Maximum Theorem (Aliprantis and Border, 2006, Thm. 17.31) implies that a is continuous and ϕ upper hemicontinuous with nonempty compact values. Moreover, define the value function $b : \Delta X^n \to \mathbb{R}$ by

$$b(\mu) := \max_{(k,p) \in \phi(\mu)} \sum_{x \geq p} \mu_k(x)(x-p).$$

As $(\mu,k,p)\mapsto\sum_{x\geq p}\mu_k(x)(x-p)$ is continuous and ϕ upper hemicontinuous with nonempty compact values, b is upper semicontinuous (see Aliprantis and Border, 2006, Lem. 17.30).

If τ is a market segmentation and ρ a strategy that satisfies (OA.1) and (OA.2), then

$$\Pi_\tau(\rho)=\sum_\mu \tau(\mu)a(\mu) \quad \text{and} \quad U_\tau(\rho)=\sum_\mu \tau(\mu)b(\mu).$$

Fix some $\pi\in[\pi_0,\pi_n^{\max}]$ for the rest of the proof. The problem of finding a greatest element in $\{u\in\mathbb{R}\mid(\pi,u)\in\hat{S}_n\}$ can be stated as maximzing $\sum_\mu \tau(\mu)b(\mu)$ over all market segmentations τ such that $\sum_\mu \tau(\mu)a(\mu)=\pi$.

We momentarily enlarge the choice set of this problem so as to obtain a compact set. Let $\tilde{\Delta}\Delta X^n$ be the set of all Borel probability measures ζ on ΔX^n.[17] Let $Z\subset\tilde{\Delta}\Delta X^n$ be the subset of probability measures ζ that average to the prior belief $\bar{\mu}$,

$$\int \mu(\mathbf{v})\mathrm{d}\zeta(\mu)=\bar{\mu}(\mathbf{v}), \quad \forall\mathbf{v}\in X^n. \tag{OA.3}$$

We endow $\tilde{\Delta}\Delta X^n$ with the weak* topology. Because ΔX^n is compact and metrizable, the space $\tilde{\Delta}\Delta X^n$ is compact (see Aliprantis and Border, 2006, Thm. 15.11). Being a closed subset, it follows that Z is compact. By the continuity of a, $\zeta\mapsto\int a(\mu)\mathrm{d}\zeta(\mu)$ is continuous. Hence, $\{\zeta\in Z\mid\int a(\mu)\mathrm{d}\zeta(\mu)=\pi\}$ is compact. Furthermore, by the upper semicontinuity of b, $\zeta\mapsto\int b(\mu)\mathrm{d}\zeta(\mu)$ is upper semicontinuous (see Aliprantis and Border, 2006, Thm. 15.5). It follows that there exists a maximizer ζ^* for the problem

$$\max_{\zeta\in Z}\int b(\mu)\mathrm{d}\zeta(\mu) \quad \text{s.t.} \quad \int a(\mu)\mathrm{d}\zeta(\mu)=\pi.$$

It remains to show that there exists a market segmentation $\tau\in\Delta\Delta X^n$ such that

$$\sum_\mu \tau(\mu)b(\mu)=\int b(\mu)\mathrm{d}\zeta^*(\mu) \quad \text{and} \quad \sum_\mu \tau(\mu)a(\mu)=\pi. \tag{OA.4}$$

The tuple $\left(\bar{\mu},\pi,\int b(\mu)\mathrm{d}\zeta^*(\mu)\right)$ lies in the convex hull of

$$\left\{(\mu,r_1,r_2)\in\Delta X^n\times\mathbb{R}^2\mid(r_1,r_2)=\big(a(\mu),b(\mu)\big)\right\}.$$

Because the dimension of this set is finite, Caratheodory's Theorem allows us to express $\left(\bar{\mu},\pi,\int b(\mu)\mathrm{d}\zeta^*(\mu)\right)$ as a convex combination of finitely many elements. Denote a generic

[17] Thus, in contrast to $\tau\in\Delta\Delta X^n$, the support of $\zeta\in\tilde{\Delta}\Delta X^n$ need not be finite.

such element by (μ^y, r_1^y, r_2^y), and let $z^y > 0$ be the corresponding weight. Then, $\tau^* \in \Delta\Delta X^n$ with $\tau^*(\mu^y) = z^y$ is a market segmentation at which (OA.4) holds.

c) The concavity of $\bar{u}_n$ follows from the convexity of $\hat{S}_n$, which we showed in the proof of part a). Being concave, $\bar{u}_n$ is continuous at every $\pi \in (\pi_0, \pi_n^{\max})$, and $\lim_{\pi\to\pi_0} \bar{u}_n(\pi) \geq \bar{u}_n(\pi_0)$ and $\lim_{\pi\to\pi_n^{\max}} \bar{u}_n(\pi) \geq \bar{u}_n(\pi_n^{\max})$. It only remains to show that these weak inequalities hold with equality and that $\bar{u}_n(\pi_n^{\max}) = 0$.

By contradiction, suppose $\lim_{\pi\to\pi_0} \bar{u}_n(\pi) > \bar{u}_n(\pi_0)$. We use again the notation from the proof of part b). Let $(\pi^s)_{s\in\mathbb{N}}$ be a sequence with $\pi^s > \pi_0$ for all s and $\lim_{s\to\infty} \pi^s = \pi_0$. Let ρ be a strategy that satisfies (OA.1) and (OA.2), and let $(\tau^s)_{s\in\mathbb{N}}$ be a sequence of market segmentations such that $\Pi_{\tau^s}(\rho) = \pi^s$ and $U_{\tau^s}(\rho) = \bar{u}_n(\pi^s)$ for all s. Then, $\tau^s \in Z$ for all s. As Z is compact and metrizable (see Aliprantis and Border, 2006, Thm. 15.11), there exists a subsequence $\left(\tau^{s(t)}\right)_{t\in\mathbb{N}}$ that converges to some $\zeta' \in Z$. By the continuity of $\zeta \mapsto \int a(\mu)\mathrm{d}\zeta(\mu)$ and the upper semicontinuity of $\zeta \mapsto \int b(\mu)\mathrm{d}\zeta(\mu)$,

$$\pi_0 = \lim_{t\to\infty} \Pi_{\tau^{s(t)}}(\rho) = \lim_{t\to\infty} \sum_{\mu} \tau^{s(t)}(\mu) a(\mu) = \int a(\mu)\mathrm{d}\zeta'(\mu),$$

$$\limsup_{t\to\infty} U_{\tau^{s(t)}}(\rho) = \limsup_{t\to\infty} \sum_{\mu} \tau^{s(t)}(\mu) b(\mu) \leq \int b(\mu)\mathrm{d}\zeta'(\mu) \leq \bar{u}_n(\pi_0).$$

As in the proof of part b), there exists a market segmentation τ such that $\sum_\mu \tau(\mu)a(\mu) = \pi_0$ and $\sum_\mu \tau(\mu)b(\mu) = \int b(\mu)\mathrm{d}\zeta'(\mu)$. This yields a contradiction to $\lim_{\pi\to\pi_0} \bar{u}_n(\pi) > \bar{u}_n(\pi_0)$. Hence, $\lim_{\pi\to\pi_0} \bar{u}_n(\pi) = \bar{u}_n(\pi_0)$.

