CONSUMER-OPTIMAL SEGMENTATION IN MULTI-PRODUCT MARKETS

By

Dirk Bergemann, Tibor Heumann and Michael C. Wang

December 2024

COWLES FOUNDATION DISCUSSION PAPER NO. 2420

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Consumer-Optimal Segmentation in Multi-Product Markets*

Dirk Bergemann† Tibor Heumann‡ Michael C. Wang§

December 22, 2024

Abstract

We analyze how market segmentation affects consumer welfare when a monopolist can engage in both second-degree price discrimination (through product differentiation) and third-degree price discrimination (through market segmentation). We characterize the consumer-optimal market segmentation and show that it has several striking properties: (1) the market segmentation displays monotonicity—higher-value customers always receive higher quality product than lower-value regardless of their segment and across any segment; and (2) when aggregate demand elasticity exceeds a threshold determined by marginal costs, no segmentation maximizes consumer surplus. Our results demonstrate that strategic market segmentation can benefit consumers even when it enables price discrimination, but these benefits depend critically on demand elasticities and cost structures. The findings have implications for regulatory policy regarding price discrimination and market segmentation practices.

JEL Classification: D42, D83, L12

Keywords: Price Discrimination, Nonlinear Pricing, Private Information, Second Degree Price Discrimination, Third Degree Price Discrimination, Pareto Distribution, Generalized Pareto Distribution, Bayesian Persuasion

*An early version of this paper working in a more limited setting appeared under the title "A Unified Approach to Second and Third Degree Price Discrimination." We acknowledge financial support from NSF grants SES-2001208 and SES-2049744. We have benefitted from many conversations and related joint work with Ben Brooks and Stephen Morris.

†Department of Economics, Yale University, dirk.bergemann@yale.edu

‡Pontificia Universidad Católica de Chile, tibor.heumann@uc.cl

§Department of Economics, Yale University, michael.wang.mcw75@yale.edu

Contents

1 Introduction

1.1 Motivation and Results

In the digital economy, firms increasingly segment their markets with unprecedented precision. Digital platforms adjust prices based on customer browsing history, streaming services offer differentiated subscription tiers, and airlines practice sophisticated yield management. While such practices traditionally raise concerns about consumer exploitation through price discrimination, their welfare implications remain ambiguous, particularly when firms can simultaneously adjust both prices and product qualities across different market segments.

The welfare implications of market segmentation have long interested economists. While third-degree price discrimination—charging different prices to distinct market segments—may either benefit or harm consumers, a similar ambiguity exists for second-degree price discrimination, where sellers screen consumers through quality-differentiated product menus. However, most research has analyzed these practices in isolation, leaving open the question of how they interact when deployed simultaneously, as is increasingly common in practice.

This paper provides a comprehensive analysis of consumer-optimal market segmentation when a monopolist can employ both forms of price discrimination. We characterize the segmentation strategies that maximize consumer welfare and show that they exhibit several surprising properties. First, despite the monopolist's ability to offer different qualities to identical consumers across segments, the optimal segmentation maintains uniform quality provision—consumers with the same value receive similar, and frequently the same quality level regardless of their segment, though they may pay different prices. Second, quality monotonicity is preserved both within and across segments— higher-value consumers always receive higher quality products. Third, when aggregate demand is sufficiently elastic relative to the seller's cost structure, no segmentation maximizes consumer surplus.

These findings significantly extend the work of Bergemann et al. (2015), who showed that an appropriate market segmentation can generate efficient allocation and maximize consumer surplus in unit demand settings. We demonstrate that with multiple products and quality differentiation, perfect efficiency is generally unattainable, but strategic segmentation can still substantially benefit consumers. Our analysis also complements recent work by Haghpanah and Siegel (2022), Haghpanah and Siegel (2023) by fully characterizing the consumer-optimal segmentation and identifying precise conditions under which segmentation improves consumer welfare.

The results have important implications for competition policy and regulation of price discrimination practices. They suggest that blanket restrictions on market segmentation may harm consumers by preventing welfare-enhancing price discrimination, while highlighting specific market conditions where segmentation is more likely to be beneficial. The findings also inform ongoing debates about big data and personalized pricing by showing how consumer heterogeneity and cost structures interact to determine optimal segmentation strategies.

We consider a seller who can engage simultaneously in second- and third- degree price discrimination. Our model consists of a monopolist that offers goods of varying quality to a continuum of buyers. The willingness-to-pay, the value is private information to each buyer and the seller only knows the distribution of values, which we refer to as the aggregate market. The seller may segment the market into submarkets, each with its own distribution of values subject only to the condition that the distribution of values across all submarkets must conform to the aggregate market. The monopolist offers an optimal pricing scheme in each submarket.

Before we proceed with describing our results it is convenient to briefly discuss how second- and third- degree price discrimination interact in our model. In a recent contribution, Haghpanah and Siegel (2022) showed already that it is impossible to implement the socially efficient surplus and allow the buyers to appropriate the gains from segmentation. Hence, there will be an inevitable trade-off between consumer surplus and social surplus. The intuition is that in the presence of second-degree price discrimination the monopolist will supply an inefficient quality, unless she can perfectly distinguish between buyers. When there is a single indivisible good for sale, this trade-off does not appear because it is suffices for the segmentation to induce the seller to not exclude any buyers, but there is no room for an inefficient quality supply.

We provide two types of results. First we describe the consumer surplus that can be attained by the consumer-optimal segmentation. Second, we provide properties of the markets that conform with this consumer-optimal segmentation. While our analysis focuses on the consumer-optimal segmentation, the analysis extends in a straightforward manner to situations in which the objective of interest is a linear combination of consumer surplus and profits, hence, the Pareto frontier that can be attained by any segmentation.

Our first main result characterizes the consumer surplus attained by the consumer-optimal segmentation (Theorem 1). In the adverse selection problem the consumer surplus corresponds to the buyers' information rents, and we can identify the contribution of each value to the consumer surplus by the product of inverse hazard rate and the marginal allocation and then take the sum (integral) over all values. We provide a convenient representation of this contribution as a function that depends only on the inverse hazard rate, we refer to this function as the *local informational rent*. Theorem 1 shows that the consumer surplus attained by the consumer-optimal segmentation is computed by twice modifying this formula. First, by taking the concavification of the local informational rents. Second, the local informational rents are not evaluated at the hazard rate of the aggregate market, but instead a distribution over hazard rates is found by solving a maximization problem over (quasi-)distributions that are majorized by the aggregate market. Interestingly, while the original problem consists of a maximization over segmentations, which are distributions over distributions of values, Theorem 1 yields a maximization over a single distribution of values. So Theorem 1 provides a much more tractable problem to find the consumer surplus generated by the consumer-optimal segmentation relative to the original problem.

Theorem 1 allows us to find necessary and sufficient conditions for the optimal segmentation to be no segmentation (see Proposition 4). We find that in many markets the allocation is inefficient but nonetheless any segmentation would (weakly) decrease consumer surplus even further. We also find that in many situations the distribution that solves the maximization problem is the aggregate market. For example, when the *marginal cost* is not too concave and the aggregate market satisfies the monotone hazard rate condition (Proposition 5). In these cases, the value is computed by taking the expectation of the concavification of the local information rents, (evaluated at the inverse hazard rates of the aggregate market as in the case without segmentation). Because of the concavification, the aggregated market itself however is not the consumer-optimal solution.

The second set of results relate to the properties of the markets that constitute the consumer-optimal segmentation. The consumer-optimal segmentation is shown to sort the consumer monotonically across all segments, see Proposition 3: (i) buyers with a given value may be offered different qualities across different segments, but (ii) the qualities all fall within a narrow bracket and these brackets are ordered monotonically and without overlap. Thus, a buyer with a higher willingness to pay will always receive a higher quality than a buyer with a lower willingness to pay, independent and across all possible market segments they find themselves in. Note that the consumption must be monotonically increasing in the value within a market due to the incentive compatibility constraint; monotonicity across markets arises only as a part of the optimization over segments.

While there might be dispersion in the quality consumed by a given value across different segments, this dispersion must be small, and it becomes negligible as the number of different values becomes large (Corollary 1). Hence, while not satisfied exactly, the consumer surplus introduces minimal dispersion in the quality consumed by any given value across segments. We can also find a lower bound on the quality consumed by any value, independent of the distribution of values in the aggregate market (Corollary 2).

After providing our general results, we focus on environments when the seller has a constant elasticity cost function. In this case, we find that the consumer-optimal segmentation has a particularly tractable form. There is a cutoff value that determines the *demand elasticity* in these markets. Demand elasticity of all market segments at values above this cutoff is the same elasticity as the aggregate market; demand elasticity of all the market segments at values below this cutoff have a constant elasticity determined by the *cost elasticity.* As the cost becomes more inelastic, the demand becomes more elastic. As the cost elasticity goes to infinity, we obtain the special case that the seller offers an indivisible good and the demand elasticity is unitary below the cutoff, thus recovering the consumer-optimal segmentation in Bergemann et al. (2015). Away from the limit, the optimal segmentation generates more inelastic demands, which increases the supply of the seller.

Finally, we move to analyze the effects of second and third degree price discrimination more broadly, not just in the consumer optimal segmentation. We extend the earlier analysis of the value of segmentation and establish that a weighted Bayesian persuasion problem can attain every point on the constrained efficient Pareto frontier of consumer and producer surplus (Theorem 2).

From a methodological perspective, our work provides novel insights into the study of third-degree price discrimination. While it has been acknowledged that the problem of third-degree price discrimination can be seen as a problem of Bayesian persuasion, previous work on third-degree price discrimination has largely *not* relied on concavification techniques. The reason for this is that the state space is the space of all demands, which typically has a large dimensionality (infinite dimensional when values are continuous), which in turn makes the concavification technique more difficult to apply. We show that one can in fact apply the concavification *pointwise* value-by-value, which allow us to use classic intuitions from one-dimensional persuasion problems. This allows us to characterize the value of the consumer-optimal segmentation and provide properties of all consumer-optimal segmentation. In Section 8, we then proceed to link the concavification argument to the arguments used in the literature and discuss our results in more detail.

1.2 Related Literature

Our results are related to a large literature on price discrimination, beginning with Pigou (1920) and now encompassing a wide range of research on the output and welfare implications of price discrimination, such as Robinson (1933); Schmalensee (1981); Varian (1985); Aguirre et al. (2010); Cowan (2012) and Bergemann et al. (2015). We use the classic model of second-degree price discrimination via quality and quantity differentiated products first presented by Mussa and Rosen (1978) and Maskin and Riley (1984). Johnson and Myatt (2003) consider a model of second degree price discrimination under monopoly and duopoly and provide conditions under which the monopolist will offer a single good.

The problem of extending the results of Bergemann et al. (2015) to a multi-product setting was analyzed earlier by Haghpanah and Siegel (2022) and Haghpanah and Siegel (2023). Haghpanah and Siegel (2022) show that when the optimal menu in the aggregate market consists of a menu of more than one item, thus a screening menu, then the consumer surplus maximizing allocation cannot attain the Pareto frontier and hence the full surplus triangle cannot be obtained (Theorem 1). Based on this insight, they provide a sufficient condition when the full surplus triangle will be attained, namely when all distribution over the values lead to the efficient single item menu (Theorem 2). By contrast, we provide necessary and sufficient conditions for there to be a segmentation that improves consumer surplus and provide the value of the consumer surplus maximizing allocation, whether it is efficient or not. Haghpanah and Siegel (2023) show that in generic markets there is always a segmentation relative to the single aggregate market that improves consumer surplus, however they do not provide properties of the consumer-optimal segmentation. In contrast to Haghpanah and Siegel (2023) (and (2022)), who work with a finite number of products, we allow for a continuum of qualities. Therefore, the results in Haghpanah and Siegel (2023) do not apply to our setting, and we find a large class of markets in which no segmentation is optimal for consumers. We completely solve for the consumer-optimal segmentation under continuous quality. Additionally, we derive some features of all Pareto efficient segmentations under both continuous and discrete quality. Our results also allow us to explicitly describe what the generic markets of Haghpanah and Siegel (2023) are, and to provide a sufficient condition under which their result fails to go through in the continuous quality limit.

Our characterization of extremal markets in the discrete quality case yields a family of distributions which solve the multi-unit generalization of the consumer surplus maximization problem considered by Condorelli and Szentes (2020). This problem and the family of distributions which underpin it are also related to the work of Roesler and Szentes (2017).

The organization of this paper is as follows. Section 2 introduces our model of non-linear pricing with market segmentation, thus integrating second and third degree price discrimination. Next, in Section 3, we characterize the consumer-optimal segmentation in a simple binary value environment, which nonetheless illustrates the main concepts we will use throughout the paper. Section 4 extends the solution to completely general conditions. In Section 5, we discuss what the consumer-optimal segmentation looks like in more detail for selected environments of economic interest. Section 6 consider the environment with constant elasticity cost function and obtains explicit solution of the optimal segmentation in terms of the demand elasticity in the segments. Section 7 extends our analysis to all Pareto efficient segmentations, and we provide a partial solution for the frontier of achievable surplus divisions. Section 8 discusses the relationship between the concavification approach pursued here and the analysis of extreme points in Bergemann et al. (2015), and Section 9 concludes. Appendix A contains omitted proof details. Appendix B links our analysis which is based on concavification with the previous literature on the single unit demand.

2 Setup

Payoffs and Pricing There is a monopolist and a continuum of consumers. The monopolist can produce a vertically differentiated good with quality

$$q \in \mathbb{R}_+.$$

The cost of producing a good of quality q is given by an increasing and convex function $c : \mathbb{R}_+ \to \mathbb{R}_+$. The monopolist posts a menu of prices $p(q) : \mathbb{R}_+ \to \mathbb{R}_+$ which specifies a price $p(q)$ for each offered quality q.

The consumer's gross surplus is their value v multiplied by the quality of the product q. The consumer purchases the quality which maximizes their net utility:

$$U(v,p) \triangleq \max_q \left[vq - p(q)\right], \tag{1}$$

and the corresponding quality choice is denoted by:

$$q(v) \triangleq \arg\max \left[vq - p(q)\right].$$

If multiple qualities maximize the utility of the buyer, ties are broken in favor of the seller. The profit of the seller from a buyer with value v and menu $p(q)$ is:

$$\Pi(v,p) \triangleq p(q(v)) - c(q(v)).$$

The seller does not know the value of any given buyer, but knows the distribution of values in a given market, as we explain next.

Markets A buyer's value is their private information (type). A *market* $x \in \Delta V$ is a probability distribution over values V that assigns probability $x(v)$ to $v \in V$. The values are drawn from a finite set:[1]

$$V = \{v_1, \ldots, v_k, \ldots, v_K\} \subset \mathbb{R}_+. \tag{2}$$

We denote by $D^x(v)$ the *demand function* associated with market x:

$$D^x(v) \triangleq \sum_{w \geq v} x(w). \tag{3}$$

As there is a bijection between a market x and its demand D^x, we frequently identify a market with its demand function. We denote by $x^* \in \Delta(V)$ the *aggregate market*, which is the distribution of values of all buyers present in the economy.

In a given market x, the profit-maximizing menu p^x in market x solves:

$$p^x \in \arg\max_{p(q)} \sum_{k=1}^{K} x(v_k)\Pi(v_k, p). \tag{4}$$

If there are multiple optimal price menus, we select the one which results in the highest consumer surplus. The aggregate consumer surplus in market x is given by:

$$U(x) \triangleq \sum_{k=1}^{K} x(v_k) U(v_k, p^x). \tag{5}$$

Segmentations The goal of this paper is to understand how profit and consumer surplus vary when consumers are divided into different submarkets, and the seller prices optimally within each one. That is, the seller engages in both second and third degree price discrimination simultaneously. Segments may be arbitrarily constructed, subject to the constraint that they aggregate together to the original market.

A *segmentation* σ is a finite distribution over markets $\sigma \in \Delta(\Delta V)$ such that

$$\sum_{x \in \text{supp}(\sigma)} \sigma(x) x = x^*,$$

where $\sigma(x)$ is the probability of market x and $\text{supp}(\sigma)$ is the support of σ, that is, the set of markets that have positive probability in this segmentation. We focus in particular on the *consumer surplus maximization problem*, which consists of finding the segmentation of

[1]The restriction to finite V is to ease the exposition technically; all results generalize naturally in the limit where V approaches a continuum.

x^* that generates the highest consumer surplus. Formally, we wish to solve:

$$\max_{\sigma \in \Delta(\Delta V)} \left[\sum_{x \in \Delta V} \sigma(x) U(x) \right] \text{ subject to } \sum_{x \in \text{supp}(\sigma)} \sigma(x) x = x^*. \tag{6}$$

We are interested in the value of this problem, as well as the segmentation that attains the maximum. We are first going to study the consumer-optimal segmentation and then study other segmentations that induce different surplus sharing between consumers and monopolist.

Notation Before we begin our analysis, we make some simplifications to our notation, and we explain how the notation is structured. Throughout, all subscripts refer to values, while superscripts refer to markets (for example, $x_k \triangleq x(v_k)$ is the probability of value v_k, while p^x is the optimal menu in market x and D^x is the demand of market x). To make notation more compact, a superscript "$*$" refers to the aggregate market x^* (so that p^* is the optimal price in the aggregate market x^*). Finally, all distribution over markets $\sigma \in \Delta(\Delta V)$ are assumed to satisfy the constraint in (6), so in later parts of the paper we write the problem of the consumer-optimal segmentation without these constraints.

3 The Binary Value Case

In this section, we characterize the consumer-optimal segmentation in a simple environment with binary values and constant elasticity cost functions. The problem in this environment is already sufficiently rich and allows us to introduce the central concepts and arguments that will lead to the solution of the general environment described next in Section 4.

Throughout this section, we suppose that there are only two values $0 < v_L < v_H < \infty$, which occur in the aggregate market with probability x_L^* and complementary probability $x_H^* = 1 - x_L^*$. The cost function is given by

$$c(q) = \frac{q^\gamma}{\gamma},$$

and thus isoelastic with cost elasticity $\gamma > 1$. We denote the inverse of the marginal cost by Q, provided that the argument value is positive:

$$Q(v) \triangleq \mathbb{I}\left[v \geq 0\right] c'^{-1}(v) = \mathbb{I}\left[v \geq 0\right] v^{\frac{1}{\gamma-1}}. \tag{7}$$

We refer to Q as the supply function, since $Q(v)$ is the quality that the monopolist would sell to a buyer of value v if the monopolist were to offer an efficient pricing scheme. The supply function takes the value 0 if the value is negative, which explains the indicator function $\mathbb{I}[v \geq 0]$ in the definition.

Roadmap The analysis proceeds as follows. We first provide the profit-maximizing pricing in a given market and derive the corresponding consumer surplus; this corresponds to the analysis found in Mussa and Rosen (1978) (focused on our binary-value setting). We then introduce a discrete version of the inverse hazard and show how to compute its distribution in any segmentation. The consumer surplus can be written completely in terms of the inverse hazard rate of the low value, and show that the consumer surplus is captured by the *local informational rent*. Next, we show that a bound on the consumer surplus attained by the consumer-optimal segmentation can be obtained by a persuasion problem where the planner chooses distributions over inverse hazard rates; it turns out that there are segmentations that can attain the same value as the persuasion problem. We conclude by analyzing how the consumer-optimal segmentation changes with the cost elasticity.