By contradiction, suppose $\lim_{\pi\to\pi_n^{\max}} \bar{u}_n(\pi) = \eta > 0$. Then, there exist $\varepsilon, \delta > 0$ and π such that $\pi_n^{\max} - \pi < \delta$, $|\eta - \bar{u}_n(\pi)| < \varepsilon$, and $\varepsilon + \delta < \eta$. Let ρ be a strategy that satisfies (OA.1) and (OA.2), and let τ be a market segmentation such that $\Pi_\tau(\rho) = \pi$ and $U_\tau(\rho) = \bar{u}_n(\pi)$. Then, $\Pi_\tau(\rho) + U_\tau(\rho) > \pi_n^{\max} - \delta + \eta - \varepsilon > \pi_n^{\max}$. But

$$\begin{aligned}
\Pi_\tau(\rho) + U_\tau(\rho) &= \sum_{\mu} \tau(\mu) \sum_{k,p} \rho_\mu(k,p) \sum_{x\geq p} \mu_k(x)x \\
&\leq \sum_{\mu} \tau(\mu) \max_k \sum_{x} \mu_k(x)x \\
&\leq \sum_{\mu} \tau(\mu) \sum_{\mathbf{v}} \mu(\mathbf{v}) \max_k v_k = \sum_{\mathbf{v}} \bar{\mu}(\mathbf{v}) \max_k v_k = \pi_n^{\max};
\end{aligned}$$

a contradiction to $\Pi_\tau(\rho)+U_\tau(\rho) > \pi_n^{\max}$. Hence, $\lim_{\pi\to\pi_n^{\max}} \overline{u}_n(\pi) = 0$. Since

$$\lim_{\pi\to\pi_n^{\max}} \overline{u}_n(\pi) \geq \overline{u}_n(\pi_n^{\max}) \geq 0,$$

this also proves that $\overline{u}_n(\pi_n^{\max}) = 0$. □

Proof of Lemma B2. The first part of the lemma holds because

$$\hat{S}_n \subseteq \left\{ \left(\Pi_\tau(\rho), U_\tau(\rho)\right) \mid \rho \text{ satisfies (OA.1)} \right\} = S_n \subseteq S.$$

Next, let $\left(\pi, \overline{u}(\pi)\right) \in S$, where we may assume $\pi < x_m$. By Theorem 1, there exists a sequence $\left((\pi_n, u_n)\right)_{n\in\mathbb{N}}$ such that $(\pi_n, u_n) \in S_n$ and $\lim_{n\to\infty}(\pi_n, u_n) = \left(\pi, \overline{u}(\pi)\right)$. Because $\lim_{n\to\infty} \pi_n^{\max} = x_m$, we have $\pi_n \in [\pi_0, \pi_n^{\max}]$ for n sufficiently large, say $n > n'$. Consider the sequence $\left((\tilde{\pi}_n, \overline{u}_n(\tilde{\pi}_n))\right)_{n\in\mathbb{N}}$, where $\tilde{\pi}_n = \pi_0$ for $n \leq n'$ and $\tilde{\pi}_n = \pi_n$ for $n > n'$. By construction, $\left(\tilde{\pi}_n, \overline{u}_n(\tilde{\pi}_n)\right) \in \hat{S}_n$. Furthermore, $\lim_{n\to\infty} \tilde{\pi}_n = \pi$. It remains to show that $\lim_{n\to\infty} \overline{u}_n(\tilde{\pi}_n) = \overline{u}(\pi)$. Because $\hat{S}_n$ differs from S_n only by the additional condition (OA.2), according to which the seller breaks ties in favor of the consumer, it holds that $\overline{u}_n(\tilde{\pi}_n) = \max\left\{u \in \mathbb{R} \mid (\tilde{\pi}_n, u) \in S_n\right\} \geq u_n$ for $n > n'$. Hence, $\liminf_{n\to\infty} \overline{u}_n(\tilde{\pi}_n) \geq \lim_{n\to\infty} u_n$. On the other hand, $\hat{S}_n \subseteq S$ and the continuity of the function $\overline{u}$ imply $\limsup_{n\to\infty} \overline{u}_n(\tilde{\pi}_n) \leq \lim_{n\to\infty} \overline{u}(\tilde{\pi}_n) = \overline{u}(\pi) = \lim_{n\to\infty} u_n$. Thus, $\lim_{n\to\infty} \overline{u}_n(\tilde{\pi}_n) = \overline{u}(\pi)$. □

OA.2 Correlated Valuations

Here, we present a generalization of the baseline model in Section 2 that allows for correlation between valuations, and show that Theorem 1 extends.

We replace the definition of the aggregate market $\bar{\mu}$ in (1) by

$$\bar{\mu}(\mathbf{v}) = f(v_1) \prod_{k=2}^{n} \left(t\delta^{v_{k-1}}(v_k) + (1-t) f(v_k)\right), \quad \forall \mathbf{v} \in X^n,$$

where $t \in [0,1)$ and $\delta^x \in \Delta X$ denotes the Dirac measure centered on $x \in X$. Thus, the valuation vector corresponds to a Markov chain. With probability t, the valuation for product k coincides with the one for product $k-1$; with probability $1-t$, the valuation for product k has distribution f, the distribution of the first product. The interpretation is that adjacent products are similar, so that consumers may have similar valuations. The correlation between the valuations, captured by t, can be arbitrarily strong; we only exclude perfect correlation. The original model assumed $t = 0$ (no correlation).

We show that Theorem 1 extends to this specification. Lemma 1 and thus the first sentence of the theorem obviously still hold. To prove the second sentence of the theorem, we only need to show that Lemma 3 still holds.

We present an adapted proof of Lemma 3. The broad idea is as follows. Because the correlation is imperfect, the aggregate market can still be segmented such that, independently for each product k, the distribution of valuations is either equal to a given g_i or some residual. In contrast to the original proof of Lemma 3, the residual now depends on the valuation for product $k-1$. We then show that the seller always prefers to offer a product for which the distribution of valuations is g_i.

Proof of Lemma 3. Let $x_i \in \{\pi_0, \dots, x_m\}$ and $p \in \{x_i, \dots, x_m\}$. Analogously to the proof of Lemma 3 for the original model, choose $\lambda \in (0,1)$ such that

$$\lambda g_i(x) \le t\delta^y(x) + (1-t)f(x), \quad \forall x, y \in X,$$
$$\lambda g_i(x) \le f(x), \quad \forall x \in X,$$

and define $h(\cdot \mid v_0) \in \Delta X$, and, for every $y \in X$, $h(\cdot \mid y) \in \Delta X$, as

$$h(\cdot \mid v_0) := \frac{1}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i,$$
$$h(\cdot \mid y) := \frac{t}{1-\lambda}\delta^y + \frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i.$$

Next, we present a market segmentation τ supported on 2^n markets. The markets in the support of τ are indexed by superscript $\mathbf{a} \in \{\mathtt{g}, \mathtt{h}\}^n$. The notation $\zeta(a_k)$ will also be used and means 1 if $a_k = \mathtt{g}$ and 0 if $a_k = \mathtt{h}$. Set

$$\tau(\mu^{\mathbf{a}}) := \prod_k \big(\zeta(a_k)\lambda + (1-\zeta(a_k))(1-\lambda)\big), \quad \forall \mathbf{a} \in \{\mathtt{g}, \mathtt{h}\}^n.$$