Optimal Screening and Consumer Surplus We consider the optimal menu offered by a profit-maximizing seller. This is classic problem analyzed by Mussa and Rosen (1978). The optimal allocation for a buyer with value v is determined by the first-order condition that balances virtual utility and cost. In the special case of binary values, and for any given market $x \in \Delta\{v_L, v_H\}$, the quality offered to the low value buyer is:

$$v_L - (v_H - v_L)\frac{x_H}{x_L} = c'(q_L). \tag{8}$$

As there is no distortion at the top, the high value buyer receives the efficient level:

$$v_H = c'\left(q_H\right).$$

The optimality condition (8) for the low value buyer, which states the marginal virtual utility is equal to the marginal cost, is the representative condition as we move to the many value environment. In the setting with finitely many values, we refer to the product term

$$h^x \triangleq (v_H - v_L)\frac{x_H}{x_L} = (v_H - v_L)\frac{1 - x_L}{x_L} \tag{9}$$

as the inverse hazard rate. It is the product of the increment between to adjacent values, and the ratio of upper cumulative probability and point probability. With binary values, the inverse hazard rate h^x of v_L is determined uniquely by the market x, that is, the probability

x_L (and $x_H = 1 - x_L$). We denote the hazard rate in the aggregate market by h^* :

$$h^* \triangleq (v_H - v_L)\frac{x_H^*}{x_L^*}.$$

The profit-maximizing allocation for a buyer with value v is determined by the first-order condition that balances virtual utility and cost:

$$q_L = Q\left(v_L - (v_H - v_L)\frac{D_H^x}{x_L}\right). \tag{10}$$

If the term inside function Q is negative, then the quality offered to the low value is zero (see (7)). We recall that in our binary value environment $D_H^x = x_H$; our notation choices are meant to make the analogies to the general case of multiple values more salient. Now, the consumer surplus in market x is earned by the high value v_H and is equal to:

$$U(x) = x_H(v_H - v_L)q_L.$$

This is the informational rent arising from the fact that v_H could pay v_L for quality q_L, which generates a surplus (for a buyer of value v_H) equal to $(v_H - v_L)q_L$.

Distribution Over Inverse Hazard Rates The inverse hazard rate h^x also represents the information rent that the buyer receives from any marginal unit of quality sold to the buyer. From the point of view of the seller, it therefore represents the *virtual cost* of selling to the low value buyer. Hence, the optimality condition (10) condition states that the marginal virtual utility $(v_L - h^x)$ is equal to the marginal cost $c'(q_L)$. We now show that for any segmentation σ, we can construct the corresponding distribution over hazard rates.

Consider any segmentation σ of the aggregate market x^* and denote its (finite) support by $\text{supp}(\sigma)$. With only two values, the feasibility constraint of the segmentation can be written simply in terms of the probability of the low value:

$$\sum_{x \in \text{supp}(\sigma)} \sigma(x)x_L = x_L^*.$$

We define for every x the weight that the market segment x has in the segmentation σ as:

$$\lambda^x \triangleq \frac{\sigma(x)x_L}{x_L^*} \in [0,1]. \tag{11}$$

Note that:

$$\sum_{x \in \text{supp}(\sigma)} \lambda^x = 1 \quad \text{and} \quad \sum_{x \in \text{supp}(\sigma)} \lambda^x h^x = h^*, \tag{12}$$

where we the equalities follow from the definitions of these objects. Hence, any segmentation σ induces a set of weights λ^x that are interpreted as a distribution over inverse hazard rates. What is remarkable is that the average hazard rate must equal the hazard rate in the aggregate market (second equation in (12)). While this last property is not general when there are many values, we will find an appropriate way to find the average hazard rate across all markets in a consumer-optimal segmentation.

Consumer Surplus and Local Informational Rents We will now show that we can rewrite the information rent entirely in terms of the hazard rate. We define a *local information rent*:

$$u_L(h) \triangleq h \cdot Q(v_L - h), \tag{13}$$

where the defining variable is now the inverse hazard rate h rather than the probability x. With this, we can write the consumer surplus generated by any segmentation σ as:

$$\sum_{x \in \text{supp}(\sigma)} \sigma(x) U(x) = \sum_{x \in \text{supp}(\sigma)} \lambda^x u_L\left(h^x\right). \tag{14}$$

We thus have that the consumer surplus generated by any segmentation is the expected value of $u_L(h)$.

We refer to $u_L(h)$ as the local information rent as it is the rent that all values above the local type v_L receive due the allocation to the low value buyer v_L. More generally, we will define a local information rent later for all intermediate values below the highest value buyer. The shape of the local informational rents is informative as to what kind of markets maximize consumer surplus. Given the isoelastic costs, $u_L(h)$ is concave whenever $h \leq v_L$ and attains a unique maximum at

$$h = \frac{\gamma - 1}{\gamma} v_L. \tag{15}$$

Figure 1 illustrates the behavior of the local information rent $u_L(h)$ associated with changes in the inverse hazard rate h. A high h tends to lower the information rent because the seller reduces the supply to the low value consumers, eventually excluding them altogether. On the other hand, when h becomes too small, then there are too few high value buyers to benefit from the informational rents generated by the low values. Hence, the pointwise surplus $u_L(h)$ is maximized at some interior value of h.

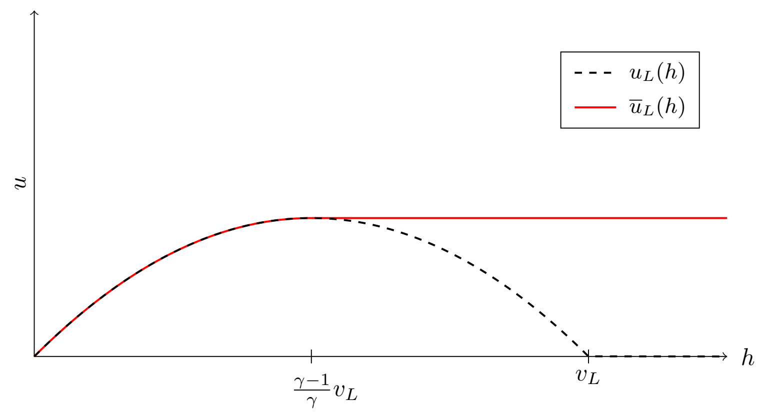

Figure 1: Local information rent $u_L(h)$ as a function of the hazard rate h.

The Segmentation Problem as Bayesian Persuasion We now transform the problem of determining an optimal segmentation into a Bayesian persuasion problem based on the local information rent as expressed by (14) and Bayes plausible distribution over inverse hazard rates h (see (12)). That is, we consider the Bayesian persuasion problem where we directly choose a distribution μ over hazard rates h:

$$\max_{\mu \in \Delta \mathbb{R}_+} \left[x_L^* \sum_{h \in \text{supp}(\mu)} \mu(h) u_L(h) \right] \text{ subject to } \sum_{h \in \text{supp}(\mu)} \mu(h) h = h^*. \tag{16}$$

Our analysis implies that every segmentation σ corresponds to a particular choice of μ, where the weight placed on each hazard rate is $\mu(h^x) = \lambda^x$.

Lemma 1 (Segmentation as One-Dimensional Bayesian Persuasion)
The consumer-optimal segmentation generates at most value (16)*. Furthermore, if a segmentation σ maximizes consumer surplus if, for every $x \in \text{supp}(\sigma)$,*

$$\lambda^x = \mu(h^x)$$

for a distribution $\mu \in \Delta \mathbb{R}_+$ that solves (16)*.*

We appeal to the standard Bayesian persuasion analysis to find the solution of (16). In particular, denote by $\overline{u}$ the concavification of u, that is, it is the smallest concave function that is pointwise larger than u. Using the analysis of the information rent following (13), the concavification of u is:

$$\overline{u}(h) \triangleq \begin{cases} u(h) & \text{if } h < \frac{\gamma-1}{\gamma}v_L; \\ u\left(\frac{\gamma-1}{\gamma}v_L\right) & \text{if } h \geq \frac{\gamma-1}{\gamma}v_L. \end{cases} \tag{17}$$

In Figure 1, $\overline{u}(h)$ is plotted as the solid red curve.

Following Kamenica and Gentzkow (2011), the solution to the Bayesian persuasion problem (16) is the value of $\overline{u}(h^*)$, namely the value of the concave envelope $\overline{u}$ at the inverse hazard rate of the aggregate market h^*. As the definition of $\overline{u}(h)$ suggests, if

$$h^* \leq \frac{\gamma-1}{\gamma}v_L, \tag{18}$$

the optimal value is attained at $\mu(h^*) = 1$. By contrast, if (18) is not satisfied, then the maximum is achieved with a binary distribution μ supported on $\widehat{h}, \widetilde{h}$:

$$\widehat{h} = \frac{\gamma-1}{\gamma}v_L, \quad \widetilde{h} = \infty. \tag{19}$$

Here $\widetilde{h} = \infty$ will mean that value v_L is not present in the second market.

From Persuasion to Segmentations We know that every segmentation corresponds to a particular choice of distribution μ over hazard rates in (16). The key question is if the converse holds—given the solution μ to (16), does there exist a segmentation σ which achieves it? The answer is yes, and we can construct this segmentation explicitly.

Proposition 1 (Optimal Segmentation—Binary Values)
The consumer-optimal segmentation σ attains the concavification bound of (16) *with equality:*

$$\sum_{x \in \text{supp}(\sigma)} \sigma(x)U(x) = x_L^* \overline{u}_L(h^*).$$

We can explicitly describe the (unique) consumer-optimal segmentation. The aggregate market x^* itself is the consumer-optimal segmentation if and only if (18) is satisfied. Otherwise, if (18) is not satisfied, the consumer-optimal segmentation is a binary segmentation supported on markets $\widehat{x}$, $\widetilde{x}$:

$$\widehat{x}_L = \frac{v_H - v_L}{v_H - v_L + \widehat{h}}, \quad \widetilde{x}_L = 0.$$

Naturally, the complements are given by $x_H = 1 - x_L$.

Notice that $\widehat{x}$ is constructed so that $h^{\widehat{x}} = \widehat{h}$, while $\widetilde{x}$ has the value v_L completely removed from its support, which we can think of as achieving $h^{\widetilde{x}} = \widetilde{h} = \infty$. Thus, h^x exactly matches the values of h in the support of μ. Since both σ and μ are binary, and every segmentation σ induces a feasible μ, it follows that the λ^x induced by this segmentation match $\mu(h^x)$.

Optimal Segmentation and Cost Elasticity We can also consider how the consumer-optimal segmentation varies with γ. As $\gamma \to \infty$, the model converges to a seller supplying an indivisible at cost 0: in this limit, the cost of supplying an infinitesimally small first unit is 0, and anything above that is infinitely costly. That is, in the limit we recover the same unit demand cost structure as Bergemann et al. (2015). The consumer-optimal segmentation given by Proposition 1 is as follows. If, in the aggregate market, the good is supplied to both consumers, then the aggregate market is efficient and this is the consumer-optimal segmentation. Instead, if the low type is excluded, the consumer-optimal segmentation generates two segments. In one of the segments the seller is left indifferent between supplying and not supplying the good to the low value. In the other market, the low value is not present, so the seller extract the full surplus from the high values.

Figure 2 illustrates this by plotting $u_L(h)$ for increasing values of γ. In the limit, $u_L(h)$ is maximized at $h = v_L$, and hence the inverse hazard rate is exactly the same as the value. So, if in the aggregate market $h_L^* > v_L$, then the market is segmented to create a segment in which the seller is exactly indifferent between supplying and not supplying the good.

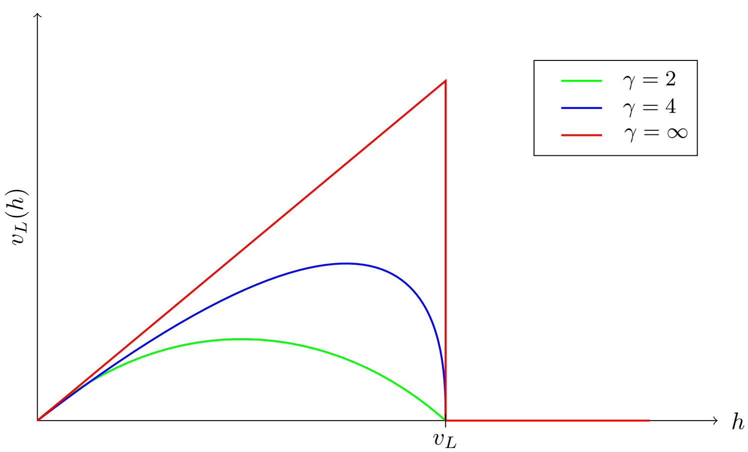

Figure 2: The local information rent $u_L(h)$ as a function of the cost elasticity γ.

When we study the model away from the limit, with $\gamma < \infty$, the problem becomes more subtle. As we saw in Proposition 1, the consumer-optimal segmentation does not lead to a socially efficient allocation for $\gamma < \infty$. In consequence, the consumer optimal segmentation balances allocative inefficiencies with informational rents. In particular, the frequency of low values has to be high enough to induce the producer to sell to low values, generating informational rents, but low enough that there are enough high values to benefit from these rents. The maximum of $u_L(h)$ gives the hazard rate at which this trade-off is exactly balanced. Since this maximum is interior, this means the resulting allocation is inefficient. Note that the shape of $u_L(h)$ does not depend on the aggregate market.

4 Consumer-Optimal Segmentations

In the previous section we consider a binary value model and showed that the consumer-optimal segmentation is equivalent to a one-dimensional Bayesian persuasion problem. With two values, this is not too surprising, as the market composition is described by a one-dimensional parameter. In this section, we show that, surprisingly, a similar logic carries over to the many value environment with general costs.

Specifically, we show that the problem of finding a consumer-optimal segmentation can be transformed into a series of $K-1$ one-dimensional Bayesian persuasion problems linked together by a single aggregate feasibility constraint. This is a significant simplification of the original maximization problem (6), which is a Bayesian persuasion problem over the $(K-1)$-dimensional simplex, and hence an infinite-dimensional optimization problem. The feasibility constraint captures which distributions of (average) hazard rates are feasible under *some* segmentation.

An additional hurdle relative to the previous section is the possibility of *support gaps*, meaning some market segments may not include every element of V in their support. These support gaps not only complicate the calculation of consumer surplus, but also allow the average hazard rate at any given value v_k to be higher or lower than it is in the aggregate market. This is in contrast to the binary value case, where the average hazard rate had to equal the aggregate hazard rate exactly. Our main challenge is then to write down an aggregate constraint which exactly pins down what sequences of hazard rates, and hence payoffs, are feasible.

4.1 Segmentation in Regular Markets

For any market x and any value v_k present in that market, $x_k > 0$, we define the gap at v_k, denoted by Δ_k^x, as the distance between v_k and the next highest value that is present in this market:

$$\Delta_k^x \triangleq \min\left\{v \in \operatorname{supp}(x) \mid v > v_k\right\} - v_k. \tag{20}$$

For completeness, if v_k is the maximum value present in market x, we define $\Delta_k^x = 0$. For any value v_k that is present in market x, we define the *discrete virtual value*

$$\phi_k^x \triangleq v_k - \Delta_k^x \frac{D_{k+1}^x}{x_k}. \tag{21}$$

A market is *regular* if ϕ_k^x is nondecreasing for all $v_k \in \text{supp}(x)$.

As in Section 3, we denote by Q the inverse of the marginal cost function:

$$Q(\phi) \triangleq \mathbb{I}[\phi \geq 0](c')^{-1}(\phi). \tag{22}$$

In regular markets, the quality supplied to every value is the supply function evaluated at the virtual value ϕ. If the virtual value is negative, then the good is not supplied, and hence the indicator function $\mathbb{I}[\phi \geq 0]$. q_k^x denotes the quality that buyer with value v_k consumes in market x under the profit-maximizing menu.

Lemma 2 (Supply in Regular Markets)
In a regular market x, the profit-maximizing menu supplies

$$q_k^x = Q(\phi_k^x), \text{ for every value } v_k.$$

We can express the consumer surplus in terms of the inverse hazard rates as in the previous section. For this, we define the inverse hazard rate at every value v_k and a given market x:

$$h_k^x \triangleq \Delta_k^x \frac{D_{k+1}^x}{x_k}, \tag{23}$$

and the *local information rent* for every value v_k :

$$u_k(h) \triangleq h \cdot Q(v_k - h). \tag{24}$$

This extends the notion of inverse hazard rate and local information rent that we encountered in the previous section in (9) and (13) from the binary to any arbitrary finite value environment. We note that the local information rent depends only on the local value v_k and the inverse hazard rate $h = h_k^x$ but not on the entire demand $D^x(\cdot)$. We can now express the consumer surplus in a compact way.

Lemma 3 (Consumer Surplus in Regular Markets)

In every regular market x, the consumer surplus is given by:

$$U(x) = \sum_{k=1}^{K-1} x_k u_k \left(h_k^x\right). \tag{25}$$

We thus represent the consumer surplus in a given regular market x by the weighted sum over the local information rents $u_k\left(h_k^x\right)$. That is, by appropriately extending the definition of the local information rent u we can write the consumer surplus in regular markets the same way as when the space of value had only two elements as in (14).

The expression of consumer surplus (25) above only holds for regular markets (possibly with gaps in the support). This restriction turns out to be without loss.

Lemma 4 (Consumer Surplus in Irregular Markets)
Every market x (possibly irregular) can be segmented into regular markets such that the seller sets the same prices for every good as in the original market x.

Hence, it is without loss of generality to restrict to segmentations which are supported only on regular markets.

4.2 Value of Segmentation

One of the main challenges in finding a consumer-optimal segmentation is that the variable of the optimization problem is a high-dimensional object: one needs to optimize over distributions over markets, but markets are in itself distributions over values. We will show that the main properties of the consumer-optimal distribution can be identified by solving a maximization problem over a single "quasi-market."

For any vector $D \in \mathbb{R}_+^K$, we write $D \prec D^*$ (and say that D is majorized by D^*) if:

$$\sum_{i=k}^{K-1}(v_{i+1} - v_i)D_{i+1} \leq \sum_{i=k}^{K-1}(v_{i+1} - v_i)D^*_{i+1} \text{ for all } 1 \leq k \leq K-1. \tag{26}$$

We do not require D to be decreasing in v_i, so in this sense we refer to D as a "quasi-market". If D were restricted to be decreasing, $\prec$ is exactly the *weak* majorization constraint studied in, for example, Kleiner et al. (2021). That is, (26) would be equivalent to D being a mean preserving spread of D^*. In the expressions we will frequently have a normalized version of the quasi-market, so to make the notation more compact, we define:

$$h_k^D \triangleq (v_{k+1} - v_k)\frac{D_{k+1}}{x_k^*}. \tag{27}$$

It is useful to optimize over D as it has a close interpretation with a classic order on distributions, with h^D being the inverse hazard rate associated with a particular market. However, note that h^D is not computed as a standard hazard rate as in the denominator is the probability of v_k in the aggregate market and not the one implied by D (which would be $D_k - D_{k-1}$). Of course, when $D = D^*$, then h^{D^*} is indeed the hazard rate of the aggregate market:

$$h_k^{D^*} = h_k^*.$$

As in Section 3, we denote the concavification of local information rent $u_k\left(h\right)$ by $\overline{u}_k\left(h\right)$. We now present an upper bound on the consumer surplus that can be attained by any segmentation σ in terms of a concavification bound expressed for the aggregate market represented by D^*. The upper bound is formed by a maximization problem over the concavified local information rents $\overline{u}_k\left(h\right)$:

$$\max_{D \prec D^*} \sum_{k=1}^{K-1} x_k^* \overline{u}_k\left(h_k^D\right). \tag{28}$$

We then show that this upper bound can indeed be obtained by the consumer-optimal segmentation σ and that the concavified maximization problem allows us to construct the optimal segmentation σ. For a regular single market, say x^*, it is clear that the above expression is an upper bound. After all, by Lemma 3 we have that

$$U\left(x\right) = \sum_{k=1}^{K-1} x_k^* u\left(h_k^*\right) \leq \sum_{k=1}^{K-1} x_k^* \overline{u}_k\left(h_k^*\right),$$

since the concavified local information rent $\overline{u}_k\left(h\right)$ is weakly higher than the local information rent $u\left(h\right)$ everywhere. But the relaxation offered in (29) goes further by allowing a maximization over all majorized vectors D rather than an evaluation just at D^*. Nonetheless, we will show that the optimal segmentation can attain this twice relaxed upper bound.