Market $\mu^{\mathbf{a}}$ is given by

$$\mu^{\mathbf{a}}(\mathbf{v}) := \frac{\prod_k \big(\zeta(a_k)\lambda g_i(v_k) + (1-\zeta(a_k))(1-\lambda)h(v_k \mid v_{k-1})\big)}{\tau(\mu^{\mathbf{a}})}$$
$$= \prod_k \big(\zeta(a_k) g_i(v_k) + (1-\zeta(a_k)) h(v_k \mid v_{k-1})\big), \quad \forall \mathbf{v} \in X^n.$$

Then τ is a market segmentation:

$$\sum_{\mathbf{a}} \tau(\mu^{\mathbf{a}})\mu^{\mathbf{a}}(\mathbf{v}) = f(v_1)\prod_{k=2}^{n}\left(t\delta^{v_{k-1}}(v_k) + (1-t)f(v_k)\right) = \bar{\mu}(\mathbf{v}), \quad \forall \mathbf{v} \in X^n.$$

Next, consider any market $\mu^{\mathbf{a}}$. If $a_k = \mathrm{g}$ for any $k \in \{1, \ldots, n\}$, then

$$\begin{aligned}
\mu_k^{\mathbf{a}}(x) &= \sum_{\mathbf{v}:v_k=x} \mu^{\mathbf{a}}(\mathbf{v}) \\
&= \sum_{\mathbf{v}:v_k=x} \prod_{k'} \left(\zeta(a_{k'})g_i(v_{k'}) + (1-\zeta(a_{k'}))h(v_{k'} \mid v_{k'-1})\right) \\
&= \sum_{v_1,\ldots,v_{k-1}} \prod_{k'<k} \left(\zeta(a_{k'})g_i(v_{k'}) + (1-\zeta(a_{k'}))h(v_{k'} \mid v_{k'-1})\right) g_i(x) \\
&= g_i(x), \quad \forall x \in X.
\end{aligned}$$

Hence,

$$\max_q q \sum_{x \geq q} \mu_k^{\mathbf{a}}(x) = \max_q q \sum_{x \geq q} g_i(x) = p \sum_{x \geq p} g_i(x) = x_i. \tag{OA.5}$$

In the following, we show that if $a_k = \mathtt{h}$ for any $k \in \{1, \ldots, n\}$, then

$$x_i \geq \max_q q \sum_{x \geq q} \mu_k^{\mathbf{a}}(x). \tag{OA.6}$$

For $k = 1$,

$$\begin{aligned}
\mu_k^{\mathbf{a}}(x) &= \sum_{\mathbf{v}:v_k=x} \mu^{\mathbf{a}}(\mathbf{v}) \\
&= \sum_{\mathbf{v}:v_k=x} \prod_{k'} \left(\zeta(a_{k'})g_i(v_{k'}) + (1-\zeta(a_{k'}))h(v_{k'} \mid v_{k'-1})\right) \\
&= h(x \mid v_0).
\end{aligned}$$

For $\mu_k^{\mathbf{a}} = h(\cdot \mid v_0)$, (OA.6) was shown as (12) in the proof of Lemma 3.

So suppose $k \in \{2, \ldots, n\}$. Let $r^* \in \{0 \ldots, k-1\}$ be the number of products such that $a_{k'} = \mathtt{h}$ for $k' < k$ and $a_{k''} \neq \mathrm{g}$ for $k' < k'' < k$. Then

$$\begin{aligned}
\mu_k^{\mathbf{a}}(x) &= \sum_{\mathbf{v}:v_k=x} \mu^{\mathbf{a}}(\mathbf{v}) \\
&= \sum_{\mathbf{v}:v_k=x} \prod_{k'} \left(\zeta(a_{k'})g_i(v_{k'}) + (1-\zeta(a_{k'}))h(v_{k'} \mid v_{k'-1})\right) \\
&= \sum_{v_1,\ldots,v_{k-1}} \prod_{k'<k} \left(\zeta(a_{k'})g_i(v_{k'}) + (1-\zeta(a_{k'}))h(v_{k'} \mid v_{k'-1})\right) h(x \mid v_{k-1}) \\
&= \sum_{v_{k-r^*-1},\ldots,v_{k-1}} e(v_{k-r^*-1}) \left(\prod_{k'=k-r^*}^{k-1} h(v_{k'} \mid v_{k'-1})\right) h(x \mid v_{k-1}),
\end{aligned}$$

where $e \in \{g_i, h(\cdot \mid v_0)\}$.

We show by induction that

$$\sum_{v_{k-r-1},\cdots,v_{k-1}} e(v_{k-r-1}) \left(\prod_{k'=k-r}^{k-1} h(v_{k'} \mid v_{k'-1}) \right) h(\cdot \mid v_{k-1}) \in \Delta X$$

is equal to

$$\left(\frac{t}{1-\lambda}\right)^{r+1} e + \frac{1-\left(\frac{t}{1-\lambda}\right)^{r+1}}{1-\frac{t}{1-\lambda}} \left(\frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i\right) \in \Delta X$$

for all $r \in \mathbb{N}$. If $r = 0$, then

$$\begin{aligned}
\sum_{v_{k-1}} e(v_{k-1}) h(\cdot \mid v_{k-1}) &= \sum_{v_{k-1}} e(v_{k-1}) \frac{t}{1-\lambda} \delta^{v_{k-1}} + \frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i \\
&= \frac{t}{1-\lambda} e + \frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i.
\end{aligned}$$

Suppose equality holds for a given $r \geq 0$. Then equality holds for $r+1$:

$$\begin{aligned}
& \sum_{v_{k-r-2},\cdots,v_{k-1}} e(v_{k-r-2}) \left(\prod_{k'=k-r-1}^{k-1} h(v_{k'} \mid v_{k'-1}) \right) h(\cdot \mid v_{k-1}) \\
= & \sum_{v_{k-r-2},\cdots,v_{k-1}} e(v_{k-r-2}) \left(\prod_{k'=k-r-1}^{k-1} h(v_{k'} \mid v_{k'-1}) \right) \frac{t}{1-\lambda} \delta^{v_{k-1}} + \frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i \\
= & \frac{t}{1-\lambda} \sum_{v_{k-r-2},\cdots,v_{k-2}} e(v_{k-r-2}) \left(\prod_{k'=k-r-1}^{k-2} h(v_{k'} \mid v_{k'-1}) \right) h(\cdot \mid v_{k-2}) + \frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i \\
= & \frac{t}{1-\lambda} \left(\left(\frac{t}{1-\lambda}\right)^{r+1} e + \frac{1-\left(\frac{t}{1-\lambda}\right)^{r+1}}{1-\frac{t}{1-\lambda}} \left(\frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i\right) \right) + \frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i \\
= & \left(\frac{t}{1-\lambda}\right)^{r+2} e + \frac{1-\left(\frac{t}{1-\lambda}\right)^{r+2}}{1-\frac{t}{1-\lambda}} \left(\frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i\right).
\end{aligned}$$

We have shown that

$$\mu_k^{\mathbf{a}} = \left(\frac{t}{1-\lambda}\right)^{r^*+1} e + \frac{1-\left(\frac{t}{1-\lambda}\right)^{r^*+1}}{1-\frac{t}{1-\lambda}} \left(\frac{1-t}{1-\lambda} f - \frac{\lambda}{1-\lambda} g_i\right).$$