Theorem 1 (Value of Segmentation)
The consumer-optimal segmentation σ attains the concavification bound:

$$\max_{\sigma \in \Delta(\Delta V)} \sum_{x \in \text{supp}(\sigma)} \sigma(x) U(x) = \max_{D \prec D^*} \sum_{k=1}^{K-1} x_k^* \overline{u}_k\left(h_k^D\right). \tag{29}$$

The theorem provides an expression for the consumer surplus in an optimal segmentation in terms of a maximization problem over a quasi-market D. We described before the statement of the theorem in which sense the concavification bound clearly represents a relaxed problem for a single market. The intriguing part of the bound is that it is stated in terms of a *single quasi-market* D while the consumer surplus on the LHS is obtained by a distribution $\sigma(x)$ over *many markets.*

The optimization problem on the RHS consists of a concave objective function, as it is the weighted sum of concavified functions, and a set of linear constraints. Thus, it can be solved using standard techniques. The maximization problem states that the value of the consumer optimal segmentation can be attained by the expectation over $K-1$ separate concavification problems, each denoted by $\overline{u}_k(h)$ for all $k = 1, ..., K-1$. The solution can be decomposed into $K-1$ local information rent problems because with regular distributions, the allocation problem for each value v_k in each market segment x can be solved independently of all the other values. This requires the earlier result of Lemma 4 that allows us to focus on regular markets.

When D^* solves (29), then the consumer-optimal segmentation is obtained by taking the expected value of $\overline{u}(h_k^*)$. This is precisely the same result we had for two values in Section 3, except we now take the expectation value by value. In Section 5 we provide conditions for D^* to be a solution of (29) (in fact, we show there is a large class of models for which this is the case). However, in general, we obtain the value of the consumer-optimal segmentation not by taking the expectation using the hazard rates in the aggregate market but by computing an optimal hazard rate. We next discuss the proof of this theorem and in the following section provide more intuition about the value of the consumer-optimal segmentation when D^* is not a solution to (29).

The proof of Theorem 1 proceeds in two major steps. First, we establish that (29) is an upper bound. Second, we show that the bound is tight by explicitly constructing a segmentation which achieves the bound. We provide the proof of the first step next as it illustrates how the concavification and the majorization constraints appear in the analysis. The second step of the proof is relegated to the Appendix. Instead, in the next subsection we provide some intuition for the structure of the optimal segmentation.

Proof (Upper Bound). We first show that the right-hand-side of (29) is an upper bound for the consumer surplus attained by the optimal segmentation. Before we proceed with the proof, we redefine the inverse hazard rate as follows:

$$h_k^x = \begin{cases} \Delta_k^x \cdot \frac{D_{k+1}^x}{x_k} & \text{if } x_k > 0; \\ v_k & \text{if } x_k = 0. \end{cases}$$

This will prevent us from having inverse hazard rates that are indeterminate, and since the supply is 0 regardless of whether the inverse hazard rate is infinite or v_k both definitions lead to the same analysis. Following (25), we get that

$$\sum_{x\in\text{supp}(\sigma)} \sigma(x)U(x) = \sum_{x\in\text{supp}(\sigma)} \sigma(x) \left[\sum_{k=1}^{K-1} x_k u_k(h_k^x)\right] = \sum_{k=1}^{K-1} \left[\sum_{x\in\text{supp}(\sigma)} \sigma(x) x_k u_k(h_k^x)\right].$$

We now define:

$$\lambda_k^x \triangleq \frac{\sigma(x)x_k}{x_k^*},$$

so we can rewrite the above as

$$\sum_{x\in\text{supp}(\sigma)} \sigma(x)U(x) = \sum_{k=1}^{K-1} x_k^* \left[\sum_{x\in\text{supp}(\sigma)} \lambda_k^x u_k(h_k^x)\right].$$

Note that:

$$\sum_{x\in\text{supp}(\sigma)} \lambda_k^x = \frac{1}{x_k^*} \sum_{x\in\text{supp}(\sigma)} \sigma(x)x_k = 1,$$

so the weights λ_k^x together form a distribution over hazard rates. We correspondingly define the average hazard rate over σ at value v_k by

$$h_k^\sigma \triangleq \sum_{x\in\text{supp}(\sigma)} \lambda_k^x h_k^x.$$

Following Jensen's inequality and the fact that $u(h) \le \overline{u}(h)$ for all h, we have:

$$\sum_{x\in\text{supp}(\sigma)} \sigma(x)U(x) \le \sum_{k=1}^{K-1} x_k^* \overline{u}_k(h_k^\sigma). \tag{30}$$

We can thus provide an upper bound on consumer surplus using the concavification of u and the average hazard rate.

The presence of support gaps introduces the possibility that the average hazard rates $h_k^\sigma \neq h_k^*$, unlike in Section 3. We need to characterize the space of feasible sequences h_k^σ. Construct a "quasi-market" D^σ which would be consistent with h_k^σ for a full support distribution:

$$D_{k+1}^\sigma \triangleq \frac{x_k^* h_k^\sigma}{v_{k+1} - v_k}.$$

Again, D^σ is not a true demand function because it is not necessarily monotone.

An important property of D^σ is that:

$$D^\sigma \prec D^*. \tag{31}$$

We now prove this inequality is satisfied. We first note that:

$$\sum_{i=k}^{K-1}(v_{i+1}-v_i)D^{\sigma}_{i+1} = \sum_{i=k}^{K-1}\sum_{x\in\text{supp}(\sigma)}\sigma(x)x_k h^x_k$$
$$=\sum_{i=k}^{K-1}\sum_{x\in\text{supp}(\sigma)}\sigma(x)\Delta^x_i D^x_{i+1}\mathbf{1}[x_i>0].$$

Here once again we use that $x_i h^x(v_i)=0$ when $x_i=0$. But, if value v_k is present in market x ($x_k>0$), then

$$\sum_{i=k}^{K-1}\Delta^x_k D^x_{i+1}\mathbf{1}[x(v_i)>0] = \sum_{i=k}^{K-1}(v_{i+1}-v_i)D^x_{i+1}, \tag{32}$$

Thus:

$$\sum_{i=k}^{K-1}(v_{i+1}-v_i)D^{\sigma}_{i+1} = \sum_{x\in\text{supp}(\sigma)}\sigma(x)\sum_{i=k}^{K-1}\mathbf{1}[x_i>0](v_{i+1}-v_i)D^x_{i+1}.$$

We conclude that:

$$\sum_{i=k}^{K-1}(v_{i+1}-v_i)D^{\sigma}_{i+1} \le \sum_{x\in\text{supp}(\sigma)}\sigma(x)\sum_{i=k}^{K-1}(v_{i+1}-v_i)D^x_{i+1} = \sum_{i=k}^{K-1}(v_{i+1}-v_i)D^*_{i+1}, \tag{33}$$

which implies that (31) is satisfied.

Finally, to conclude the proof, we note that (30) and (31) imply that right-hand-side of (29) is an upper bound for the consumer surplus attained by the optimal segmentation. ∎

We thus proved that (29) is an upper bound. Proving that this bound is tight, i.e. that there exists a segmentation over regular markets which achieves it, is more complicated, and we relegate the details to the Appendix. Instead, in the next subsection, we provide the basic elements for the construction in the proof and explain the different ways in which segmentations can improve consumer surplus.

4.3 Two Forms of Persuasion

The proof that the concavification bound can be attained by an optimal segmentation σ has two parts. In the first part, we show that it is possible to construct a segmentation σ in which in every market $x\in\text{supp}(\sigma)$ the inverse hazard rate of the distribution x at every value v_k in this market is:

$$h^x_k = h^D_k,$$

where D solves (29). Since, in general, D will differ from D^*, this requires constructing segmentations where the average hazard rates differ from the aggregate market. To change the hazard rate, we need to introduce gaps in the segments: that is, there will be markets where $x_k > 0$ and $x_{k+1} = 0$. This allows to increase the hazard rate of high values at the expense of the hazard rate at low values.

The second part of the proof consists in segmenting the markets at every value v_k where the local information rent $u_k(h_k)$ lies below its concavification $\overline{u}_k(h_k)$. In this second step, we do not introduce gaps, but instead introduce variation across markets in the hazard rates at a given value. However, the average hazard rate stays the same as in the first step.

We refer to the segmentation we produce in the first part of the proof as *between-value persuasion*: it consists of changing the distribution of average hazard rates across values. By contrast, we refer to the segmentation in the second part of the proof as *within-value persuasion*: it consists of changing the distribution of the hazard rates across market segments while keeping the average hazard rate of a value constant.

When D^* is a solution to (29), we can then say there is only within-value persuasion; when the solution to (29) is $D \neq D^*$ and

$$\overline{u}_k(h_k^D) = u_k(h_k^D) \text{ for all } v_k,$$

then there is only between-value persuasion. We illustrate this with two different examples.[2] We begin with an example in which the aggregate market D^* solves (29) so there is only within-value persuasion.

Example 1 (Within-Value Persuasion). Suppose there are two values $V = \{1, 2\}$ and the cost function is:

$$c(q) = \begin{cases} 0 & \text{if } q \in [0,1]; \\ \frac{3}{4}(q-1) & \text{if } q \in [1,2]; \\ \infty & \text{if } q > 2. \end{cases}$$

This model can be interpreted as a seller that can supply a unit of the good at 0 cost and can supply a second unit at cost $\frac{3}{4}$.[3] We can solve this model the same way as in Section 3, but appropriately changing the u function, which we illustrate in Figure 3. If $h_1^* > 1$ we get

[2]In Section 3, we studied markets with binary support. We found that the aggregate market was segmented into two markets; in one market, type v_L consumes a positive quality, while in the other market,v_L is not present. With two values, D^* always solves (29), so there is never between-value persuasion. Additionally, there is no variation in the quality supplied to different values. In this sense, the example is too simple to illustrate the construction of the consumer-optimal segmentation effectively.

[3]We interpret qualities $q \in (0,1)$ as a seller that offers 1 unit with probability q and 0 units with probability $(1-q)$. Qualities $q \in (1,2)$ can be interpreted in an analogous way.

the same optimal segmentation as when $c(q) = q^\gamma/\gamma$ and we took the limit $\gamma \to \infty$ (in this case, for the purpose of the consumer-optimal segmentation, it is irrelevant that the seller can supply a second good).

If $h_1^* \in (1/4, 1)$ then D^* is the unique solution to (29). In this case, the consumer-optimal segmentation segments the aggregate market into two segments; in one segment the hazard rate will be $1/4$ (and thus the low value is supplied both units of the good), and in the other segment the hazard rate will be 1 (and thus the low value is supplied only one unit of the good). This clearly provides higher consumer surplus than the aggregate market as sometimes the low value is supplied both units of the good, which increases the consumer surplus from high values. We can then see how providing some variability across the hazard rates of a specific value across markets and help increase consumer surplus.

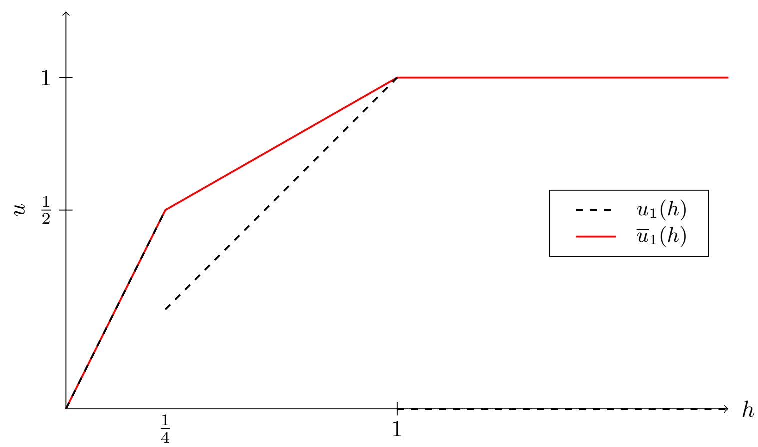

Figure 3: u_1 and $\overline{u}_1$ for Example 1.

Next, we provide an example where there is only between-value persuasion, and show how it can improve consumer surplus.

Example 2 (Between-Value Persuasion). Suppose the cost function is $c(q) = q^2/2$, there are three values $V = \{1, 2, 3\}$ and the aggregate market is:

$$x_1^* = \frac{17}{24};\ x_2^* = \frac{1}{8};\ x_3^* = \frac{1}{6}.$$

The hazard rates are:

$$h_1^* = \frac{x_2^* + x_3^*}{x_1^*} = \frac{7}{17}; \quad h_2^* = \frac{x_3^*}{x_2^*} = \frac{4}{3}; \quad h_3^* = 0.$$

Note that the inverse hazard rate is increasing between values 1 and 2, but the distribution is regular.

The solution to (29) is:

$$h_1^D = \frac{1}{2} \text{ and } h_2^D = 1.$$

Hence, these will be the average hazard rates in the aggregate market. Since we have that:

$$u_1\left(\frac{1}{2}\right) = \overline{u}_1\left(\frac{1}{2}\right) \text{ and } u_2(1) = \overline{u}_2(1),$$

there will be no within-value segmentation. In particular, in this example, a consumer-optimal segmentation consists of the following two markets:

$$\widehat{x}_1 = \frac{40}{59};\ \widehat{x}_2 = \frac{760}{4661};\ \widehat{x}_3 = \frac{741}{4661};$$

$$\widetilde{x}_1 = \frac{80}{99};\ \widetilde{x}_2 = 0;\ \widetilde{x}_3 = \frac{19}{99}.$$

The weights on the markets are $\sigma(\widehat{x}) = 4661/6080$ and $\sigma(\widetilde{x}) = 1419/6080$.

We plot the u functions, their concavifications, and the inverse hazard rates in Figure 4. Observe that in this example, the solid dots (the hazard rates at the consumer-optimal segmentation) are shifted away from the white dots (the hazard rate in the aggregate market).

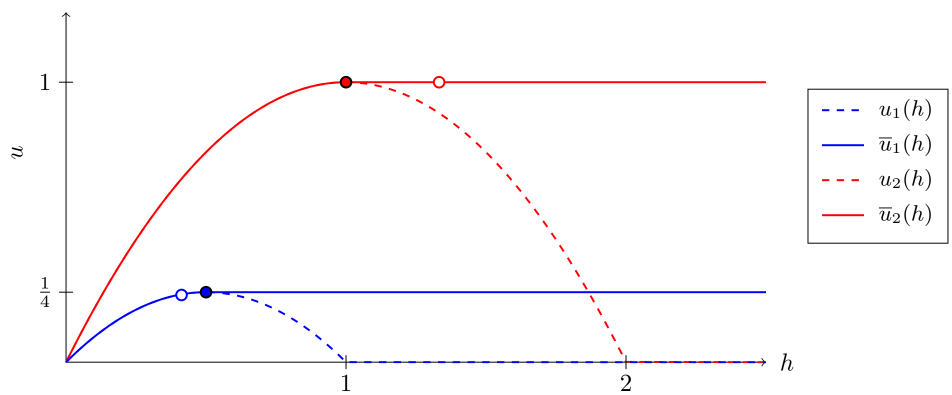

Figure 4: u_k and $\overline{u}_k$ for Example 2.

4.4 Properties of the Consumer-Optimal Segmentation

Although we explicitly construct a market segmentation that attains the upper bound of (29), the consumer-optimal segmentation that we obtain by construction is not easy to express in closed form. However, we can find a sharp description of the qualities that buyers will consume in this segmentation. Furthermore, the properties we provide next hold across all consumer-optimal segmentations, not just the one we obtain from following the procedure described in the proof of Theorem 1.

We first characterize the hazard rates of the demands in the support of a consumer-optimal segmentation in terms of a solution to (29). We denote by $\text{supp}(\overline{u}_k(h))$ the support of the concavification of u at h:

$$\text{supp}(\overline{u}_k(h)) \triangleq \{h' \in \mathbb{R}_+ : u_k(h') = \overline{u}_k(h') \text{ and } \overline{u}_k(\omega h + (1-\omega)h') \text{ is linear in } \omega \in [0,1]\}.$$

In other words, the support consists of the hazard rates h' where u_k and $\overline{u}_k$ coincide, and $\overline{u}_k$ is linear on the interval between h and h'.

Proposition 2 (Properties of Consumption)
A segmentation σ solves (6) if and only if there exists some D solving (29) such that for every $x \in \text{supp}(\sigma)$ and $v_k \in \text{supp}(x)$,

$$h_k^x \in \text{supp}\left(\overline{u}_k\left(h_{k+1}^D\right)\right), \tag{34}$$

and, for all k,

$$\sum_{\{x|x_k>0\}} \sigma(x) h_k^x = h_k^D. \tag{35}$$

This proposition provides general properties of the hazard rates of the demand at any given value v_k in a consumer-optimal segmentation in terms of the solution to (29). The expected hazard rate is h_k^D (see (35)) and the hazard rates are in the support of $\overline{u}$ (see (34)). We recall that in any market buyer v_k will consume quality:

$$q = Q(v_k - h_k^x).$$

So, we can now translate these properties about hazard rates to properties about the qualities consumed by buyers in a consumer-optimal segmentation.

Let q_k^σ denote the qualities consumed by value v_k in some market of segmentation σ :

$$q_k^\sigma \triangleq \left\{q \in \mathbb{R}_+ \mid q_k^x = q \text{ for some } x \in \text{supp}(\sigma)\right\}.$$

Proposition 3 (Monotonicity of Quality)
Let σ be a consumer-optimal segmentation. Then, for all k,

$$\max\{q : q \in q_k^\sigma\} \leq \min\{q : q \in q_{k+1}^\sigma\}.$$

That is, the quality consumption of different values is totally ordered between segments.

This proposition follows from the proof of Theorem 1 and the result that every h_k^x must be in the support of $\overline{u}_k(h_k^\sigma)$. That consumption is monotone within segments is immediate from incentive compatibility. However, the fact that monotonicity also holds across segments is a special property of the consumer-optimal segmentation.

We can then provide a limit on how much dispersion there is in the consumption across values for any given value. We denote by

$$d_k^\sigma \triangleq \max\{q : q \in q_k^\sigma\} - \min\{q : q \in q_k^\sigma\},$$

the dispersion of consumption in segmentation σ for value v_k. This is the difference between the maximum and minimum quality purchased by value v_k across all markets in segmentation σ. We denote $\overline{Q}$ the maximum efficient quality supplied to any value:

$$\overline{Q} \triangleq \arg\max_{q \in \mathbb{R}_+} \big\{ v_K q - c(q) \big\} < \infty.$$

Since the seller will never supply an inefficiently high quality, the quality supplied to every value will be below $\overline{Q}$. We now use these definitions to bound the consumption dispersion.

Corollary 1 (Quality Dispersion)
For any consumer-optimal segmentation and $\epsilon > 0$, there exists at most $\overline{Q}/\epsilon$ different values v_k such that

$$d_k^\sigma > \epsilon.$$

Hence, for any fixed level of dispersion ϵ, there is a bound on the number of values that has a dispersion larger than ϵ. Importantly, the bound does not depend on the aggregate market. An immediate implication is that, if we approximate a absolutely continuous distribution with a limit of increasingly finer discrete distributions, then in the limit almost every type will consume a unique quality. In the following section, we provide conditions such that the spirit of the result is satisfied exactly, that is, each buyer consumes only one quality across all markets in a consumer-optimal segmentation.

Finally, Proposition 2 allows us to obtain a upper bound on the prices of any quality level in any consumer-optimal segmentation. For this, first define:

$$\overline{h}_k = \arg\max_h \big[u_k(h) \big] \tag{36}$$

Following (34) we have that value v_k always buys a quality weakly larger than:

$$\underline{q}_k \triangleq Q(v_k - h_k).$$

We formalize this in the following corollary.

Corollary 2 (Minimum Quality)
In any consumer-optimal segmentation σ*, any buyer of type* v_k *present in market* $x \in \operatorname{supp}(\sigma)$ *consumes at least quality* $\underline{q}_k$*. Furthermore, the bound is tight: there exists an aggregate market under which no segmentation is optimal and every value* v_k *consumes* $\underline{q}_k$*.*

We can thus bound the size of the inefficiencies in any consumer-optimal segmentation. The surprising aspect of the result is that the bound can be derived using only the value of a buyer and the cost function.

5 Computing the Value of Segmentation

The results in the previous section provided a characterization of the value of the consumer-optimal segmentation and provided properties of the qualities consumed by different buyers in the optimal segmentation. However, in this characterization, the value of the consumer-optimal segmentation is expressed as a solution to a maximization problem. This makes it difficult to gauge when the gains can be large or small, or even when no segmentation is indeed the consumer-optimal segmentation. We thus now provide conditions under which the value of segmentation can be characterized more sharply. First, we provide sufficient conditions for D^* to be the solution of (29). These conditions simplify the calculation of the value of the consumer-optimal segmentation. However, even when D^* solves (29), segmentation may improve the consumer surplus through within-value persuasion (for example, as in Section 3). Hence, we also characterize when no segmentation is optimal.

We begin by providing conditions in terms of the local value of segmentation u and the hazard rate in the aggregate market h_k^*. These conditions will be easy-to-verify and help characterize the value of segmentation. We then provide conditions on the cost function and the distribution of values, which will be less general but easier to interpret.

5.1 When No Segmentation is Optimal

First, we provide a condition which guarantees D^* is a solution to (29).

Lemma 5 (No Between-Value Segmentation)
The solution to (29) *is* D^* *if and only if* $\overline{u}_k'(h_k^*)$ *is nondecreasing. In particular, in this case, the consumer-optimal segmentation generates:*

$$\max_{\sigma \in \Delta(\Delta V)} \sum_{x \in \operatorname{supp}(\sigma)} \sigma(x) U(x) = \sum_{k=1}^{K-1} x_k^* \overline{u}_k\left(h_k^*\right). \tag{37}$$

Proof. (⇐) Consider any segmentation σ and recall that $D^\sigma \prec D^*$. The difference between the RHS of (29) at D^σ and at D^* can be bounded as

$$\begin{aligned}\sum_{k=1}^{K-1} x_k^*(\overline{u}_k(h_k^\sigma) - \overline{u}_k(h_k^*)) &\leq \sum_{k=1}^{K-1} x_k^*(\overline{u}_k'(h_k^*)(h_k^\sigma - h_k^*)) \\ &= \sum_{k=1}^{K-1} \overline{u}_k'(h_k^*)(v_{k+1} - v_k)(D_{k+1}^\sigma - D_{k+1}^*)\end{aligned}$$

where the inequality uses the fact that $\overline{u}_k$ is concave. Following Theorem 2 of Fan and Lorentz (1954) we have that:

$$\sum_{k=1}^{K-1} \overline{u}_k'(h_k^*)(v_{k+1} - v_k)(D_{k+1}^\sigma - D_{k+1}^*) \leq 0,$$

which proves the result. To be precise in the use of Theorem 2 of Fan and Lorentz (1954), we express our terms in their notation. We have the function

$$\Phi(z, k) = -c_k z_k$$

which is linear and submodular in (z_k, k). The constants $c_k = \overline{u}_k'(h_k^*)$ are nondecreasing and $D^\sigma \prec D^*$, and hence

$$\sum_{j=1}^{K-1} \Phi((v_{j+1} - v_j)D_{j+1}^*, j) \leq \sum_{j=1}^{K-1} \Phi((v_{j+1} - v_j)D_{j+1}^\sigma, j).$$

Multiplying this inequality by -1 yields

$$\sum_{k=1}^{K-1} \overline{u}_k'(h_k^*)(v_{k+1} - v_k)(D_{k+1}^\sigma - D_{k+1}^*) \leq 0$$

proving the lemma.

(⇒) In the proof of Theorem 1, we show that for any solution D of (29), $u'(h_k^D)$ is nondecreasing; see discussion surrounding (A.9). ∎

The proof basically uses the concavity of $\overline{u}$ and a classic inequality for majorization constraints. The use of inequalities that rely on majorizations has been used recently by Kleiner et al. (2021). Unlike these papers, we do not have a monotonicity constraint (D^σ might be nondecreasing), but we can still use Theorem 2 in Fan and Lorentz (1954). Hence, the essence of the result is the same.

We now characterize when no segmentation is optimal. For this, we define:

$$O = \Big\{ x \in \Delta V : u_k(h_k^x) = \overline{u}_k(h_k^x) \text{ for all } k, \text{ and } u_k'(h_k^x) \text{ is nondecreasing in } k \Big\}.$$

This exactly characterizes the set of markets where no segmentation is optimal.

Proposition 4 (No Segmentation)
The consumer-optimal segmentation is no segmentation if and only if $x^* \in O$.

Proof. ($\Leftarrow$). Suppose $x^* \in O$. Then, following Lemma 5, we have that D^* solves (29). Furthermore, since $u_k(h_k^*) = \overline{u}_k(h_k^*)$,

$$\max_{D \prec D^*} \sum_{k=1}^{K-1} x_k^* \overline{u}_k(h_k^D) = \sum_{k=1}^{K-1} x_k^* \overline{u}_k(h_k^*) = \sum_{k=1}^{K-1} x_k^* u_k(h_k^*).$$

The second equality follows from the definition of O. Thus, no segmentation is optimal.

($\Rightarrow$). If the first condition is not satisfied, then:

$$\sum_{k=1}^{K-1} x_k^* u_k(h_k^*) < \sum_{k=1}^{K-1} x_k^* \overline{u}_k(h_k^*) \leq \max_{D \prec D^*} \sum_{k=1}^{K-1} x_k^* \overline{u}_k(h_k^D),$$

contradicting that no segmentation is optimal. If the second condition is not satisfied, then by Lemma 5,

$$\sum_{k=1}^{K-1} x_k^* u_k(h_k^*) \leq \sum_{k=1}^{K-1} x_k^* \overline{u}_k(h_k^*) < \max_{D \prec D^*} \sum_{k=1}^{K-1} x_k^* \overline{u}_k(h_k^D),$$

which again contradicts that no segmentation is optimal. ∎

We illustrate the set O in Figure 5 for when $c(q) = q^2/2$ and $V = \{3, 4, 5\}$. Since there are three values, the simplex is a subset of $\mathbb{R}_+^2$. The three vertices of the triangles represent the three markets in which only one of the values is present. Then, any point in the triangle represents a market where the composition of values is represented by the distance to the vertices. In this graph, we also illustrate the set of regular distributions that have positive virtual values (that is, ϕ_k^x is increasing in k and non-negative for all $k \in \{1, ..., K\}$). For reference, these are the distributions under which there is no exclusion of any value and there is full separation. We can see there is a large class of markets under which no segmentation is optimal. This is despite the fact that in all these markets the allocation is inefficient. However, we find that in some markets, segmentation can improve consumer surplus even when the distribution is regular and there is no exclusion of any values.

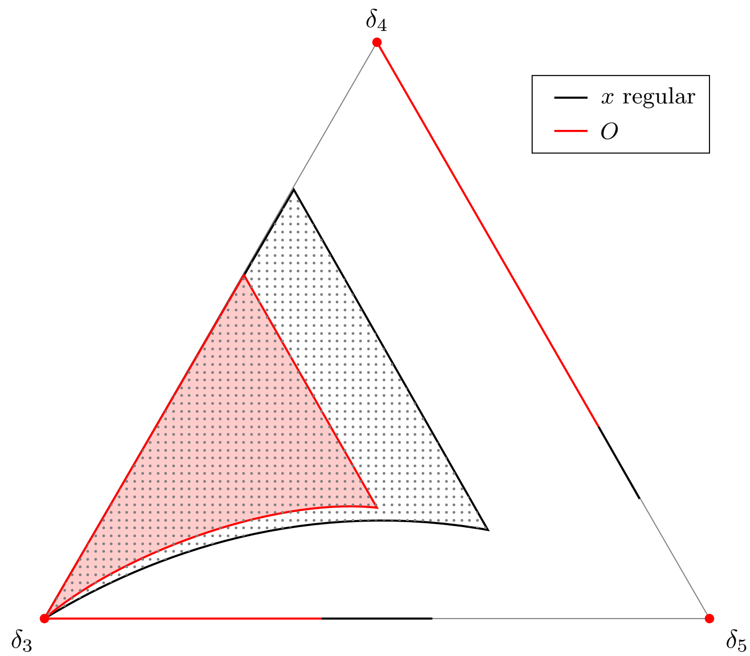

Figure 5: Illustration of O and regular markets with quadratic costs.

5.2 Conditions on Cost Function and Aggregate Market

We now provide sufficient conditions for $\overline{u}'_k(h^*_k)$ to be nondecreasing in terms of the primitives, the cost function and the distribution of values in the aggregate market. Assume that $c(q)$ is thrice differentiable, and that:

$$\inf_q \frac{c'''(q)q}{c''(q)} \geq -1. \tag{38}$$

This condition holds whenever $c'(q)$ is convex or not too concave. This assumption allows us to characterize the shape of $\overline{u}_k$.

Lemma 6 (Concave u_k)
Under (38)*, the concavification of u_k is given by:*

$$\overline{u}_k(h) = \begin{cases} u_k(h) & \text{if } h \leq \overline{h}_k; \\ u_k(\overline{h}_k) & \text{if } h > \overline{h}_k. \end{cases}$$

where $\overline{h}_k$ is defined in (36)*. Furthermore, $\overline{u}_k(h)$ is strictly concave for all $h < \overline{h}_k$.*

In other words, the concavification is defined by a cutoff $\overline{h}_k$, where the concavification is constant for hazard rates above the cutoff and is equal to u_k below the cutoff. An immediate corollary of this is that in the consumer-optimal segmentation, the quality consumed is unique.

Corollary 3 (Unique Consumption)
Under (38)*, in any consumer-optimal segmentation* σ*,* q_k^{σ} *is a singleton for every* k*.*

Hence, we obtain that each buyer consumes only one quality. This refines Corollary 1, which states that with many different buyer values, almost all of them consumed essentially the same quality.

For the aggregate market, we assume that:

$$(v_{k+1} - v_k)\frac{D_k^*}{x_k^*} \text{ is decreasing in } k. \tag{39}$$

If $(v_{k+1} - v_k)$ is constant, i.e. the values are a uniform grid, we recover the well-known monotone hazard rate (MHR) condition. For continuous random variables, this is also equivalent to requiring the distribution to be log-concave (see Bagnoli and Bergstrom (2005) for common distributions that satisfy this condition and other properties of these distributions).

Lemma 7 (No Between-Value Segmentation with Convex Marginal Cost)
If (38)-(39) *are satisfied, then* $\overline{u}_k'(h_k^*)$ *is nondecreasing and hence the consumer surplus generated by the consumer-optimal segmentation is* (37)*.*

Following Lemma 5-7, we thus have that under (38)-(39), the solution to (29) is D^*. Note that the aggregate market in Example 2 did not satisfy the monotone hazard rate condition, which explains why D^* did not solve (29).

We can now provide a sharp characterization of the consumer-optimal segmentation.

Proposition 5 (Gains From Segmentation)
Under (38)-(39)*, the gains from the consumer-optimal segmentation are:*

$$\max_{\sigma}\left[\sum_{x\in \mathrm{supp}(\sigma)} \sigma(x)U(x)\right] - U(x^*) = \sum_{k=1}^{\widehat{k}-1} x_k^*(u_k(\overline{h}_k) - u_k(h_k^*)),$$

where $\widehat{k} = \min\left\{k \mid h_k^* \leq \overline{h}_k\right\}$.

The proposition states that the consumer-optimal segmentation is given by the expected difference between u and $\overline{u}$. Furthermore, this difference will be non-zero for values below a cutoff. Hence, the only gains for consumers come from changes in the consumption of relatively low values.

The proof of Proposition 5 explicitly constructs one consumer-optimal segmentation, which is helpful for understanding how the segmentation is affecting consumption. In this segmentation σ, each market $x \in \text{supp}(\sigma)$ has support $\{v_j, v_{j+1}, \ldots, v_K\}$ for some $j \leq \widehat{k}$. Within every market x, for every $k > \widehat{k}$,

$$h_k^x = h_k^*,$$

meaning high-value buyers consume the same quality as in the aggregate market. On the other hand, for all $k \leq \widehat{k}$,

$$h_k^x = \overline{h}_k \leq h_k^*,$$

so low-value buyers consume a higher quality than they would in the absence of segmentation. We know there is no within-value persuasion, so all buyers of a given type consume the same quality across segments. Whether the consumer-optimal segmentation raises the consumption of type v_k depends on how h_k^* compares to $\overline{h}_k$.

Proposition 5 also allows us to identify conditions so that the cutoff is trivial, so there are no benefits from segmentation at all.

Lemma 8 (No Segmentation—Convex MC)
Under (38)-(39)*, if $h_1^* \leq \overline{h}_1$, then no segmentation is the consumer-optimal segmentation.*

This result is in contrast to Haghpanah and Siegel (2023), where with finite goods, segmentation is generically (in the space of possible markets) beneficial to consumers. Here, with a continuum of qualities, we obtain a large class of aggregate markets in which no segmentation is optimal.

6 Isoelastic Cost

A particular functional form satisfying (38) which gives us interesting comparative statics is the isoelastic cost case:

$$c(q) = \frac{q^\gamma}{\gamma}, \ \gamma > 1. \tag{40}$$

The parameter $\gamma = c'(q)q/c(q)$ indicates the cost elasticity. This parametric form for the cost function will allow us to simplify expressions and provide expression that permit easy comparative statics with respect to γ. We continue to assume (39), the monotone hazard rate condition on the aggregate market.[4]

[4]With isoelastic costs, this assumption can be mildly relaxed to the weaker condition that, for all k,

$$\frac{\phi_{k+1}^* - \phi_k^*}{v_{k+1} - v_k} \geq \frac{1}{\gamma}.$$

We can describe the consumer-optimal segmentation using a discrete version of the familiar notion of demand elasticity η_k^x:

$$\eta_k^x \triangleq \frac{x_k}{D_{k+1}} \cdot \frac{v_k}{v_{k+1} - v_k} = \frac{v_k}{h_k^x}.$$

That is, in market x at price v_x if the price of a good increases by a fraction $(v_{k+1} - v_k)/v_k$ of the original price, then demand will decrease by a fraction x_k/D_{k+1}, which is η_k^x. Note that for measuring the price increase we use as base the pre-increase price while to measure the demand decrease we use the post-increase demand. When working with continuous demands this is obviously irrelevant, while for discrete demands there are many natural ways to extend the definition. We take the definition as it is more convenient for the notation and algebra, but obviously the choice becomes irrelevant as the grid of possible values become fine enough. We now describe the demand elasticity in the segments of a consumer-optimal segmentation.

Proposition 6 (Optimal Segmentation—Isoelastic Cost)
Under (39)-(40)*, in any consumer-optimal segmentation σ, for every $x \in \text{supp}(\sigma)$ and every k such that $x_k > 0$:*

$$\eta_k^x = \begin{cases} \frac{\gamma}{\gamma-1} & \text{if } k < \widehat{k}; \\ \eta_k^* & \text{if } k \geq \widehat{k}. \end{cases}$$

where $\widehat{k} = \min\left\{k \mid h_k^* \leq \frac{\gamma-1}{\gamma}\right\}$

Proof of Proposition 6. The isoelastic cost function satisfies (38). Following Lemma 7, we know that $\overline{u}_k'(h_k^*)$ is nondecreasing and following Lemma 5 we know that this implies that the solution to (29) is D^*. Finally, since $\overline{u}_k'(h_k^*)$ is nondecreasing, we know that $h_k^* \leq \overline{h}_k$ if and only if $k \leq \widehat{k}$, where $\widehat{k}$ is defined in Proposition 5 and $\overline{h}_k$ is defined in (36).

Following Lemma 6 we know that $u_k(h) = \overline{u}_k(h)$ for all $h \leq \overline{h}_k$. Following Proposition 2, we have that in any consumer-optimal segmentation the hazard rates satisfy:

$$h_k^x = \begin{cases} \overline{h}_k & \text{if } k < \widehat{k}; \\ h_k^* & \text{if } k \geq \widehat{k}. \end{cases}$$

Finally, we have that:

$$\overline{h}_k = \frac{v_k(\gamma - 1)}{\gamma}.$$

Re-arranging terms, we get the result. ∎

We thus have that in the consumer-optimal segmentations the demand elasticity in every segment takes two forms. Either the demand has the same elasticity as the aggregate market, or it has a constant elasticity determined by the cost function. Note that we might still have that the support of values differs across markets, so the price of goods might change across markets in the consumer-optimal segmentation. As γ increases, the demand elasticity falls, and the range of consumer values consuming more than they would under no segmentation shrinks.

The demand elasticity in the consumer-optimal segmentation is weakly decreasing in the cost elasticity. The intuition is that with a more elastic cost, the demand needs to be more inelastic for the seller to provide a relative higher quality. Hence, a more inelastic demand is necessary to reduce the inefficiencies.

We can also recover the unit demand case of Bergemann et al. (2015) by taking the limit as $\gamma \to \infty$. Then, we get that in every segment, demand is unit-elastic, and the set of values who consume more than under no segmentation is

$$\left\{ v_k \mid \eta_k^* \leq 1 \right\}$$

which, with the MHR assumption, is the set of values excluded in the aggregate market. Additionally, observe that

$$Q(v_k - \overline{h}_k) = \left(\frac{v_k}{\gamma} \right)^{\frac{1}{\gamma - 1}} \xrightarrow{\gamma \to \infty} 1.$$

Thus, in the consumer-optimal segmentation all consumers are allocated the good, and the segmentation is socially efficient.