So

$$
\begin{aligned}
& \max_q q \sum_{x \geq q} \mu_k^{\mathbf{a}}(x) \\
= \quad & \max_q \sum_{x \geq g} \left(\left(\frac{t}{1-\lambda} \right)^{r^*+1} e(x) + \frac{1 - \left(\frac{t}{1-\lambda} \right)^{r^*+1}}{1 - \frac{t}{1-\lambda}} \left(\frac{1-t}{1-\lambda} f(x) - \frac{\lambda}{1-\lambda} g_i(x) \right) \right) \\
\leq \quad & \left(\frac{t}{1-\lambda} \right)^{r^*+1} \max_q q \sum_{x \geq g} e(x) + \frac{1 - \left(\frac{t}{1-\lambda} \right)^{r^*+1}}{1 - \frac{t}{1-\lambda}} \max_q q \sum_{x \geq g} \left(\frac{1-t}{1-\lambda} f(x) - \frac{\lambda}{1-\lambda} g_i(x) \right) \\
\leq \quad & \left(\frac{t}{1-\lambda} \right)^{r^*+1} x_i + \frac{1 - \left(\frac{t}{1-\lambda} \right)^{r^*+1}}{1 - \frac{t}{1-\lambda}} \max_q q \sum_{x \geq g} \left(\frac{1-t}{1-\lambda} f(x) - \frac{\lambda}{1-\lambda} g_i(x) \right).
\end{aligned}
$$

For any $q \in \{x_1, \ldots, x_{i-1}\}$:

$$
q \sum_{x \geq g} \left(\frac{1-t}{1-\lambda} f(x) - \frac{\lambda}{1-\lambda} g_i(x) \right) \leq \frac{1-t}{1-\lambda} q - \frac{\lambda}{1-\lambda} q,
$$

implying

$$
\begin{aligned}
q \sum_{x \geq q} \mu_k^{\mathbf{a}}(x) \quad & \leq \quad \left(\frac{t}{1-\lambda} \right)^{r^*+1} x_i + \frac{1 - \left(\frac{t}{1-\lambda} \right)^{r^*+1}}{1 - \frac{t}{1-\lambda}} \frac{1-t-\lambda}{1-\lambda} q \\
& = \quad \left(\frac{t}{1-\lambda} \right)^{r^*+1} x_i + \left(1 - \left(\frac{t}{1-\lambda} \right)^{r^*+1} \right) q \\
& \leq \quad x_i.
\end{aligned}
$$

For any $q \in \{x_i, \ldots, x_m\}$:

$$
q \sum_{x \geq g} \left(\frac{1-t}{1-\lambda} f(x) - \frac{\lambda}{1-\lambda} g_i(x) \right) \leq \frac{1-t}{1-\lambda} \pi_0 - \frac{\lambda}{1-\lambda} x_i,
$$

implying

$$
\begin{aligned}
q \sum_{x \geq q} \mu_k^{\mathbf{a}}(x) \quad & \leq \quad \left(\frac{t}{1-\lambda} \right)^{r^*+1} x_i + \frac{1 - \left(\frac{t}{1-\lambda} \right)^{r^*+1}}{1 - \frac{t}{1-\lambda}} \frac{1-t-\lambda}{1-\lambda} x_i \\
& = \quad x_i.
\end{aligned}
$$

Thus, (OA.6) holds.

By (OA.5) and (OA.6), there exists an optimal strategy ρ for the seller with the following property: for every market $\mu^{\mathbf{a}}$ such that $a_k = \mathrm{g}$ for some product k, offer such a product at price p.

Lastly, observe that the only market $\mu^{\mathbf{a}}$ with $a_k \neq \mathrm{g}$ for all $k \in \{1, \ldots, n\}$ has

$$\tau(\mu^{\mathbf{a}}) = (1-\lambda)^n.$$

Let π_n be the surplus of the seller, and u_n the consumer surplus, under this market segmentation and such an optimal strategy. Then

$$\lim_{n\to\infty} \pi_n = \lim_{n\to\infty} (1-(1-\lambda)^n) x_i = x_i,$$
$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} (1-(1-\lambda)^n) \sum_{x \geq p} g_i(x)(x-p) = \sum_{x \geq p} g_i(x)(x-p). \qquad \square$$

OA.3 Continuous Valuations

Here, we show that Theorem 1 extends to continuous valuations ($X = [0,1]$).

In this section, ΔY denotes the set of all distributions (i.e., Borel probability measures) on space Y. $\mathscr{B}(\cdot)$ denotes the Borel σ-algebra. Sets of distributions are endowed with the weak* topology. All distributions on product spaces in this section are uniquely defined by its values on the products of the Borel σ-algebras (see Aliprantis and Border, 2006, Thms. 4.44, 10.10, and 15.11). We write "for all products of Borel sets" rather than "for all Borel sets of the product space" where convenient.

Let $X = [0,1]$, and let f be an atomless distribution in ΔX that has full support. The aggregate market $\bar{\mu} \in \Delta X^n$ is now defined by

$$\bar{\mu}(B_1 \times \cdots \times B_L) = \prod_k f(B_k), \quad \forall B_1 \times \cdots \times B_n \in \prod_k \mathscr{B}(X).$$

A market segmentation is a distribution $\tau \in \Delta\Delta V$ of markets $\mu \in \Delta X^n$ that averages to $\bar{\mu}$:

$$\int \mu(B_1 \times \cdots \times B_L) \tau(\mathrm{d}\mu) = \prod_k f(B_k), \quad \forall B_1 \times \cdots \times B_n \in \prod_k \mathscr{B}(X). \tag{OA.7}$$

The k-marginal of a market $\mu \in \Delta X^n$ is $\mu_k \in \Delta X$, given by:

$$\mu_k(B) = \int_{\mathbf{v}: v_k \in B} \mu(\mathrm{d}\mathbf{v}), \quad \forall B \in \mathscr{B}(X).$$

A strategy of the seller is a mapping $\rho : \Delta X^n \times \mathscr{B}(\{1,\dots,n\} \times X) \to [0,1]$ such that $\rho(\mu,\cdot) \in \Delta(\{1,\dots,n\} \times X)$ for all $\mu \in \Delta X^n$ and $\mu \mapsto \rho(\mu,\{k\} \times B)$ is measurable for all $\{k\} \times B \in \mathscr{B}(\{1,\dots,n\} \times X)$. Thus, a strategy selects, potentially randomly, a product $k \in \{1,\dots,n\}$ to be offered and a price $p \in X$ to be charged for any market $\mu \in \Delta X^n$.