Finally, we describe how the potential gains from segmentation changes with γ. For this, we denote by O_γ the set of markets under which no segmentation is optimal when the cost elasticity is γ.

Proposition 7 (Comparative Statics)
For any $\gamma' < \gamma$, $O_{\gamma'} \subset O_\gamma$.

Hence, as the cost becomes more elastic (γ decreases), the potential gains from segmentation increases. This is despite the fact that even in some inefficient markets, no segmentation can be optimal. In contrast, in the limit $\gamma \to \infty$, no segmentation is optimal only if the allocation in the aggregate market is efficient. One could have thought that this means that it is relatively unlikely to find a market in which no segmentation is optimal. However, the conclusion is the opposite: in the limit $\gamma \to \infty$ the cost is very inelastic which reduces the potential benefits from segmentation.

7 Surplus-Sharing Segmentations

So far, this paper has focused on the consumer-optimal segmentation, the segmentation which maximizes consumer surplus. One might naturally wonder whether our approach can be used to characterize the entire surplus-sharing frontier, i.e. the set of all possible divisions of surplus under *some* segmentation. This would produce a generalization of the "surplus triangle" of Bergemann et al. (2015). Of course, the segmentation will also change how the seller prices the different goods, and so it will also change the inefficiencies.

It turns out that by appropriately modifying the objective function, our same approach can be used to produce the entire Pareto frontier, representing all Pareto-efficient outcomes. For the remainder of the surplus frontier, while the (modified) concavification problems remains a valid outer bound, this bound may not be tight. However, under binary values $K = 2$, or if we assume isoelastic costs and MHR, the bound once again becomes tight and our concavification bound characterizes the entire set of achievable outcomes.

7.1 Pareto-Efficient Segmentations

Consider the problem of finding, for some $\lambda \in [0, 1]$, a segmentation which solves

$$\max_{\sigma \in \Delta(\Delta V)} \sum_{x \in \operatorname{supp}(\sigma)} \sigma(x) \Big[\lambda \Pi(x) + (1 - \lambda) U(x) \Big].$$

It will be convenient to rewrite this as

$$\max_{\sigma \in \Delta(\Delta V)} \sum_{x \in \operatorname{supp}(\sigma)} \sigma(x) \Big[\lambda W(x) + (1 - 2\lambda) U(x) \Big],$$

where $W(x)$ is the total welfare in a regular market x. By Lemma 4, it continues to be without loss to restrict attention to regular markets, so we can write W as

$$W(x) \triangleq \sum_{k=1}^{K} w_k(h_k^x) = \sum_{k=1}^{K} v_k Q(v_k - h_k^x) - c(Q(v_k - h_k^x)).$$

We then define the local objective function $\omega_{\lambda,k}$ in a manner similar to the consumer surplus:

$$\omega_{k,\lambda}(h) \triangleq \lambda w_k(h) + (1 - 2\lambda) u_k(h).$$

Then, the problem can be written as

$$\max_{\sigma\in\Delta(\Delta V)} \sum_{x\in\operatorname{supp}(\sigma)} \left[\sum_{k=1}^{K-1} x_k \omega_{\lambda,k}(h_k^x)\right] + \lambda x_K^* w_K(0),$$

where the last term comes from the fact that $h_K^x = 0$ always. Since Section 4 uses no assumptions on the shape of u_k, it is no surprise that the same result goes through by just replacing u_k with $\omega_{\lambda,k}$.

Theorem 2 (Pareto-Efficient Segmentations)
Every Pareto-efficient segmentation σ satisfies, for some $\lambda \in [0,1]$,

$$\sum_{x\in\operatorname{supp}(\sigma)} \sigma(x)\Big[\lambda\Pi(x) + (1-\lambda)U(x)\Big] = \max_{D\prec D^*} \sum_{k=1}^{K-1} x_k^* \overline{\omega}_{\lambda,k}(h_k^D) + \lambda x_K^* w_K(0). \qquad (41)$$

The proof of this theorem completely mirrors that of Theorem 1. Note that for $\lambda \geq \frac{1}{2}$, the solution will be first degree (perfect) price discrimination, which maximizes social surplus and produces zero consumer surplus.

7.2 Surplus-Sharing Frontier

In general, the entire frontier can be found by finding segmentations which solve

$$\max_{\sigma\in\Delta(\Delta V)} \sum_{x\in\operatorname{supp}(\sigma)} \sigma(x)\Big[e_1\lambda\Pi(x) + e_2(1-\lambda)U(x)\Big],$$

for $e_1, e_2 \in \{-1, +1\}$ and $\lambda \in [0,1]$. As before, it will be convenient to rewrite this as

$$\max_{\sigma\in\Delta(\Delta V)} \sum_{x\in\operatorname{supp}(\sigma)} \sigma(x)\Big[e_1\lambda\Pi(x) + \left(e_2(1-\lambda) - e_1\lambda\right)U(x)\Big].$$

We thus define our local objective function

$$\omega_{\lambda,e,k}(h) = e_1\lambda w_k(h) + \Big[e_2(1-\lambda) - e_1\lambda\Big]u_k(h).$$

The first step of the proof of Theorem 1, showing that the concavification bound is an upper bound, still applies, giving us an "outer bound" on the surplus frontier.

Lemma 9 (Outer Bound on Surplus-Sharing Frontier)

For any $e \in \{-1,+1\}^2$ *and* $\lambda \in [0,1]$,

$$\max_{\sigma \in \Delta(\Delta V)} \sum_{x \in \text{supp}(\sigma)} \sigma(x)\Big[e_1 \lambda \Pi(x) + e_2(1-\lambda)U(x)\Big] \le \max_{D \prec D^*} \sum_{k=1}^{K-1} \overline{\omega}_{\lambda,e,k}(h_k^D) + \omega_{\lambda,e,K}(0). \quad (42)$$

Theorem 2 tells us that for $e = (+1,+1)$, the boundary of the surplus frontier and the outcome corresponding to the solution of (42) coincide. However, in general, this is not true. The reason is that in the proofs of Theorem 1 and 2, we show the upper bound is implementable with segmentations over regular markets. A crucial step of the proof is noticing that the first-order condition governing the optimal D implies a submodularity property (Lemma 5) which guarantees regularity. For general (e, λ), we lose this submodularity, and hence the regularity guarantee.

The concavification bound of (42) is tight whenever $K = 2$, since all binary distributions are regular. We can also recover the tightness of the bound by imposing the isoelastic functional form, along with monotone hazard rate. In particular, suppose that the cost function is:

$$c(q) = \frac{q^\gamma}{\gamma}, \ \gamma \ge 2, \quad (43)$$

and again impose the MHR condition (39). With these two assumptions, the entire surplus boundary can be recovered from the concavification bound (42). In fact, the solution is attained by $D = D^*$, as it was in Proposition 6.

Proposition 8 (Surplus-Sharing Frontier)
Under (39) *and* (43)*, the surplus-sharing frontier attains the concavification bound of* (42) *with equality at* $D = D^*$*:*

$$\max_{\sigma \in \Delta(\Delta V)} \sum_{x \in \text{supp}(\sigma)} \sigma(x)\Big[e_1 \lambda \Pi(x) + e_2(1-\lambda)U(x)\Big] = \sum_{k=1}^{K-1} \overline{\omega}_{\lambda,e,k}(h_k^*) + \omega_{\lambda,e,K}(0)$$

for all $e \in \{-1,+1\}^2$ *and* $\lambda \in [0,1]$.

The functional form assumption (43) is only used to prove tightness of the concavification bound when $e = (-1,+1)$, corresponding to the "lower right" section of the boundary, which is the most difficult case to deal with. The weaker requirement that $c'''(q) \ge 0$ is sufficient for the other cases.

Figure 6 shows the surplus-sharing frontier for quadratic costs, $c(q) = q^2/2$, when $V = \{1,2\}$ and $x^* = (\frac{3}{5}, \frac{2}{5})$. Note that the set of possible surplus divisions is a convex subset of the "surplus triangle" of Bergemann et al. (2015), which is the triangle formed by the constraints that (1) $U \geq 0$, (2) $\Pi \geq \Pi^*$, the profit in the aggregate market, and (3) $U + \Pi$ does not exceed the total available surplus in the market.

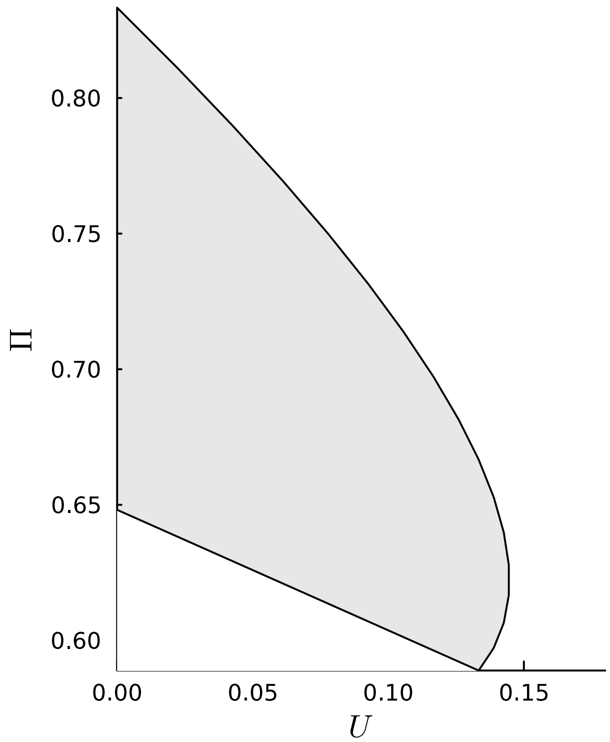

Figure 6: Surplus-sharing frontier with quadratic cost.

8 Concavification and Extreme Points

Our results arise as a result of concavification arguments; this stands in contrast to the existing literature on price discrimination, which has taken different routes to understanding how market segmentation can benefit consumers. In this section, we discuss some of the main results in the literature and contrast them to our results.

8.1 Extreme Points

Given a menu p, we can consider the (possibly empty) set of markets where p is optimal:

$$X_p \triangleq \{x \in \Delta V \mid \Pi(x,p) \geq \Pi(x,p') \; \forall p'\} \,. \tag{44}$$

The set X_p is compact and convex. An *extreme point* of X_p is a point which cannot be represented by convex combination of other points in X_p. We will also refer to markets in the topological interior of X_p, meaning they are neither extreme points nor on the boundary of X_p. Consequently, we call such an x an *interior market*. By the Minkowski-Caratheodory Theorem (Simon (2011), Theorem 8.11), X_p is equal to the convex hull of its extreme points.

We denote by p^* the optimal price in the aggregate market, which we assume in this section that is uniquely defined. We also assume that the allocation in the aggregate market is inefficient to avoid analyzing trivial cases. The set X_{p^*} identifies the set of all markets $x' \in X_{p^*}$ where the seller's optimal price remains the same as in the aggregate market. In particular, the extreme points of this set can be useful to identify the consumer-optimal segmentation in some situations.

To illustrate this, analyze the case in which the seller has an indivisible good for sale. We obtain this by assuming that:

$$c(q) = \begin{cases} 0 & \text{if } q \leq 1; \\ \infty & \text{if } q > 0. \end{cases} \tag{45}$$

This is the case studied by Bergemann et al. (2015). In this case we have that the extreme points are sufficient to understand the consumer-optimal segmentation (and, in fact, they are sufficient to understand the welfare implications of third-degree price discrimination).

Proposition 9 (Optimal Segmentation—Single Good)
If the cost function is (45)*, then there exists a consumer-optimal segmentation that places weight only on the extreme points of* X_{p^*}.

Bergemann et al. (2015) also show that using the extremal points of X_{p^*} allow implementing an efficient allocation without increasing the seller's profits (relative to the aggregate market), and hence it must be the consumer-optimal segmentation. A remarkable aspect of this result is that the extreme points also have a very simple structure.

With more than one good one can continue to follow this approach. In particular consider a cost function $c(q)$ which is piecewise linear with kinks at integer values of q, i.e. when $q \in \{1, 2, \ldots, \}$. We can interpret this as the cost function of a seller that has many (but discrete) goods for sale, where a non-integer q is a randomization between the neighboring two integer values.[5] We can then try to apply the same logic as in Bergemann et al. (2015), but we obtain a much weaker result.

Proposition 10 (Pareto Improvements)
If x^* *is in the interior of* X_{p^*} *and the allocation is inefficient in the aggregate market, then there is a segmentation that places weight only on the extreme points of* X_{p^*} *that increases consumer surplus and keeps profits constant.*

[5]That is, the buyer gets quality $\lceil q \rceil$ with probability $(q - \lfloor q \rfloor)$ and $\lfloor q \rfloor$ with probability $(\lceil q \rceil - q)$, where $\lfloor \cdot \rfloor$ is the floor function and $\lceil \cdot \rceil$ is the ceiling function.

An immediate corollary is that for generic aggregate markets, there is a segmentation that Pareto improves the welfare generated by the aggregate market. This was also proved in much greater generality by Haghpanah and Siegel (2023).

There are two drawbacks with this approach. First, segmentations that place weight only on the extreme points of X_{p^*} are not sufficient to generate the consumer-optimal segmentation. Hence, it is necessary to study segmentations that place weight on all extreme points of all sets of the form X_p, and not just X_{p^*}. Second, the number of extreme points grows quickly with the number of goods. To illustrate this, we exemplify what happens as the number of goods grows to infinite.

Proposition 11 (Extreme Markets with a Continuum of Goods)
If c' is strictly increasing for all q, then for every strictly regular market x that induces no exclusion (that is, ϕ_k^x is strictly increasing in k and positive for all k), X_{p^x} is a singleton. In particular, x is a extreme market of X_{p^x}.

We thus have that as the number of markets grow, essentially every relevant market is an extreme market (recall from Lemma 4 we can focus on regular markets). Since we do not know whether the consumer-optimal segmentation will place weights only on the extreme points of X_{p^*}, this approach does not allow narrowing the class of markets that need to be considered when there are many goods. Furthermore, it is clear that in general the gains from segmentations will become negligible as the number of goods increase, unless we allow for non-local segmentations. Thus, while Proposition 10 remains valid for any finite number of goods, its usefulness becomes negligible as the number of goods grow.

Appendix B contains additional discussion and results for the discrete good case. We also show how concavification can be used to derive the extreme points in closed form.

8.2 Local Segmentations

We can use the set O to characterize the set of markets that are in the support of a consumer-optimal segmentation. For the second part we first observe that in any consumer-optimal segmentation σ, if $x \in \text{supp}(\sigma)$ then no segmentation must be the consumer-optimal segmentation when the aggregate market is x. After all, if there is an improvement for consumers, these gains could have been also realized when the aggregate market is x^*.

Corollary 4 (Support of Consumer-Optimal Segmentations)
Any consumer-optimal segmentation σ has support $\text{supp}(\sigma) \subset O$.

While the set O restricts the set of possible markets, this can still be a larger set. However, one can build a consumer-optimal segmentation in which one restricts attention to the closest markets to x^*. To formalize this, for any subset of markets $M \subset \Delta V$, we denote by $\text{co}(M)$ the convex hull of M. And, denote by $\overline{\text{co}}(M)$ the convex hull of M without the points M themselves:

$$\overline{\text{co}}(M) \triangleq \text{co}(M) \backslash M.$$

We say M are local extreme points of x^* if $M \subset O$, $x \in \text{co}(M)$ and $\overline{co}(M) \cap O = \{\emptyset\}$. In other words, the markets M are local extreme points of x^* if markets M belong to O, they allow segmenting x^* and there is no market in O that can nontrivially be segmented by markets in M.

Proposition 12 (Constructing Consumer-Optimal Segmentations)
For all $x^* \notin O$, there exists a consumer-optimal segmentation that has support only on local extreme markets of x^*.

Proof. Let $\preceq$ be the Blackwell order on $\Delta(\Delta V)$, that is, for any $\sigma, \sigma' \in \Delta(\Delta V)$, $\sigma \preceq \sigma'$ if σ' is a mean-preserving spread of σ. Let S^* be the set of all consumer-optimal segmentations. Let $\sigma \in S^*$ be such that $\sigma \preceq \sigma'$ for all $\sigma' \in S^*$. Following Proposition 4, we know that $\text{supp}(\sigma) \subset O$. Now, suppose $\text{supp}(\sigma)$ is not a local extreme market of x^*. Then, there exists $\widehat{x}$ such that $\widehat{x} \in \overline{\text{co}}(\text{supp}(\sigma))$. Let $\widehat{\sigma}$ be a non-trivial segmentation of $\widehat{x}$ with support $\text{supp}(\sigma') \subset \text{supp}(\sigma)$. We now consider the following segmentation:

$$\widetilde{\sigma}(x) = \begin{cases} \sigma(x) - \epsilon\widehat{\sigma}(x) & \text{if } x \neq \widehat{x}; \\ \sigma(\widehat{x}) + \epsilon & \text{if } x = \widehat{x}, \end{cases}$$

where ϵ is small enough such that $\widetilde{\sigma}(x) \geq 0$ for all x. The utility of segmentation $\widetilde{\sigma}$ can be written as follows:

$$\sum_{x \in \text{supp}(\sigma)} \widetilde{\sigma}(x) U(x) = \sum_{x \in \text{supp}(\sigma)} \sigma(x) U(x) + \epsilon \left(U(\widehat{x}) - \sum_{x \in \text{supp}(\sigma)} \widehat{\sigma}(x) U(x) \right).$$

Following Proposition 4 we have that:

$$U(\widehat{x}) - \sum_{x \in \text{supp}(\sigma)} \widehat{\sigma}(x) U(x) \geq 0.$$

Hence, $\widetilde{\sigma}$ is also a consumer-optimal segmentation. However, we also have that $\widetilde{\sigma} \preceq \sigma$, so we reach a contradiction. ∎

The proposition identifies a set of candidates markets that can be part of a consumer-optimal segmentation and then states that one can reduce the search to the closest markets among the candidate ones (that is, closest to the aggregate market). As a way to illustrate how much Proposition 12 narrows the search of a consumer-optimal segmentation, we note that it is possible to restrict attention to consumer-optimal segmentations that place positive weight on the boundary of O. The boundary of O is highlighted in Figure 5. In this sense, it shares a similar spirit as the work of Bergemann et al. (2015) (Proposition 9): we identify potential markets that can be in the support of an optimal segmentation and then show we just need to search among the closest ones.

9 Conclusion

This paper has characterized how market segmentation affects consumer welfare when monopolists can engage in both second- and third-degree price discrimination. Our analysis yields several key insights. First, consumer-optimal segmentation maintains consistent quality provision across segments while allowing price variation. Second, the benefits of segmentation depend critically on demand elasticities and cost structures, with no segmentation being optimal when aggregate demand is sufficiently elastic.

These theoretical results have direct practical implications. For competition authorities, they suggest that market segmentation should be evaluated based on observable market characteristics rather than prohibited categorically. The finding that quality provisions remain consistent across segments provides a potential metric for identifying harmful segmentation practices. For firms, our characterization of optimal segmentation strategies offers guidance for designing market segmentation policies that balance profit maximization with consumer welfare.

Our analysis also connects to broader debates about big data and personalized pricing in digital markets. While enhanced ability to segment markets could enable more sophisticated price discrimination, our results suggest this may benefit consumers when properly structured. However, the conditions we identify for beneficial segmentation - particularly regarding demand elasticities and cost structures - may help guide regulatory policy.

Several important directions remain for future research. First, extending the analysis to competitive markets as in the recent analysis of Bergemann et al. (2023) for single unit demand would provide insight into how market structure affects optimal segmentation. Second, empirical work testing our theoretical predictions about the relationship between demand elasticities and optimal segmentation would be valuable.

References

Aguirre, I., S. Cowan, and J. Vickers (2010): "Monopoly Price Discrimination and Demand Curvature," *American Economic Review*, 100, 1601–1615.

Bagnoli, M. and T. Bergstrom (2005): "Log-concave Probability and its Applications," *Economic Theory*, 26 (2), 445–469.

Bergemann, Dirk, Ben Brooks, and Stephen Morris (2015): "The Limits of Price Discrimination," *American Economic Review*, 105, 921–957.

Bergemann, D., B. Brooks, and S. Morris (2023): "On the Alignment of Consumer Surplus and Total Surplus under Competitive Price Discrimination," Tech. Rep. CFDP 2373, Cowles Foundation for Research in Economics.

Condorelli, D. and B. Szentes (2020): "Information Design in the Hold-Up Problem," *Journal of Political Economy*, 128, 681–709.

Cowan, S. (2012): "Third-Degree Price Discrimination and Consumer Surplus," *Journal of Industrial Economics*, 60, 333–345.

Fan, K. and G. Lorentz (1954): "An Integral Inequality," *American Mathematical Monthly*.

Haghpanah, Nima and Jason Hartline (2021): "When Is Pure Bundling Optimal?" *Review of Economic Studies*, 88, 1127–1156.

Haghpanah, N. and R. Siegel (2022): "The Limits of Multi-Product Discrimination," *American Economic Review: Insights*, 4, 443–458.

——— (2023): "Pareto Improving Segmentation of Multi-Product Markets," *Journal of Political Economy*, 131 (6), 1546–1575.

Johnson, J. and D. Myatt (2003): "Multiproduct Quality Competition: Fighting Brands and Product Line Pruning," *American Economic Review*, 93, 748–774.

Kamenica, Emir and Matt Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack (2021): "Extreme Points and Majorization: Economic Applications," *Econometrica*, 89, 1557–1593.

Maskin, E. and J. Riley (1984): "Monopoly with Incomplete Information," *RAND Journal of Economics*, 15, 171–196.

Mussa, M. and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

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

Robinson, J. (1933): *The Economics of Imperfect Competition*, London: Macmillan.

Roesler, A. and B. Szentes (2017): "Buyer-Optimal Learning and Monopoly Pricing," *American Economic Review*, 107, 2072–2080.

Schmalensee, R. (1981): "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination," *American Economic Review*, 71, 242–247.

Simon, Barry (2011): *Convexity: An Analytic Viewpoint*, Cambridge University Press.

Varian, H. (1985): "Price Discrimination and Social Welfare," *American Economic Review*, 75, 870–875.

A Proof Details

In these proofs, we will at times refer to an indexed list of markets x^ℓ. In these cases, a superscript ℓ is shorthand for market x^ℓ, e.g. h^1 means h^{x^1}.

Proof of Lemma 4. We fix a market x and let p be a seller-optimal price menu. Let X_p be the set of markets where p is optimal:

$$X_p \triangleq \left\{ x' \in \Delta V : p \in \arg\max \sum_k x'_k \Pi(v_k, p) \right\}.$$

It is clear that X is convex and compact. Following the Krein-Milman theorem, we can write x as a linear combination of extreme points of X. Furthermore, by Caratheodory's theorem, we can write it as a linear combination of at most K extreme points.

It thus suffices to show that every extreme point of X_p is regular. Suppose that y is an extreme point of X_p and is not regular. Then there exists v_k such that

$$\varphi^D_{k+1} < \varphi^D_k.$$

Consider the binary segmentation supported on the following markets:

$$y_i^+ = \begin{cases} y_i & \text{if } i \notin \{k, k+1\}; \\ y_i + \epsilon & \text{if } i = k; \\ y_i - \epsilon & \text{if } i = k+1, \end{cases} \qquad y_i^- = \begin{cases} y_i & \text{if } i \notin \{k, k+1\}' \\ y_i - \epsilon & \text{if } i = k; \\ y_i + \epsilon & \text{if } i = k+1, \end{cases}$$

where ϵ is small enough such that $y_+(v_{k+1}), y_-(v_k) \geq 0$. It is easy to verify that $y_+, y_- \in X$ and that

$$y = \frac{y_+ + y_-}{2}.$$

Thus, y is not an extreme point of X_p, a contradiction. ∎

Proof of Theorem 1. In the main text, we showed that (29) is an upper bound. Next, we show that there exists a segmentation which exactly achieves the average hazard rate at every point, i.e., for every $x \in \text{supp}(\sigma)$ and $v_k \in \text{supp}(x)$, $h^x_k = h^D_k$.

Lemma A.1 (Majorized D are Feasible)
For any D such that $D \prec D^*$, there exists a segmentation σ such that for every $x \in \text{supp}(\sigma)$ and $v_k \in \text{supp}(x)$, $h^x_k = h^D_k$.

Proof. For ease of notation, in this proof we work with unscaled markets $z : V \to \mathbb{R}_+$ which we treat like distributions without imposing $\sum_k z_k = 1$. We only require that for all k,

$$\sum_z \sigma(z) z_k = x_k^*. \tag{A.1}$$

At the end, we can convert these unscaled markets into actual markets by re-scaling:

$$x_k = \frac{z_k}{\sum_i z_i}, \quad \hat{\sigma}(x) = \sigma(z) \cdot \sum_i z_i.$$

This rescaling does not affect the hazard rates (and hence the implied allocations) in any way. Additionally,

$$\sum_k \left[\sum_z \sigma(z) z_k \right] = 1 \implies \sum_k \left[\sum_x \hat{\sigma}(x) x_k \right] = 1 \implies \sum_x \hat{\sigma}(x) = 1$$

so $\hat{\sigma}$ is in fact a valid segmentation.

Let D be the function we want to implement. This proof works by induction, going decreasing down the support. Start with some segmentation σ over $\{z^1, \ldots, z^L\}$ with the property that for some value j, for all $k \leq j$,

$$z_k^\ell = x_k^*. \tag{A.2}$$

Furthermore, for all $k > j$ and z_ℓ such that $v_k \in \operatorname{supp}(z_\ell)$,

$$h_k^\ell = \frac{\Delta_k^\ell D_{k+1}^\ell}{z_k^\ell} = h_k^D = \frac{(v_{k+1} - v_k) D_{k+1}}{x_k^*}. \tag{A.3}$$

(A.1)-(A.3) are clearly satisfied by the trivial segmentation for $j = K - 1$. We produce a segmentation $\hat{\sigma}$ supported on at most one extra market which preserves (A.1)-(A.3) at $j - 1$.

The construction is as follows. Suppose that we keep the same markets z^ℓ and weights $\sigma(z^\ell)$, but we modify them only at v_j so that either (A.3) holds, or $z_j^\ell = 0$. That is,

$$\hat{z}_j^\ell = \begin{cases} \frac{h_j^\ell}{h_j^D} \cdot x_j^* & v_j \in \operatorname{supp}(\hat{z}_\ell) \\ 0 & \text{otherwise} \end{cases}.$$

Since $\hat{z}_k^\ell = z_k^\ell$ for all $k \neq j$, (A.1)-(A.2) continue to hold except at $k = j$. Clearly, if we remove v_j from every market, then $\sum_\ell \sigma(z^\ell) \hat{z}_j^\ell = 0 < x_j^*$. We claim that if v_j is included in the support of every market, then

$$\sum_{\ell} \sigma(z^\ell) \hat{z}_j^\ell \geq x_j^*.$$

To prove this, we first note that:

$$\sum_{\ell} \sigma(z^\ell) \hat{z}_j^\ell = \sum_{\ell} \frac{\sigma(z^\ell) h_j^\ell}{h_j^D} \cdot x_j^* = \frac{h_j^\sigma}{h_j^D} \cdot x_j^* = \frac{D_{j+1}^\sigma}{D_{j+1}} \cdot x_j^*.$$

We now show that

$$\frac{D_{j+1}^\sigma}{D_{j+1}} \geq 1.$$

Observe that v_j is in the support of all z^ℓ. Thus, by (32) and (A.1),

$$\sum_{i=j}^{K-1} (v_{i+1} - v_i) D_{j+1}^\sigma = \sum_{i=j}^{K-1} (v_{i+1} - v_i) D_{j+1}^*. \tag{A.4}$$

The majorization constraint implies that:

$$\sum_{i=j}^{K-1} (v_{i+1} - v_i) D_{j+1}^* \geq \sum_{i=j}^{K-1} (v_{i+1} - v_i) D_{i+1}. \tag{A.5}$$

Combining (A.4) and (A.5), we have

$$(v_{j+1} - v_j) D_{j+1}^\sigma + \sum_{i=j+1}^{K-1} (v_{i+1} - v_i) D_{i+1}^\sigma \geq \sum_{i=j}^{K-1} (v_{i+1} - v_i) D_{i+1}. \tag{A.6}$$

Furthermore, for every $k > j$, by (A.3)

$$D_{k+1}^\sigma = \frac{x_k^* h_k^\sigma}{v_{k+1} - v_k} = \frac{x_k^* h_k^D}{v_{k+1} - v_k} = D_{k+1}$$

and hence (A.6) implies that

$$D_{j+1}^\sigma \geq D_{j+1}.$$

Clearly, the value of $\sum_\ell \sigma(z^\ell) z_j^\ell$ is strictly decreasing as we remove v_j from the support of markets one by one. The analysis above shows that as some point, we cross over x_j^*. At the crossing point we split the market into two segments, with appropriate weights so that (A.1) is satisfied with equality. This completes the construction of σ'.

There are two important things to point out from this proof. First, the number of total segments is at most K. Second, the proof is agnostic about the order in which we remove v_j from the support of the markets in each step. In particular, we can order them in a way that produces markets with more natural supports. For example, if we always order the markets by whether they include v_{j+1}, we get the feature that at every v_j, either every market with a gap at v_{j+1} also has a gap at v_j, or every market with v_{j+1} in the support also has v_j in the support. Together, these imply that the construction has a well-defined limit when V approaches a continuum of values. ∎

In the end, we are left with a segmentation such that in every market,

$$U(x) = \sum_{k=1}^{K-1} x_k^* u_k(h_k^\sigma).$$

However, we want to achieve the concavified value $\overline{u}_k$, not u_k. Take any segment x and v_k such that

$$\overline{u}_k(h_k^\sigma) = \lambda u_k(h_1) + (1-\lambda) u_k(h_2), \quad \lambda h_1 + (1-\lambda) h_2 = h_k^\sigma,$$

where $h_1 < h_k^\sigma < h_2$.

Let us segment x into two markets x^1, x^2 with the same support as x such that: (1) for all $v_j \neq v_k$, $h_j^1 = h_j^2 = h_j^x$, and (2) $h_j^1 = h_1^x$, $h_j^2 = h_2^x$. Since demands are equal up to v_k, $x_k^1 > x > x_k^2$. Hence, there exists some $\mu \in [0,1]$ such that $\mu x_k^1 + (1-\mu) x_k^2 = x$. The payoff across this segmentation is

$$\mu x_k^1 u_k(h_1) + (1-\mu) x_k^2 u_k(h_2) = \overline{u}_k(h_k^\sigma)$$

as desired.

The last step of the proof is to check that this constructed segmentation respects the regularity assumption necessary for (25) to hold. To do so, we rely on the following lemma.

Lemma A.2 (Submodularity Implies Regularity)
If $u_k'(h_k) = \overline{u}_k'(h_k) \leq \overline{u}_{k+1}'(h_{k+1}) = u_{k+1}'(h_{k+1})$, *then* $v_{k+1} - h_{k+1} \geq v_k - h_k$.

Proof. Decompose the difference

$$0 \leq \overline{u}_{k+1}'(h_{k+1}) - \overline{u}_k'(h_k) = \left[\overline{u}_{k+1}'(h_{k+1}) - \overline{u}_{k+1}'(h_k + (v_{k+1} - v_k))\right] + \left[\overline{u}_{k+1}'(h_k + (v_{k+1} - v_k)) - \overline{u}_k'(h_k)\right].$$

We claim that the second bracketed difference is negative. If so, then the first difference must be positive, which by concavity is true and only if

$$h_{k+1} \leq (h_k + (v_{k+1} - v_k)) \iff v_{k+1} - h_{k+1} \geq v_k - h_k.$$

To prove this, consider the analytical extension of u_k onto all v:

$$u_v(h) \triangleq hQ(v-h) \implies u_v'(h) = Q(v-h) - hQ'(v-h).$$

Let $h(v) = h_k + (v - v_k)$, so that

$$\overline{u}_{k+1}'(h_k + (v_{k+1} - v_k)) - \overline{u}_k'(h_k) = \int_{v_k}^{v_{k+1}} \frac{d}{dv}\left[\overline{u}_v'(h(v))\right] dv. \tag{A.7}$$

There are two distinct cases: when $u_v(h) = \overline{u}_v(h)$, and where $u_v(h) < \overline{u}_v(h)$. First, for any open region where $u_v(h(v)) = \overline{u}_v(h(v))$,

$$\begin{aligned} \frac{d}{dv}\left[\overline{u}_v'(h(v))\right] = \frac{d}{dv}\left[u_v'(h(v))\right] &= (1 - 2h'(v))Q'(v-h) - h(v)(1-h'(v))Q''(v-h) \\ &= -Q'(v-h(v)) - (1-h'(v))u_v''(h) = -Q'(v-h(v)) \leq 0, \end{aligned}$$

where the second line uses that

$$u_v''(h) = hQ''(v-h) - 2Q'(v-h).$$

Next, we claim that for any differentiable $h(v)$,

$$u_v(h(v)) < \overline{u}_v(h(v)) \implies \frac{d}{dv}\left[\overline{u}_v'(h(v))\right] \leq 0.$$

To see this, let $T(v)$ and $B(v)$ be the top and bottom support points of $\overline{u}_v(h(v))$, respectively, so that

$$\overline{u}_v'(h(v)) = u_v'(T(v)) = u_v'(B(v)) = \frac{u_v(T(v)) - u_v(B(v))}{T(v) - B(v)}. \tag{A.8}$$

We can compute

$$\frac{d}{dv}\left[\overline{u}_v'(h(v))\right] = \frac{u_v'(T(v)) \cdot T'(v) + \frac{\partial}{\partial v}u_v(T(v)) - u_v'(B(v)) \cdot B'(v) - \frac{\partial}{\partial v}u_v(B(v))}{T(v) - B(v)} + \frac{(T'(v) - B'(v)) \cdot (u_v(T(v)) - u_v(B(v)))}{(T(v) - B(v))^2}.$$

Substituting in (A.8), the second term cancels with part of the first, resulting in

$$\frac{d}{dv}\left[\overline{u}_v'(h(v))\right] = \frac{1}{T(v)-B(v)}\left(\frac{\partial}{\partial v}u_v(T(v)) - \frac{\partial}{\partial v}u_v(B(v))\right).$$

We can simplify this by using

$$\frac{\partial}{\partial v}u_v(h) = hQ'(v-h).$$

But, note that

$$u_v'(T(v)) = Q(v-T(v)) - T(v)Q'(v-T(v)) = Q(v-B(v)) - B(v)Q'(v-B(v)) = u_v'(B(v)).$$

This, in turn, means that

$$\frac{d}{dv}u_v(T(v)) - \frac{\partial}{\partial v}u_v(B(v)) = \frac{1}{T(v)-B(v)}\Big(Q(v-T(v)) - Q(v-B(v))\Big) \leq 0.$$

Hence, the integrand of (A.7) is weakly negative on any open interval, completing the proof. ∎

To see how we can use this lemma, note that the KKT conditions for (29) are:

$$\overline{u}_k'(h_k^\sigma) = \sum_{i=1}^{k-1}\mu_i \tag{A.9}$$

where μ_i is the multiplier associated with the majorization constraint starting at i. All $\mu_i \geq 0$, so at the optimal D^σ, $\overline{u}_k'(h_k^\sigma) \leq \overline{u}_{k+1}'(h_{k+1}^\sigma)$. But, by construction, for all $x \in \operatorname{supp}(\sigma)$, $h_k^x \in \operatorname{supp}(\overline{u}_k(h_k^\sigma))$. In particular, this means that u_k is concave at h_k^x and $u_k'(h_k^x) = \overline{u}_k'(h_k^\sigma)$. Thus, we can apply Lemma A.2, implying that the segmentation we constructed in Lemma A.1 is regular, completing the proof of Theorem 1. ∎

Proof of Proposition 3. Recall that for any $x \in \operatorname{supp}(\sigma)$ and $v_k \in \operatorname{supp}(x)$,

$$\overline{u}_k'(h_k^x) = \overline{u}_k'(h_k^\sigma)$$

which is constant across x. By Lemma A.2, this means that for all k and $x, x' \in \operatorname{supp}(\sigma)$,

$$v_{k+1} - h_{k+1}^x \geq v_k - h_k^{x'}.$$

This means

$$q_{k+1}^x = Q(v_{k+1} - h_{k+1}^x) \geq Q(v_k - h_k^{x'}) = q_k^{x'}$$

for all x, x' and instances of v_k, v_{k+1}. ∎

Proof of Lemma 6. Recall that

$$u_k'(h) = Q(v_k - h) - hQ'(v_k - h), \quad u_k''(h) = hQ''(v_k - h) - 2Q'(v_k - h).$$

u_k being increasing at h means that

$$u_k'(h) \geq 0 \implies h \leq \frac{Q(v_k - h)}{Q'(v_k - h)}.$$

We would like this to be a sufficient condition for u_k to be concave at h, that is, either $Q''(v_k - h) \leq 0$, or

$$u_k''(h) \leq 0 \implies h \leq \frac{2Q'(v_k - h)}{Q''(v_k - h)}.$$

A sufficient condition is that

$$\frac{Q(v_k - h)}{Q'(v_k - h)} \leq \frac{2Q'(v_k - h)}{Q''(v_k - h)} \iff \frac{Q''(v_k - h)Q(v_k - h)}{Q'(v_k - h)^2} \leq 2.$$

Defining $q = Q(v_k - h)$, then recalling that $Q(\varphi) = (c')^{-1}(\varphi)$ and applying the inverse function theorem, yields

$$\frac{c'''(q)q}{c''(q)} \geq -2$$

which is obviously implied by (38). ∎

Proof of Lemma 7. This proof consists of two parts. First, we show that $\overline{h}_k$ is increasing while h_k^* is decreasing, meaning that for some threshold $\widehat{k}$, $k < \widehat{k} \implies \overline{u}_k'(h_k^*) = 0$, and for $k \geq \widehat{k}$, $\overline{u}_k'(h_k^*) = u_k'(h_k^*)$. $\overline{h}_k$ is the discretization of

$$h(v) = \arg\max_h \left[hQ(v - h)\right].$$

The FOC characterizing $h(v)$ is

$$q(v - h(v)) - h(v)Q'(v - h(v)) = 0$$

Applying the implicit function theorem, we get

$$Q'(v - h(v))(1 - h'(v)) - h'(v)Q'(v - h(v)) - h(v)Q''(v - h(v))(1 - h'(v)) = 0$$
$$\implies h'(v) = 1 - \frac{Q'(v - h(v))}{2Q'(v - h(v)) - h(v)Q''(v - h(v))}.$$

Under (38),

$$h(v) \leq \frac{Q'(v-h(v))}{|Q''(v-h(v))|} \implies 2 - \frac{h(v)Q''(v-h(v))}{Q'(v-h(v))} \geq 1 \implies h'(v) \in [0,1].$$

Hence, $\overline{h}_k$ is increasing, and so is $v_k - \overline{h}_k$ (a fact we need for the proof of Proposition 5).