The producer surplus under market segmentation τ and strategy ρ is

$$\Pi_\tau(\rho) := \int\int p\mu_k([p,1])\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu),$$

and the consumer surplus is

$$\begin{aligned} U_\tau(\rho) \quad &:= \quad \int\int\int_p^1 (x-p)\mu_k(\mathrm{d}x)\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu) \\ &= \quad \int\int\int_p^1 \mu_k([x,1])\mathrm{d}x\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu), \end{aligned}$$

where the second equality uses integration by parts. A strategy ρ^* is optimal for the seller under market segmentation τ if

$$\rho^* \in \operatorname*{argmax}_\rho \Pi_\tau(\rho), \tag{OA.8}$$

A combination of producer and consumer surplus (π,c) is feasible if there exists a market segmentation τ and an optimal strategy ρ such that $(\pi,u) = (\Pi_\tau(\rho),U_\tau(\rho))$. The set of feasible surplus pairs is denoted by

$$S_n := \big\{\Pi_\tau(\rho),U_\tau(\rho)) \mid \tau \text{ satisfies (OA.7)}, \ \rho \text{ satisfies (OA.8)}\big\}.$$

For any $\pi \in (0,1]$, define the distribution $g_\pi \in \Delta X$ by

$$g_\pi([0,x]) := \begin{cases} 0 & \text{if } x \in [0,\pi), \\ 1-\frac{\pi}{x} & \text{if } x \in [\pi,1), \\ 1 & \text{if } x = 1. \end{cases}$$

It is readily checked that

$$pg_\pi([p,1]) = \begin{cases} p & \text{for all } p \in [0,\pi), \\ \pi & \text{for all } p \in [\pi,1]. \end{cases} \tag{OA.9}$$

For $\pi \in [0,1]$, let

$$\overline{u}(\pi) := \int_{\pi}^{1} g_{\pi}([x,1])\mathrm{d}x = \int_{\pi}^{1} \pi/x\mathrm{d}x = -\pi \ln \pi,$$

and define the set

$$S := \left\{(\pi, c) \mid \pi \in [\pi_0, 1], u \in [0, \overline{u}(\pi)]\right\}.$$

We now show that Theorem 1 extends to this model. Our proof uses three lemmas. The first lemma is analogous to Lemma 1.

Lemma OA.1. *If $(\pi, u) \in S_n$, then $u \leq \overline{u}(\pi)$.*

Proof. Let τ be any market segmentation, and let ρ be any strategy that is optimal given τ such that $\Pi_\tau(\rho) = \pi$ and $U_\tau(\rho) = u$. Define the distribution $h \in \Delta X$ by

$$h(B) := \int \mu_k(B)\rho(\mu, \mathrm{d}(k,p))\tau(\mathrm{d}\mu), \quad \forall B \in \mathscr{B}(X).$$

By the optimality of ρ, we have for any $q \in [\pi, 1]$

$$\begin{aligned} \int\int q\mu_k([q,1])\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu) \quad &\leq \quad \int\int p\mu_k([q,1])\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu) \\ &= \quad \Pi_\tau(\rho) = \pi = qg_\pi([q,1]). \end{aligned}$$

Dividing through by q, we see that g_π first-order stochastically dominates h. Hence,

$$\begin{aligned} u \quad &= \quad \int\int\int_p^1 \mu_k([x,1])\mathrm{d}x\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu) \\ &= \quad \int\int\int \mu_k([x,1])\mathrm{d}x\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu) - \int\int\int_0^p \mu_k([x,1])\mathrm{d}x\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu) \\ &\leq \quad \int\int\int \mu_k([x,1])\mathrm{d}x\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu) - \int\int\int_0^p \mu_k([p,1])\mathrm{d}x\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu) \\ &= \quad \int\int\int \mu_k([x,1])\mathrm{d}x\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu) - \pi \\ &= \quad \int\int\int \mu_k([x,1])\rho(\mu,\mathrm{d}(k,p))\tau(\mathrm{d}\mu)\mathrm{d}x - \pi \\ &= \quad \int h([x,1])\mathrm{d}x - \pi \\ &\leq \quad \int g_\pi([x,1])\mathrm{d}x - \pi \\ &= \quad \overline{u}(\pi), \end{aligned}$$

where we used Fubini's Theorem for the fifth row. □

The next lemma is analogous to Lemma 3.

Lemma OA.2. *Let $e,h \in \Delta X$ and $\lambda \in (0,1)$ such that*

$$\lambda e(B) + (1-\lambda)h(B) = f(B) \quad \forall B \in \mathscr{B}(X), \tag{OA.10}$$

$$\max_x xe([x,1]) \geq \max_x xh([x,1]). \tag{OA.11}$$

Let $p \in \operatorname{argmax}_x xe([x,1])$. There exists a sequence $\big((\pi_n,u_n)\big)_{n\in\mathbb{N}}$ such that $(\pi_n,u_n) \in S_n$ and

$$(\pi_n,u_n) \underset{n\to\infty}{\longrightarrow} \left(pe([p,1]), \int_p^1 e([x,1])\mathrm{d}x\right). \tag{OA.12}$$

Proof. Fix some $n \in \mathbb{N}$. We present a market segmentation τ supported on 2^n markets. The markets in the support of τ are indexed by superscript $\mathbf{a} \in \{\mathtt{e},\mathtt{h}\}^n$. The notation $\zeta(a_k)$ will also be used and means 1 if $a_k = \mathtt{e}$ and 0 if $a_k = \mathtt{h}$. Set

$$\tau(\mu^{\mathbf{a}}) := \prod_k \big(\zeta(a_k)\lambda + (1-\zeta(a_k))(1-\lambda)\big), \quad \forall \mathbf{a} \in \{\mathtt{e},\mathtt{h}\}^n.$$

Market $\mu^{\mathbf{a}}$ is given by

$$\begin{aligned}\mu^{\mathbf{a}}(B_1\times\cdots\times B_n) &:= \frac{\prod_k \big(\zeta(a_k)\lambda e(B_k) + (1-\zeta(a_k))(1-\lambda)h(B_k)\big)}{\tau(\mu^{\mathbf{a}})} \\ &= \prod_k \big(\zeta(a_k)e(B_k) + (1-\zeta(a_k))h(B_k)\big), \quad \forall B_1\times\cdots\times B_n \in \prod_k \mathscr{B}(X).\end{aligned}$$

Then τ is a market segmentation:

$$\sum_{\mathbf{a}} \tau(\mu^{\mathbf{a}})\mu^{\mathbf{a}}(\mathbf{v}) = \prod_k f(B_k) = \bar{\mu}(\mathbf{v}), \quad \forall B_1\times\cdots\times B_n \in \prod_k \mathscr{B}(X).$$

Consider any market $\mu^{\mathbf{a}}$. If $a_k = \mathtt{e}$ for any $k \in \{1,\dots,n\}$, then

$$\mu^{\mathbf{a}}_k(B) = \int_{\mathbf{v}:v_k\in B} \mu(\mathrm{d}\mathbf{v}) = e(B), \quad \forall B \in \mathscr{B}(X).$$

If $a_k = \mathtt{h}$ for any $k \in \{1,\dots,n\}$, then $\mu^{\mathbf{a}}_k = h$.

Next, we describe a strategy ρ as follows:

- For every market $\mu^{\mathbf{a}}$ in the support of τ such that $a_k = \mathtt{e}$ for some product k, offer any such product at price p. Note that for such markets, $\mu^{\mathbf{a}}_{k'} \in \{e,h\}$ for all $k' \in \{1,\dots,n\}$.

- For the unique market $\mu^{\mathbf{a}}$ in the support of τ such that $a_k = \mathtt{h}$ for all products k, offer product $k = 1$ at some price $p_0 \in \operatorname{argmax}_x xh([x,1])$.

- For every posterior outside of the support of τ, offer product $k = 1$ at price p_0.

We have not specified how ρ selects among products k with $a_k = \mathtt{e}$, but this indeterminacy will not matter. Note furthermore that V is compact and metrizable, which implies that ΔX^n is metrizable (Aliprantis and Border, 2006, Thm. 15.11). Hence, any finite subset of ΔX^n is a Borel set. Because the support of τ is finite, this implies that $\mu \mapsto \rho(\mu, \{k\} \times B)$ is measurable for all $k \in \{1, \ldots, n\}$ and all $B \in \mathscr{B}(X)$, as required by the definition of a strategy. By (OA.10), ρ satisfies optimality (OA.8).