Next, we show that $u_k'(h_k^*)$ is increasing for all $k \geq \widehat{k}$. We can decompose

$$u'_{k+1}(h^*_{k+1}) - u'_k(h^*_k) = \big(u'_{k+1}(h^*_k) - u'_k(h^*_k)\big) + \big(u'_{k+1}(h^*_{k+1}) - u'_{k+1}(h^*_k)\big).$$

The first term is equal to

$$\int_{v_k}^{v_{k+1}} \frac{\partial}{\partial v}\big[u'_v(h^*_k)\big]\ dv = \int_{v_k}^{v_{k+1}} Q'(v-h^*_k) - h^*_k Q''(v-h^*_k)\ dv.$$

Under (38), for any h such that $h \leq \overline{h}_k$,

$$hQ''(v-h) \leq Q'(v-h) \implies \int_{v_k}^{v_{k+1}} \frac{\partial}{\partial v}\big[u'_v(h^*_k)\big]\ dv \geq 0.$$

This also means, in particular, that $u'_{k+1}(h) \geq 0$ for all $h \in (h^x_{k+1}, h^x_k)$. The second term is

$$\int_{h^x_k}^{h^*_{k+1}} u''_{k+1}(h)\ dh \geq 0$$

using the fact that $h^*_k \leq h^*_{k+1}$ by (39) and that, by Lemma 6, $k \geq \widehat{k} \implies u''_{k+1}(h) \geq 0$. Thus, $u'_k(h^*_k)$ is increasing in k. ∎

Proof of Proposition 5. What remains is to show that there exists a solution in which every segment x is supported on $\{v_j, \ldots, v_K\}$ for some $j \leq \hat{k}$. Consider the following demand $\hat{D}$ supported on V, with hazard rates:

$$\hat{h}_k = \min\{\overline{h}_k, h^*_k\}.$$

Since $\overline{h}_k$ is increasing and h^*_k is decreasing, $\hat{h}$ switches from the first argument to the second exactly once, and $\overline{h}_k - h^*_k = (v_k - h^*_k) - (v_k - \overline{h}_k)$ satisfies increasing differences. Furthermore, from the proof of Lemma 7, we know that $v_k - \overline{h}_k$ is increasing, so $\hat{D}$ is regular.

We are now ready to construct the segmentation. Define the family of distributions

$$\hat{D}^j(v) = \begin{cases} 1 & \text{if } v < v_j; \\ \frac{\hat{D}(v)}{\hat{D}} & \text{if } v \geq v_j. \end{cases}$$

Since $\hat{D}$ is regular, so is each $\hat{D}^j$. What we now need is to find weights $\hat{\sigma}^j$ such that $\sum_j \hat{\sigma}^j \hat{D}^j = D^*$ and $\sum_j \hat{\sigma}^j = 1$. We construct these weights recursively.

Begin by making $\hat{\sigma}^1$ the largest value until $\hat{\sigma}^1 \hat{D}^1(v) \leq D^*(v)$ is binding for some v. Our previous argument shows that $\varphi_k^* - (v_k - h_k)$ is decreasing. This is the amount of "over-weighting" in $\hat{D}$ relative to D^*, meaning that $\frac{\hat{x}/\hat{D}}{x_k^*/D_k^*}$ is decreasing, so the binding v for $\hat{D}^1$ is, in fact, v_1.

Now, consider the remainder distribution $\widetilde{D} = D^* - \hat{\sigma}^1 \hat{D}^1$. Notice that

$$\frac{\hat{x}_k/\hat{D}_k}{\widetilde{x}_k/\widetilde{D}_k} = \frac{\hat{x}_k/\hat{D}_k}{(x_k^* - \hat{\sigma}^1 \hat{x}_k^1)/(D_k^* - \hat{\sigma}^1 \hat{D}_k^1)}.$$

Compute

$$\frac{x_k^* - \hat{\sigma}^1 \hat{x}_k^1}{D_k^* - \hat{\sigma}^1 \hat{D}_k^1} = \left[\frac{1 - \hat{\sigma}^1 \hat{x}_k^1 / x_k^*}{1 - \hat{\sigma}^1 \hat{D}_k^1 / D_k^*}\right] \frac{x_k^*}{D_k^*}.$$

Combining these two together, we get that

$$\frac{\hat{x}_k/\hat{D}_k}{\widetilde{x}_k/\widetilde{D}_k} = \left[\frac{1 - \hat{\sigma}^1 \hat{x}_k^1 / x_k^*}{1 - \hat{\sigma}^1 \hat{D}_k^1 / D_k^*}\right]^{-1} \cdot \frac{\hat{x}_k/\hat{D}_k}{x_k^*/D_k^*}.$$

Now, $\frac{\hat{x}_k/\hat{D}_k}{x_k^*/D_k^*}$ is decreasing in k, which means that the bracketed term is increasing in k, and hence its inverse is also decreasing. Thus, $\frac{\widetilde{x}_k/\widetilde{D}_k}{x_k^*/D_k^*}$ is overall decreasing in k.

We now repeat the same construction on $\widetilde{D}$, and again until we reach $v_{\hat{k}}$, the crossing point of $\overline{h}_k$ and h_k^*, at which point we set $\hat{\sigma}^{\hat{k}}$ to be all remaining weight. This completes the construction of $\hat{\sigma}^j$, proving the proposition. ∎

Proof of Theorem 2. The proof mirrors that of Theorem 1. We first apply Lemma A.1, so all we need to do is to show that the D solving the maximization problem can be implemented by a segmentation over regular markets. The KKT conditions on h_k^D imply that

$$\frac{d}{dh}[\overline{\omega}'_{k,\lambda}(h_k^\sigma)] = \frac{d}{dh}[\omega'_{k,\lambda}(h_k^\sigma)] = \lambda w'_k(h_k^\sigma) + (1-\lambda) u'_k(h_k^\sigma)$$

is increasing in k. As before, decompose

$$\overline{\omega}'_{k+1,\lambda}(h_{k+1}) - \overline{\omega}'_{k,\lambda}(h_k) = \left[\overline{\omega}'_{k+1,\lambda}(h_{k+1}) - \overline{\omega}'_{k+1,\lambda}(h_k + (v_{k+1} - v_k))\right] + \left[\overline{\omega}'_{k+1,\lambda}(h_k + (v_{k+1} - v_k)) - \overline{\omega}'_{k,\lambda}(h_k)\right].$$

We wish to show that the second bracketed term is negative. Again, take $h(v) = h_k + (v - v_k)$, so that

$$\left[\overline{\omega}'_{k+1,\lambda}(h_k+(v_{k+1}-v_k))-\overline{\omega}'_{k,\lambda}(h_k)\right]=\int_{v_k}^{v_{k+1}}\frac{d}{dv}\left[\overline{\omega}'_{v,\lambda}(h(v))\right]\,dv$$

where $\omega_{v,\lambda}$ is the continuous extension of $\omega_{k,\lambda}$ onto all v and similarly for $w_v(h)$. Note that

$$w'_v(h)=-vQ'(v-h)+c'(Q(v-h))Q'(v-h)=-hQ'(v-h).$$

Taking the total differential with respect to v,

$$\begin{aligned}\frac{d}{dv}\left[w'_v(h(v))\right]&=-h'(v)Q'(v-h(v))-h(v)(1-h'(v))Q''(v-h(v))\\&=-Q'(v-h(v))-(1-h'(v))w''_v(h(v))=-Q'(v-h(v))\le 0.\end{aligned}$$

Thus, in any open interval where $\overline{\omega}'_{v,\lambda}(h(v))=\omega'_{v,\lambda}(h(v))$,

$$\frac{d}{dv}\overline{\omega}'_{v,\lambda}(h(v))=\lambda\frac{d}{dv}w'_v(h(v))+(1-2\lambda)\frac{d}{dv}u'_v(h(v))\le 0.$$

On the other hand, whenever $\omega_{v,\lambda}(h(v))<\overline{\omega}_{v,\lambda}(h(v))$, we have that

$$\begin{aligned}\overline{\omega}'_{v,\lambda}(h(v))=\frac{\omega_{v,\lambda}(T(v))-\omega_{v,\lambda}(B(v))}{T(v)-B(v)}=\lambda\cdot\frac{w_v(T(v))-w_v(B(v))}{T(v)-B(v)}+\\(1-2\lambda)\cdot\frac{u_v(T(v))-u_v(B(v))}{T(v)-B(v)}.\end{aligned}$$

We then follow the proof of Lemma A.2, which goes through some slight modifications. For the case where $\omega_{v,\lambda}(h(v))<\overline{\omega}_{v,\lambda}(h(v))$, we are left with

$$\frac{d}{dv}\overline{\omega}'_{v,\lambda}(h(v))=\frac{\partial}{\partial v}\omega_{v,\lambda}(T(v))-\frac{\partial}{\partial v}\omega_{v,\lambda}(B(v))\tag{A.10}$$

where

$$\frac{\partial}{\partial v}\omega_{v,\lambda}(h)=\lambda\left[\frac{\partial}{\partial v}w_v(h)\right]+(1-2\lambda)\left[\frac{\partial}{\partial v}u_v(h)\right]=\lambda Q(v-h)+(1-\lambda)hQ'(v-h).$$

By construction, $\omega'_{v,\lambda}(T(v))=\omega'_{v,\lambda}(B(v))$, which means

$$(1-2\lambda)Q(v-T(v))-(1-\lambda)T(v)Q'(v-T(v))=(1-2\lambda)Q(v-B(v))-(1-\lambda)B(v)Q'(v-B(v)).$$

Substituting this equation into the previous, (A.10) becomes

$$\frac{d}{dv}\overline{\omega}'_{v,\lambda}(h(v)) = (1-\lambda)\big[Q(v-T(v)) - Q(v-B(v))\big] \leq 0,$$

completing the proof. ∎

Proof of Proposition 7. We prove that for any market $x \notin O_\gamma$, $x \notin O_{\gamma'}$. Denote by $u_v(h;\gamma)$ the local information rate function with isoelastic costs:

$$u_v(h;\gamma) = h(v-h)^{\frac{1}{\gamma-1}-1},$$

By Lemma 6, with isoelastic costs, $u_v(h;\gamma) = \overline{u}_v(h;\gamma)$ if and only if

$$h \leq \frac{\gamma-1}{\gamma}v.$$

This condition obviously becomes weaker as γ increases; hence, if it is not satisfied at γ, then it is also not satisfied by $\gamma' < \gamma$.

Now, take $v_k \in \text{supp}(x)$ such that $u'_v(h^x_{k+1};\gamma) < u'_v(h^x_k;\gamma)$. Consider the function $h(v,\gamma)$ defined on $[v_k, v_{k+1}]$ which maintains $u'_v(h(v,\gamma);\gamma) = u'_k(h^x_k;\gamma)$. To characterize h, first compute the scalar derivative:

$$u'_v(h,\gamma) = (v-h)^{\frac{1}{\gamma-1}-1}\left(v - \frac{\gamma}{\gamma-1}h\right).$$

Then, take the total derivative of this expression with respect to v when $h = h(v)$:

$$\frac{d}{dv}\Big[u'_v(h(v);\gamma)\Big] = \left(\frac{1}{\gamma-1}-1\right)(v-h(v))^{\frac{1}{\gamma-1}-2}(1-h'(v))\left(v-\frac{\gamma}{\gamma-1}h(v)\right) + (v-h(v))^{\frac{1}{\gamma-1}-1}\left(1-\frac{\gamma}{\gamma-1}h'(v)\right).$$

The differential equation characterizing $h(v,\gamma)$ is:

$$\frac{d}{dv}\Big[u'_v(h(v,\gamma);\gamma)\Big] = 0 \iff h'(v,\gamma) = \frac{(v-h(v,\gamma)) - \frac{2-\gamma}{\gamma-1}h(v,\gamma)}{2(v-h(v,\gamma)) - \frac{2-\gamma}{\gamma-1}h(v,\gamma)}. \tag{A.11}$$

Note that

$$h(v) \leq \frac{\gamma-1}{\gamma}v \implies h'(v) \geq 0.$$

Furthermore, observe that (A.11) is increasing in γ. Since $x \notin O_\gamma$,

$$h^x_{k+1} > h(v_{k+1};\gamma) \implies h^x_{k+1} - h^x_k > \int_{v_k}^{v_{k+1}} h'(v,\gamma)\, dv > \int_{v_k}^{v_{k+1}} h'(v,\gamma')\, dv. \tag{A.12}$$

Hence, $x \notin O_{\gamma'}$. The strict inclusion follows from observing that we can replace the first inequality of (A.12) with a weak inequality, so if x is on the border of O_γ, then $x \notin O_{\gamma'}$. ∎

Proof of Proposition 8. If $e = (+1, -1)$, then the problem is to maximize profits and minimize consumer surplus, which is achieved by first degree (perfect) price discrimination. So, it suffices to consider the case where $e_1 = -1$. Rewrite the objective function as

$$-\lambda W + (e_2(1-\lambda) + \lambda)U = U + \mu W$$

where

$$\mu \triangleq \frac{-\lambda}{e_2(1-\lambda)+\lambda} \in (-\infty, \infty).$$

Correspondingly, define the local objective

$$\omega_{\mu,k}(h) = u_k(h) + \mu w_k(h).$$

We can compute

$$\omega'_{\mu,k}(h) = Q(v_k - h) - (1+\mu)hQ'(v_k - h),$$

and

$$\omega''_{\mu,k}(h) = (1+\mu)hQ''(v_k - h) - (2+\mu)Q'(v_k - h).$$

We divide the analysis into three cases. First, if $\mu \leq -2$, then $\omega''_{\mu,k}(h) \geq 0$, meaning that the objective is convex everywhere. The concavification is then given by the line segment between 0 and v_k:

$$\overline{\omega}_{\mu,k}(h) = \mu\left(1 - \frac{h}{v_k}\right) w_k(0).$$

This has derivative

$$\overline{\omega}'_{\mu,k}(h) = -\frac{\mu}{v_k} w_k(0) = -\frac{\mu}{v_k}\big(v_k Q(v_k) - c(Q(v_k))\big).$$

This expression is increasing in k, so under (39) we have submodularity and the solution to the concavification bound is $D = D^*$, by an argument identical to that of Lemma 5. The solution is implementable by putting as much weight as possible on the distribution where $h^x_k = v_k$, and then dividing the remaining mass into degenerate distributions with all weight on a single value.

Next, suppose $\mu \geq -1$. We then have $\omega''_{\mu,k} \leq 0$, and $\omega_{\mu,k}$ is concave everywhere. The rest of the analysis is identical to that of Proposition 5; $\omega'_{\mu,k}(h)$ is increasing in v_k, and hence we have both submodularity and implementability in segmentations with gapless support.

The final case is when $\mu \in (-2, -1)$. Now, $\omega_{\mu,k}$ is increasing everywhere, but is concave at h if and only if

$$h \leq \frac{2+\mu}{1+\mu} \cdot \frac{Q'(v_k - h)}{Q''(v_k - h)}.$$

We claim that $\omega_{\mu,k}$ switches from concave to convex exactly once, then remains convex. A sufficient condition for this is that

$$\frac{Q'(v_k - h)}{Q''(v_k - h)} \text{ is decreasing in } h \iff \sup_{\phi} \frac{Q'(\phi)Q'''(\phi)}{(Q''(\phi))^2} \leq 1.$$

This is satisfied by the isoelastic cost function with $\gamma \geq 2$. Thus:

$$\overline{\omega}_{\mu,k}(h) = \begin{cases} \omega_{\mu,k}(h) & \text{if } h \leq \hat{h}_k; \\ \omega_{\mu,k}(\hat{h}_k) & \text{otherwise,} \end{cases}$$

where

$$\omega'_{\mu,k}(\hat{h}_k) = -\frac{\omega_{\mu,k}(\hat{h}_k)}{v_k - \hat{h}_k}.$$

We now claim that $\hat{h}_k$ is increasing in k, and hence we have the single-crossing property used in the proof of Proposition 5. Here, we work directly with the functional form:

$$(v_k - \hat{h}_k)^{\frac{1}{\gamma-1}} - \frac{1+\mu}{\gamma-1} \cdot \hat{h}_k (v_k - \hat{h}_k)^{\frac{1}{\gamma-1}-1} = -\frac{(\hat{h}_k + \mu v_k)(v_k - \hat{h}_k)^{\frac{1}{\gamma-1}} - \mu/\gamma (v_k - \hat{h}_k)^{\frac{\gamma}{\gamma-1}}}{v_k - \hat{h}_k}.$$

This simplifies to

$$\hat{h}_k = \frac{(1+\mu) - \frac{\mu}{\gamma}}{\frac{1+\mu}{\gamma-1} - \frac{\mu}{\gamma}} v_k.$$

Importantly, $\hat{h}_k$ only matters when it is positive and less than v_k, and when it is positive it must be an increasing function of v_k.

Lastly, we need to establish that submodularity holds. It is easy to verify that $\omega'_{\mu,k}$ is increasing in v_k whenever $\omega''_{\mu,k} \leq 0$, and hence when $h^*_k \leq \hat{h}_k$ we have submodularity. Otherwise, we need to compute the derivatives at $\hat{h}_k$:

$$\omega'_{\mu,k}(\hat{h}_k) = (v_k - \hat{h}_k)^{\frac{1}{\gamma-1}-1} \left(v_k - \hat{h}_k - \frac{1+\mu}{\gamma-1} \hat{h}_k \right).$$

Since $\hat{h}_k$ is a linear function of v_k, and $v_k - \hat{h}_k \geq 0$ in the relevant parameter region, this is increasing in v_k. This completes the proof. ∎

B Additional Results: Discrete Goods

B.1 Consumer-Optimal Segmentation

We model the discrete good environment by taking $c(q)$ to be convex and piecewise linear between integer values of q; equivalently, we take $c'(q)$ to be a step function. For notational ease, assume that $Q \in \mathbb{N}$, and let $\{\kappa_i\}$, $i \in \{1, \ldots, Q\}$, denote the different values of $c'(q)$.

Such a $c(q)$ is a piecewise linear approximation of some smooth function $\hat{c}(q)$. In particular, we can think of u_k as being equal to the associated $\hat{u}_k$ when $v_k - h \in \{\kappa_i\}$, and otherwise is equal to the linear interpolation between these points. That makes $\overline{u}_k$ itself a piecewise linear approximation of $\text{cav}[\hat{u}_k]$.

This observation by itself is not particularly useful. However, combined with an assumption that replicates the conditions of Section 5, we are able to recover many of the results from before. In particular, if:

$$\kappa_{i+1} - \kappa_i \text{ is increasing} \tag{B.1}$$

then c is a linear approximation of some $\hat{c}$ satisfying (38). Thus, the concavification is equal to a piecewise linear approximation of $\text{cav}[\hat{u}_k]$ up to the maximum, which is

$$i_k^* = \arg\max_{i \in \mathbb{N}} \left[i \cdot (v_k - \kappa_i) \right].$$

We can invert this condition by noting that

$$i_k^* = \max_{i \in \mathbb{N}} \left\{ i \mid i \cdot (v_k - \kappa_i) \geq (i-1) \cdot (v_k - \kappa_{i-1}) \right\} = \max_{i \in \mathbb{N}} \left\{ i \mid v_k - \kappa_i \geq (i-1)(\kappa_i - \kappa_{i-1}) \right\}.$$

This gives a discrete counterpart to Corollary 2.