Lastly, observe that the only market $\mu^{\mathbf{a}}$ with $a_k = \mathtt{h}$ for all $k \in \{1, \ldots, n\}$ has

$$\tau(\mu^{\mathbf{a}}) = (1 - \lambda)^n.$$

Let π_n be the surplus of the seller, and u_n the consumer surplus, under this market segmentation and such an optimal strategy. Then

$$\lim_{n\to\infty} \pi_n = \lim_{n\to\infty} (1 - (1-\lambda)^n) pe([p,1]) = pe([p,1]),$$
$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} (1 - (1-\lambda)^n) \int_p^1 e([x,1]) \mathrm{d}x = \int_p^1 e([x,1]) \mathrm{d}x. \qquad \square$$

Lemma OA.2 is not directly useful because the distributions g_π have an atom at $x = 1$ whereas f is atomless. We will approximate the respective g_π by atomless distributions. To this end, we now introduce a third lemma, which has no analog in the original model.

Fix some $\pi \in (\pi_0, 1)$, some $p \in [\pi, 1]$, and some $\varepsilon \in [0, 1]$. For $N \in \mathbb{N}$, let

$$\pi = x_0^N < x_1^N < \ldots < x_N^N = 1$$

be a collection of points in $[\pi, 1]$ of equal distance. For each $i = 1, \ldots, N$, define

$$\alpha_{\pi,p}^{\varepsilon,N}(i) := \frac{(1-\varepsilon) g_\pi([x_{i-1}^N, x_i^N]) + \varepsilon \delta^p([x_{i-1}^N, x_i^N])}{f([x_{i-1}^N, x_i^N])},$$

where $\delta^p \in \Delta X$ denotes the Dirac measure centered on p. Let

$$\lambda_{\pi,p}^N := \min_{\substack{\varepsilon \in [0,1] \\ i \in \{1,\ldots,N\}}} \frac{1}{a_{\pi,p}^{\varepsilon,N}(i)}$$

and note that $\lambda_{\pi,p}^N \in (0,1)$ because g_π and δ^p are distributions and assign probability one to $[\pi, 1]$ whereas f assigns probability strictly less than one to $[\pi, 1]$. For each $i = 1, \ldots, N$, let furthermore

$$\beta^{\varepsilon,N}_{\pi,p}(i) = \frac{1}{1-\lambda^N_{\pi,p}} - \frac{\lambda^N_{\pi,p}}{1-\lambda^N_{\pi,p}}\alpha^{\varepsilon,N}_{\pi,p}(i).$$

We now introduce two distributions in ΔX: the distribution $e^{\varepsilon,N}_{\pi,p}$, which has support $[\pi,1]$ and is given by

$$e^{\varepsilon,N}_{\pi,p}([x,x^N_i]) := \alpha^{\varepsilon,N}_{\pi,p}(i)f([x,x^N_i]), \quad \forall x \in [x^N_{i-1},x^N_i], \forall i = 1,\dots,N,$$

and the distribution $h^{\varepsilon,N}_{\pi,p}$ given by

$$h^{\varepsilon,N}_{\pi,p}([0,x_i]) := \frac{1}{1-\lambda^N_{\pi,p}} f([0,x]), \quad \forall x \in [0,\pi),$$
$$h^{\varepsilon,N}_{\pi,p}([x,x_i]) := \beta^{\varepsilon,N}_{\pi,p}(i)f([x,x^N_i]), \quad \forall x \in [x^N_{i-1},x^N_i], \forall i = 1,\dots,N.$$

Then, f is a mixture of $e^{\varepsilon,N}_{\pi,p}$ and $h^{\varepsilon,N}_{\pi,p}$:

$$\lambda^N_{\pi,p}e^{\varepsilon,N}_{\pi,p}(B) + (1-\lambda^N_{\pi,p})h^{\varepsilon,N}_{\pi,p}(B) = f(B) \quad \forall B \in \mathscr{B}(X). \tag{OA.13}$$

Lemma OA.3. *Let $\pi \in (\pi_0,1)$ and $p \in [\pi,1]$. For every $\varepsilon \in (0,1]$, let $(p^{\varepsilon,N})_{N\in\mathbb{N}}$ be a sequence of prices such that*

$$p^{\varepsilon,N} \in \operatorname*{argmax}_x x e^{\varepsilon,N}_{\pi,p}([x,1]), \quad \forall N \in \mathbb{N}.$$

Then it holds that

$$\lim_{\varepsilon\to 0}\lim_{N\to\infty} p^{\varepsilon,N} e^{\varepsilon,N}_{\pi,p}([p^{\varepsilon,N},1]) = \pi, \tag{OA.14}$$
$$\lim_{\varepsilon\to 0}\lim_{N\to\infty} \int_{p^{\varepsilon,N}}^1 e^{\varepsilon,N}_{\pi,p}([x,1])\mathrm{d}x = \int_p^1 g_\pi([x,1])\mathrm{d}x. \tag{OA.15}$$

Moreover,

$$\lim_{\varepsilon\to 0}\max_x x h^{\varepsilon,N}_{\pi,p}([x,1]) < \pi_0, \quad \forall N \in \mathbb{N}. \tag{OA.16}$$

Proof. Let $\hat{e}^{\varepsilon}_{\pi,p} \in \Delta X$ be the distribution given by

$$\hat{e}^{\varepsilon}_{\pi,p}(B) := (1-\varepsilon)g_\pi(B) + \varepsilon\delta^p(B), \quad \forall B \in \mathscr{B}(X).$$

To prove (OA.14), we first show

$$\lim_{N\to\infty} p^{\varepsilon,N} = p, \tag{OA.17}$$

$$\lim_{N\to\infty} e^{\varepsilon,N}_{\pi,p}([p^{\varepsilon,N},1]) = \hat{e}^{\varepsilon}_{\pi,p}([p,1]). \tag{OA.18}$$

Fix some $\varepsilon \in (0,1]$. For all $x \in [x_{i-1}^N, x_i^N]$ and all $i = 1, \dots, N-1$,

$$e_{\pi,p}^{\varepsilon,N}([x,1]) = \frac{\hat{e}_{\pi,p}^{\varepsilon}([x_{i-1}^N, x_i^N])}{f([x_{i-1}^N, x_i^N])} f([x, x_i^N]) + \hat{e}_{\pi,p}^{\varepsilon}([x_i^N, 1]).$$

Hence,

$$\lim_{N\to\infty} e_{\pi,p}^{\varepsilon,N}([x,1]) = \hat{e}_{\pi,p}^{\varepsilon}([x,1]), \quad \forall x \in [\pi, 1). \tag{OA.19}$$

It follows that if $p < 1$, then $\lim_{N\to\infty} p^{\varepsilon,N} = p$ because $x \mapsto x\hat{e}_{\pi,p}^{\varepsilon}([x,1])$ is uniquely maximized at $x = p$ by (OA.9). If $p = 1$, then $\lim_{N\to\infty} p^{\varepsilon,N} = p$ because

$$\lim_{N\to\infty} x_{N-1}^N e_{\pi,p}^{\varepsilon,N}([x_{N-1}^N, 1]) = \lim_{N\to\infty} x_{N-1}^N \hat{e}_{\pi,p}^{\varepsilon}([x_{N-1}^N, 1]) = \hat{e}_{\pi,p}^{\varepsilon}(\{1\})$$

This shows (OA.17).

If $p^{\varepsilon,N} \in [x_{i-1}^N, x_i^N]$, then

$$e_{\pi,p}^{\varepsilon,N}([p^{\varepsilon,N}, 1]) \le e_{\pi,p}^{\varepsilon,N}([x_{i-1}^N, 1]) = \hat{e}_{\pi,p}^{\varepsilon}([x_{i-1}^N, 1]) \le \hat{e}_{\pi,p}^{\varepsilon}\left(\left[p^{\varepsilon,N} - \frac{1-\pi}{N}, 1\right]\right),$$

where the last inequality holds because $[x_{i-1}^N, x_i^N]$ has length $(1-\pi)/N$. Hence,

$$\limsup_{N\to\infty} e_{\pi,p}^{\varepsilon,N}([p^{\varepsilon,N}, 1]) \le \limsup_{N\to\infty} \hat{e}_{\pi,p}^{\varepsilon}\left(\left[p^{\varepsilon,N} - \frac{1-\pi}{N}, 1\right]\right) \le \hat{e}_{\pi,p}^{\varepsilon}([p,1])$$

because $x \mapsto \hat{e}_{\pi,p}^{\varepsilon}([x,1])$ is upper semicontinuous. On the other hand,

$$\liminf_{N\to\infty} e_{\pi,p}^{\varepsilon,N}([p^{\varepsilon,N}, 1]) \ge \hat{e}_{\pi,p}^{\varepsilon}([p,1])$$

because otherwise

$$\liminf_{N\to\infty} p^{\varepsilon,N} e_{\pi,p}^{\varepsilon,N}([p^{\varepsilon,N}, 1]) = p \liminf_{N\to\infty} e_{\pi,p}^{\varepsilon,N}([p^{\varepsilon,N}, 1]) < p\hat{e}_{\pi,p}^{\varepsilon}([p,1]),$$

which contradicts the optimality of $p^{\varepsilon,N}$. This shows (OA.18). Together, (OA.17) and (OA.18) imply

$$\lim_{N\to\infty} p^{\varepsilon,N} e_{\pi,p}^{\varepsilon,N}([p^{\varepsilon,N}, 1]) = p\hat{e}_{\pi,p}^{\varepsilon}([p,1]).$$

Letting ε go to zero concludes the proof of (OA.14):

$$\lim_{\varepsilon\to 0} \lim_{N\to\infty} p^{\varepsilon,N} e_{\pi,p}^{\varepsilon,N}([p^{\varepsilon,N}, 1]) = p g_{\pi}([p,1]) = \pi.$$

Next, we show (OA.15). For given $\varepsilon \in (0,1]$, the Dominated Convergence Theorem implies

$$\begin{aligned}\lim_{N\to\infty}\int_{p^{\varepsilon,N}}^{1} e^{\varepsilon,N}_{\pi,p}([x,1])\mathrm{d}x &= \int \lim_{N\to\infty}\mathbf{1}_{[p^{\varepsilon,N},1]}(x)e^{\varepsilon,N}_{\pi,p}([x,1])\mathrm{d}x \\ &= \int \mathbf{1}_{[p,1]}(v)\hat{e}^{\varepsilon}_{\pi,p}([x,1])\mathrm{d}x \\ &= \int_{p}^{1}\hat{e}^{\varepsilon}_{\pi,p}([x,1])\mathrm{d}x\end{aligned}$$

where the second equality holds by (OA.19) and (OA.17). Letting ε go to zero yields (OA.15).

Finally, we prove (OA.16). Fix $N \in \mathbb{N}$. For any $\varepsilon \in (0,1]$, we have

$$h^{\varepsilon,N}_{\pi,p}([x,1]) = 1 - \frac{1}{1-\lambda^{N}_{\pi,p}} f([0,x]) < f([x,1]), \quad \forall x \in [0,\pi).$$

Hence

$$\lim_{\varepsilon\to 0}\max_{x\in[0,\pi]} xh^{\varepsilon,N}_{\pi,p}([x,1]) < \max_{x\in[0,\pi]} xf([x,1]) \le \pi_0.$$

On the other hand, if $x \in [x^N_{i-1}, x^N_i]$, $i = 1,\dots,N-1$, then

$$h^{\varepsilon,N}_{\pi,p}([x,1]) = \frac{1}{1-\lambda^{N}_{\pi,p}} f([x,1]) - \frac{\lambda^{N}_{\pi,p}}{1-\lambda^{N}_{\pi,p}}\left(\frac{\hat{e}^{\varepsilon}_{\pi,p}([x^N_{i-1},x^N_i])}{f([x^N_{i-1},x^N_i])} f([x,x^N_i]) + \hat{e}^{\varepsilon}_{\pi,p}([x^N_i,1])\right).$$

By the Maximum Theorem,

$$\lim_{\varepsilon\to 0}\max_{x\in[x^N_{i-1},x^N_i]} xh^{\varepsilon,N}_{\pi,p}([x,1]) = \max_{x\in[x^N_{i-1},x^N_i]} xh^{0,N}_{\pi,p}([x,1]), \quad \forall i = 1,\dots,N-1.$$

Analogously,

$$\lim_{\varepsilon\to 0}\max_{x\in[x^N_{N-1},1]} xh^{\varepsilon,N}_{\pi,p}([x,1]) = \max_{x\in[x^N_{N-1},1]} xh^{0,N}_{\pi,p}([x,1]).$$

It remains to show that

$$\max_{x\in[\pi,1]} xh^{0,N}_{\pi,p}([x,1]) < \pi_0. \tag{OA.20}$$

Note that if $x \in \{x^N_0,\dots,x^N_{N-1}\}$, then

$$xe^{0,N}_{\pi,p}([x,1]) = xg_{\pi}([x,1]) = \pi > \pi_0 \ge xf([x,1])$$

and thus

$$e^{0,N}_{\pi,p}([x,1]) > f([x,1]).$$

Because the ratio

$$\frac{e^{0,N}_{\pi,p}([x,x_i^N])}{f([x,x_i^N])} = \frac{g_\pi([x_{i-1}^N,x_i^N])}{f([x_{i-1}^N,x_i^N])}$$

is the same for all $x \in (x_{i-1}^N, x_i^N)$, $i = 1,\dots,N$, it follows that

$$e^{0,N}_{\pi,p}([x,1]) > f([x,1]), \quad \forall x \in [\pi,1].$$

Using (OA.13), we conclude

$$h^{0,N}_{\pi,p}([x,1]) < f([x,1]), \quad \forall x \in [\pi,1],$$

and thus

$$x h^{0,N}_{\pi,p}([x,1]) < x f([x,1]) \leq \pi_0, \quad \forall x \in [\pi,1],$$

which implies (OA.20). □

We are now ready to prove Theorem 1.

Proof of Theorem 1. Let $(\pi,u) \in S$. Because the function $p \mapsto \int_p^1 g_\pi([x,1])\mathrm{d}x$ is continuous, there exists a price $p \in [\pi,1]$ such that

$$u = \int_p^1 g_\pi([x,1])\mathrm{d}x.$$

The function $\pi \mapsto \overline{u}(\pi) = -\pi \ln \pi$ is continuous. To prove the theorem, we may therefore assume $1 > \pi > \pi_0$, because if for any such π and any $u_\pi \in [0,\overline{u}(\pi)]$ there exists a sequence $\big((\pi_n,u_n)\big)_{n\in\mathbb{N}}$ converging to (π,u_π), then there also exists a sequence converging to (π,u) for $\pi \in \{\pi_0,1\}$ and any $u \in [0,\overline{u}(\pi)]$.

We now apply Lemma OA.3. For every $\varepsilon \in (0,1]$, let $(p^{\varepsilon,N})_{N\in\mathbb{N}}$ be a sequence of prices such that

$$p^{\varepsilon,N} \in \operatorname*{argmax}_x x e^{\varepsilon,N}_{\pi,p}([x,1]), \quad \forall N \in \mathbb{N}.$$

Then by Lemma OA.3,

$$\lim_{\varepsilon\to 0}\lim_{N\to\infty} p^{\varepsilon,N} e^{\varepsilon,N}_{\pi,p}([p^{\varepsilon,N},1]) = \pi,$$
$$\lim_{\varepsilon\to 0}\lim_{N\to\infty} \int_{p^{\varepsilon,N}}^1 e^{\varepsilon,N}_{\pi,p}([x,1])\mathrm{d}x = \int_p^1 g_\pi([x,1])\mathrm{d}x.$$

To prove the theorem, it therefore suffices to show that for any ε below some cutoff and for any N above some cutoff, there exists a sequence $\big((\pi_n, u_n)\big)_{n\in\mathbb{N}}$ such that $(\pi_n, u_n) \in S_n$ and

$$(\pi_n, u_n) \underset{n\to\infty}{\longrightarrow} \left(p^{\varepsilon,N} e^{\varepsilon,N}_{\pi,p}([p^{\varepsilon,N}, 1]), \int_{p^{\varepsilon,N}}^{1} e^{\varepsilon,N}_{\pi,p}([x,1])\mathrm{d}x\right).$$

Invoking Lemma OA.3, let ε be small enough and N big enough such that

$$p^{\varepsilon,N} e^{\varepsilon,N}_{\pi,p}([p^{\varepsilon,N}, 1]) \geq \pi_0 \geq \max_x x h^{\varepsilon,N}_{\pi,p}([x,1]).$$

Then, such a sequence $\big((\pi_n, u_n)\big)_{n\in\mathbb{N}}$ exists by Lemma OA.2. □

References

ACQUISTI, A., C. TAYLOR, AND L. WAGMAN (2016): "The Economics of Privacy," *Journal of Economic Literature*, 54(2), 442–492.

ACQUISTI, A., AND H. VARIAN (2005): "Conditioning Prices on Purchase History," *Marketing Science*, 24(3), 367–381.

ALIPRANTIS, C. D., AND K. C. BORDER (2006): *Infinite Dimensional Analysis*. Springer, Berlin, third edn.

ANDERSON, C. (2006): *The Long Tail: Why the Future of Business Is Selling Less of More.* Hyperion, New York.

ARMSTRONG, M., AND J. ZHOU (2011): "Paying for Prominence," *Economic Journal*, 121(556), F368–F395.

BERGEMANN, D., AND A. BONATTI (2015): "Selling Cookies," *American Economic Journal: Microeconomics*, 7(3), 259–94.

——— (2019): "Markets for Information: An Introduction," *Annual Review of Economics*, 11, 85–107.

BERGEMANN, D., B. BROOKS, AND S. MORRIS (2015): "The Limits of Price Discrimination," *American Economic Review*, 105(3), 921–957.

BLACKWELL, D. (1953): "Equivalent Comparison of Experiments," *Annals of Mathematical Statistics*, 24(2), 265–272.

BONATTI, A., AND G. CISTERNAS (2020): "Consumer Scores and Price Discrimination," *Review of Economics Studies*, 87(2), 750–791.

BOUNIE, D., A. DUBUS, AND P. WAELBROECK (2022): "Competition and Mergers with Strategic Data Intermediaries," *working paper*.

BRYNJOLFSSON, E., Y. J. HU, AND M. D. SMITH (2003): "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science*, 49(11), 1580–96.

CAVALLO, A. (2017): "Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers," *American Economic Review*, 107(1), 283–303.

CONDORELLI, D., AND B. SZENTES (2020): "Information Design in the Hold-Up Problem," *Journal of Political Economy*, 128(2), 681–709.

CONITZER, V., C. R. TAYLOR, AND L. WAGMAN (2012): "Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases," *Marketing Science*, 31(2), 277–292.

DE CORNIÈRE, A., AND R. DE NIJS (2016): "Online Advertising and Privacy," *RAND Journal of Economics*, 47(1), 48–72.

DE CORNIÈRE, A., AND G. TAYLOR (2019): "A Model of Biased Intermediation," *RAND Journal of Economics*, 50(4), 854–882.

DELLAVIGNA, S., AND M. GENTZKOW (2019): "Uniform Pricing in US Retail Chains," *Quarterly Journal of Economics*, 134(4), 2011–2084.

ECONOMIST (2021): "Shein Exemplifies a New Style of Chinese Multinational," *The Economist*, October 7, https://www.economist.com/business/shein-exemplifies-a-new-style-of-chinese-multinational/21805217.

FEDERAL TRADE COMMISSION (2014): "*Data Brokers: A Call for Transparency and Accountability*," Washington, DC.

HAGHPANAH, N., AND R. SIEGEL (2022a): "The Limits of Multi-Product Price Discrimination," *American Economic Review: Insights*, 4(4), 443–459.

——— (2022b): "Pareto Improving Segmentation of Multi-Product Markets," *working paper*.

HAGIU, A., AND B. JULLIEN (2011): "Why do Intermediaries Divert Search?," *RAND Journal of Economics*, 42(2), 337–362.

HIDIR, S., AND N. VELLODI (2021): "Privacy, Personalization, and Price Discrimination," *Journal of the European Economic Association*, 19(2), 1342–1363.

ICHIHASHI, S. (2020): "Online Privacy and Information Disclosure by Consumers," *American Economic Review*, 110(2), 569–95.

——— (2021): "Competing Data Intermediaries," *RAND Journal of Economics*, 52(3), 515–537.

INDERST, R., AND M. OTTAVIANI (2012): "Competition Through Commissions and Kickbacks," *American Economic Review*, 102(2), 780–809.

KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101(6), 2590–2615.

NOCKE, V., AND P. REY (2022): "Consumer Search and Choice Overload," *working paper*.

OECD (2018): "Personalised Pricing in the Digital Era – Note by the European Union," Directorate for Financial and Enterprise Affairs, Competition Committee, https://one.oecd.org/document/DAF/COMP/WD(2018)128/en/pdf.

PRAM, K. (2021): "Disclosure, Welfare and Adverse Selection," *Journal of Economic Theory*, 197, 105327.

QUAN, T. W., AND K. R. WILLIAMS (2018): "Product Variety, Across-Market Demand Heterogeneity, and the Value of Online Retail," *RAND Journal of Economics*, 49(4), 877–913.

SHAKED, M., AND J. G. SHANTHIKUMAR (2007): *Stochastic Orders*, Springer Series in Statistics. Springer New York.

TAYLOR, C. R. (2004): "Consumer Privacy and the Market for Customer Information," *RAND Journal of Economics*, 35(4), 631–650.

TEH, T.-H., AND J. WRIGHT (2022): "Intermediation and Steering: Competition in Prices and Commissions," *American Economic Journal: Microeconomics*, 14(2), 281–321.

VILLAS-BOAS, J. M. (2004): "Price Cycles in Markets with Customer Recognition," *RAND Journal of Economics*, 35(3), 486–501.