Corollary 5 (Minimum Quality with Convex Discrete Marginal Cost)
Under (B.1)*, every v_k such that*

$$v_k - \kappa_i \geq (i-1)(\kappa_i - \kappa_{i-1})$$

consumes quality at least i.

By imposing the MHR condition (39), we can also extend Proposition 5.

Proposition B.1 (Optimal Segmentation with Convex Discrete Marginal Cost)
Under (39) *and* (B.1)*, in any consumer-optimal segmentation, for every market* $x \in \text{supp}(\sigma)$ *and every consumer of type* v_k*,*

$$q_k^x \in \begin{cases} \{i^*\} & \text{if } h_k^* \geq v_k - \kappa_{i^*}; \\ \{q_k^*, q_k^* + 1\} & \text{if } h_k^* < v_k - \kappa_{i^*}. \end{cases}$$

Proposition 4 does not go through—this can be seen immediately from Proposition B.1, as the quality provided to the same type may differ by 1 across segments. As a result, generically, some segmentation is always beneficial. This fact is also the main result of Haghpanah and Siegel (2023). However, if we consider a discrete approximation of a continuous cost function that satisfies Proposition 4, the benefits of segmentation are small: equal to the gain from concavifying the hazard rate around the two nearby points of $v_k - \kappa_i$. As the number of goods grows, i.e. as $c(q)$ approaches a smooth function, this gain goes to 0.

B.2 From Concavification to Extreme Points

In this section, we show that concavification can be used to explicitly derive the extreme points in a discrete good setting. We will make some slight adaptations to our notation. With discrete quality, the optimal price menu $p(q)$ will be constant between integer values of q. Thus, we take prices to be an increasing vector $p^x \in \mathbb{R}_+^Q$, where element $p_i^x = p^x(i)$, $i \in \{1, \ldots, Q\}$. Rather than working with p, it is actually easier to work with the *incremental price* $\rho_i^x = p_i^x - p_{i-1}^x$. This is because we know $\rho_i^x = \min\{v \mid \phi^x(v) \geq \kappa_i\}$. Note that ρ_i^x is the lowest consumer type who purchases quality i, and thus is also increasing in i.

This model is equivalent to one of a multi-good monopolist, where the seller sells Q goods which are identical to the consumer but have increasing marginal cost of production κ_i. Then, ρ_i corresponds to the price charged for each good, which is a "quality increment" in our original model. We will sometimes refer to ρ_i as the price of increment or unit i.

Consider the analysis of Section 7 in this setting. Given any increasing sequence κ_i, $w_k(h)$ is constant except at the jump points of $Q(v_k - h)$. Hence, to achieve the Pareto efficient frontier, it is without loss to restrict to distributions such that $v_k - h_k^x \in \{\kappa_i\}$; these are the only hazard rates necessary to obtain the concavification of $\omega_{k,\lambda}$, no matter what λ is.

Lemma B.3 (Sufficient Distributions)
When c is discrete, the Pareto frontier can be achieved with segmentations supported only on distributions x such that x is regular and for every k, $v_k - h_k^x \in \{\kappa_i\}$*.*

These distributions have a very particular structure: for every k, $h_k^x = v_k - \kappa_i$ for some

$i \in \{1, \ldots, Q\}$, where κ_i increases with k. Equivalently, there is some partition of V into i different (possibly empty) sections $R_i = [r_i, r_{i+1})$, such that for any $v_k \in R_i$,

$$v_k - h_k^x = \kappa_i \implies D_{k+1}^x = \frac{v_k - \kappa_i}{v_{k+1} - \kappa_i} \cdot D_k^x.$$

This distribution can be written explicitly as

$$D_k^x = \frac{A_i}{v_k - c_i} \quad \forall v_k \in (r_i, r_{i+1}]$$

where $A_1 \triangleq r_1 - c_1$ and

$$A_{i+1} \triangleq A_i \frac{r_{i+1} - c_{i+1}}{r_{i+1} - c_i}.$$

Within the sections R_i, this distribution is what is known as the *generalized Pareto distribution*, defined by:

$$F(x) = 1 - \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi}$$

where the parameters μ, σ, ξ are called the location, scale, and shape, respectively.

Let D_r^Q denote the set of piecewise generalized Pareto distributions with support Q, where r is the vector of cutoffs $r_i \geq \kappa_i$ which define the regions R_i. In Figure B.1 we illustrate a few examples of the piecewise Pareto demand functions on the continuous value set $V = [0, 1]$. By construction, within each segment, the marginal revenue of selling an additional unit of quality to type v_k is κ_i when $v_k \in R_i$, meaning that $\rho_i^x = r_i$.

Intuitively, these distributions place the greatest amount of mass possible to the right of r_i without violating the requirement that $\rho_i^x = r_i$. This intuition is formalized below.

Proposition B.2 (Stochastic Dominance)
Take any $x \in \Delta S$ and r such that $r_i \geq \rho_i^x$ for all i. Then:

$$D^x(v_k) \leq D_r^Q(v_k), \ \forall v_k \in Q.$$

That is, x is first-order stochastically dominated by D_r^Q.

Proof. For our discussion of discrete qualities, we need some additional notation. Let

$$\Pi_i(x, \rho_i) \triangleq D^x(\rho_i)(\rho_i - \kappa_i)$$

be the profit attributable to selling the ith quality increment at the incremental price ρ_i. We

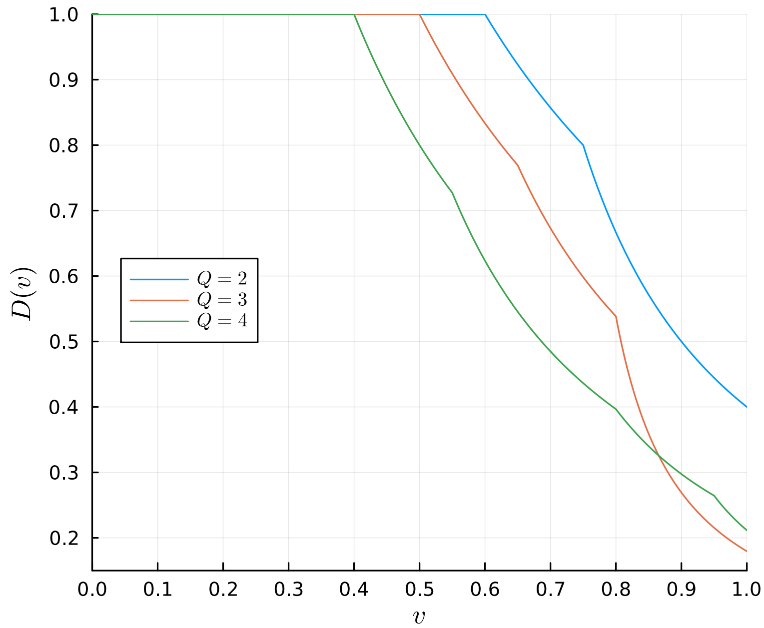

Figure B.1: Examples of piecewise Pareto demands.

also denote the total profit by

$$\Pi(x, \rho) \triangleq \sum_i \Pi_i(x, \rho_i).$$

We claim that given any two vectors $\rho, \hat{\rho}$ with $\rho_i \leq \hat{\rho}_i$ for all i, $D_{\hat{\rho}}$ first-order stochastically dominates any $x \in X_\rho$. That is, for all $v_k \in V$, $D^x_k \leq D_{\hat{\rho}}(v_k)$. Suppose, by contradiction, there exists $\hat{v}$ such that $D^x(\hat{v}) > D_{\hat{\rho}}(\hat{v})$, and without loss assume $\hat{v}$ is the smallest such value. Let ρ_i be such that $\hat{v} \in (\rho_i, \rho_{i+1}]$. Note that $\hat{v} > \hat{\rho}_1 \geq \rho_1$, so such ρ_i exists. We then have that:

$$\Pi_i(x, \hat{v}) > \Pi_i(D_{\hat{\rho}}, \hat{v}) \geq \Pi_i(D_{\hat{\rho}}, \rho_i) \geq \Pi_i(x, \rho_i).$$

The first inequality follows from $D^x(\hat{v}) > D_{\hat{\rho}}(\hat{v})$. The second inequality follows from Lemma B.4 and using that $\hat{v} \in [\rho_i, \rho_{i+1}]$ so $\hat{v} < \hat{\rho}_{i+1}$. The third inequality follows from the fact that $D^x(v) \leq D_{\hat{\rho}}(v)$ for $v < \hat{v}$. But, this implies that ρ_i is not an optimal price for increment i, a contradiction. Hence, for any $x \in X_\rho$, $D_{\hat{\rho}}$ first-order stochastically dominates D^x. Finally, it is immediate that D_r first-order stochastic dominates any D^Q_r where $Q \subset V$. ∎

This property of the piecewise generalized Pareto also means that it solves the multi-unit version of the consumer maximization problem studied in Condorelli and Szentes (2020), a feature we discuss in Section B.3.

Next, we define the extreme point problem under consideration. Given ρ, we can define the (possibly empty) set of markets where ρ is optimal:

$$X_\rho = \{x \mid \rho^x = \rho\} = \left\{x \mid p_i^x = \sum_{j \le i} \rho_j\right\}.$$

As mentioned previously, the set X_ρ is a compact, convex subset of $\mathbb{R}_+^K$, and equal to the convex hull of its extreme points.

Lemma B.3 states that the set of all such D_r^Q, where we vary over all $Q \subseteq V$ and increasing sequences $r \in Q^Q$, is sufficient to achieve all Pareto efficient segmentations. This is because they are the only distributions needed to support $\overline{\omega}_{k,\lambda}$ for any λ. It turns out that these distributions are also exactly the extreme points of X_ρ.

Proposition B.3 (Extreme points of X_p)
With discrete qualities, x is an extreme point of X_ρ if and only if $D^x = D_r^Q$ for some $Q \subseteq V$ and $r_i \le \rho_i \le r_{i+1}$ for all $i \in \{1, 2, \ldots, Q\}$.

Proof. We begin with two lemmas. The first tells us how profits co-move with prices and costs. The second establishes that X_ρ is nonempty if and only if D_ρ exists.

Lemma B.4 (Co-monotonicity of Prices and Costs)
For any i, j such that $c_i > c_j$ and $\rho_i, \hat{\rho}_i$ such that $\rho_i > \hat{\rho}_i$ (or $c_i < c_j$ and $\rho_i < \hat{\rho}_i$):

$$\Pi_i(x, \rho_i) \le \Pi_i(x, \hat{\rho}_i) \implies \Pi_j(x, \rho_i) < \Pi_j(x, \hat{\rho}_i).$$

Proof. We have:

$$\Pi_i(x, \rho_i) \le \Pi_i(x, \hat{\rho}_i) \implies \frac{D^x(\rho_i)}{D^x(\hat{\rho}_i)} \le \frac{\hat{\rho}_i - c_i}{\rho_i - c_i} < \frac{\hat{\rho}_i - c_j}{\rho_i - c_j} \implies \Pi_j(x, \rho_i) < \Pi_j(x, \hat{\rho}_i),$$

as desired. The case of $c_i < c_j$ can be proven similarly by inverting the ratios. ∎

Lemma B.5 (Existence of X_ρ)
X_ρ is non-empty if and only if for all $i \in \{1, .., Q\}$: (i) $\rho_i \le \rho_{i+1}$ and (ii) $c_i < \rho_i$.

Proof. For prices satisfying these conditions, $D_\rho \in X_\rho$ and so X_ρ is non-empty. For necessity, $\rho_i < c_i$ is obviously never optimal, and $\rho_i > \rho_{i+1}$ cannot be optimal by Lemma B.4. ∎

We are now ready to prove the proposition.

($\Rightarrow$). First, we show that if $x \in X_\rho$ is an extreme point, then every $v_k \in \text{supp}(x)$ is an optimal price for some unit i. To see this, suppose there is a v_k such that $x_k > 0$, but, for every i,

$$v_k \notin \arg\max_{\rho_i} \Pi_i(x, \rho_i).$$

Consider the following market segmentation with uniform distribution and binary support in markets $\{x_-, x_+\}$ defined as follows:

$$D_{x_-}(v) = \begin{cases} D(v) & \text{if } v \neq v_k; \\ D(v_k) - \epsilon & \text{if } v = v_k, \end{cases} \quad D_{x_+}(v) = \begin{cases} D(v) & \text{if } v \neq v_k; \\ D(v_k) + \epsilon & \text{if } v = v_k. \end{cases}$$

The segmentations clearly conform to the aggregate market. Furthermore, if ϵ is small enough, then we continue to have that for all i,

$$v_k \notin \arg\max_{\rho_i} \Pi_i(x_-, \rho_i) \text{ and } v_k \notin \arg\max_{\rho_i} \Pi_i(x_+, \rho_i).$$

Hence, $x_+,\ x_- \in X_\rho$, contradicting that x is an extreme point. Thus, every extreme point satisfies:

$$D_k^x = \frac{\Pi_i(x, \rho_i)}{v_k - c_i},$$

for some i. To finish the proof, we need to show that the constants $\Pi_i(x, \rho_i)$ are consistent with the piecewise Pareto. This is equivalent to showing that $\overline{\rho}_i^x = \underline{\rho}_{i+1}^x$ whenever $c_i < c_{i+1}$, where

$$\overline{\rho}_i^x = \max\left\{\arg\max_{\rho_i} \Pi_i(x, \rho_i)\right\}$$

is the largest optimal price for increment i, and similarly for $\underline{\rho}_i^x$. Suppose $\overline{\rho}_i^x < \underline{\rho}_{i+1}^x$, and let i be the smallest index where this fails. Consider the following segmentation, again with uniform distribution over binary support:

$$D_{x_-}(v) = \begin{cases} D^x(v) & \text{if } v < \underline{\rho}_{i+1}^x; \\ \kappa D^x(v) & \text{if } v \geq \underline{\rho}_{i+1}^x, \end{cases} \quad D_{x_+}(v) = \begin{cases} D^x(v) & \text{if } v < \underline{\rho}_{i+1}^x; \\ (2-\kappa) D^x(v) & \text{if } v \geq \underline{\rho}_{i+1}^x, \end{cases}$$

where $\kappa < 1$. We claim that if κ is sufficiently close to 1, $x_-, x_+ \in X_\rho$.

Take x_- (a similar argument works for x_+). If any prices changed, it is of some unit in $\{i+1, \ldots, n\}$. But, for any unit $j \geq i+1$ and $v_k < \underline{\rho}_{i+1}^x$,

$$\Pi_j(x, \rho_j) \geq \Pi_j(x, \underline{\rho}_{i+1}^x) > \Pi_j(x, \overline{\rho}_i^x) \geq \Pi_j(x, v_k),$$

where the first inequality comes from optimality of price ρ_j, and the next two follow from Lemma B.4. Thus, for κ sufficiently close to 1, we have $x_-, x_+ \in X_\rho$, a contradiction.

($\Leftarrow$). By Proposition B.2, D_ρ first-order stochastically dominates every element of X_ρ. Hence, it cannot be written as the convex combination of two elements in X_ρ. Additionally, by Lemma B.5, D_ρ exists if and only if X_ρ is non-empty, and hence has extreme points. ∎

We can visualize this result using the simplex. Figure B.2 represents the simplex ΔV when there are three values and two products. The different regions X_ρ are identified by a vector that identifies the optimal (incremental) prices in each region.

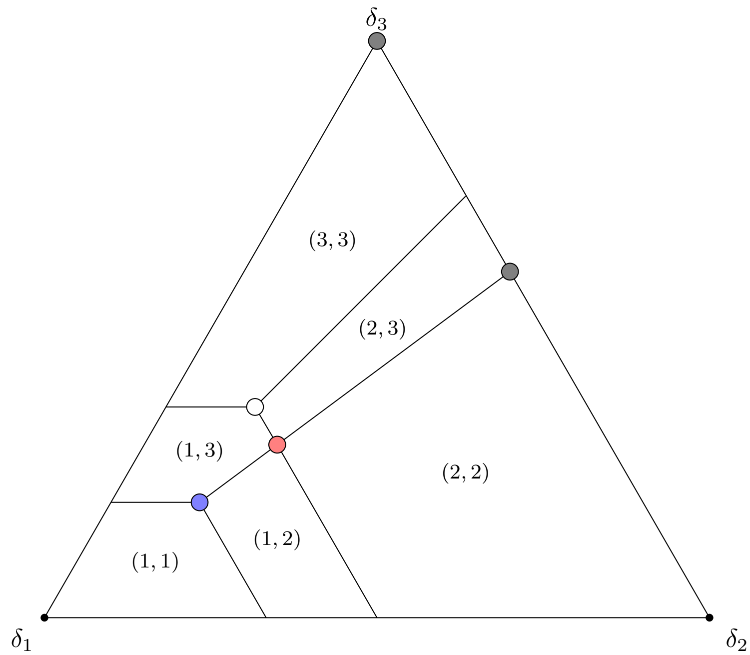

Figure B.2: Division of ΔV into X_ρ with $V = \{1, 2, 3\}$ and $c = (0, \frac{1}{3})$.

The extreme points of X_ρ are the vertices of each price region. In Figure B.2 there are three interior extremal markets, illustrated with white, red and blue dots. The white dot is a market in which the seller is indifferent between the three prices $\rho \in \{1, 2, 3\}$ for quality 1 and the optimal price for the second unit of quality is $\rho = 3$. The blue dot is a market in which the seller is indifferent between the three prices $\rho \in \{1, 2, 3\}$ for the second unit of quality and the optimal price for the quality 1 is $\rho = 1$. The red dot is a market such that the seller is indifferent between the prices $\rho \in \{1, 2\}$ for quality 1 and is indifferent between the prices $\rho \in \{2, 3\}$ for the second unit. There are other extremal markets in which the support of the distribution is smaller. The extremal markets marked by gray dots illustrate points in which all buyers have value 3, and another one in which the buyer never has a value of 2, but the seller is indifferent between the prices $\rho \in \{2, 3\}$ for the second unit.

Looking at Figure B.2, it is clear that the intersections of the lines are the only points which are not the convex combination of any two other points in the price region. The proof makes this observation rigorous by explicitly checking that every point apart from these intersections can be written as the convex combination of two other points within the same price region. For the points where some value $v_k \in \text{supp}(x)$ is not an optimal price for some unit, we can easily perturb x in opposite directions without changing the optimal price vector. This leaves us with just those markets x which are along the boundary. For these markets, we verify that there always exists two "neighboring" extremal markets such that x lies in the convex combination of the two.

B.3 Unconstrained Consumer Maximization

The *unconstrained consumer maximization* problem asks which distribution x leads to the highest consumer surplus $U(x)$. Formally, it solves

$$\max_{x \in \Delta V} \Big[U(x, \rho^x) \Big] \text{ such that } \rho^x \in \arg\max_{\rho \in V^Q} \big[\Pi(x, \rho) \big]. \tag{B.2}$$

We call this problem *unconstrained* because it is essentially the same as (6) without the majorization constraint.

It is easy to see that increasing the distribution of values will increase profit. In fact, 0 and v_K are the profit when the distribution of values is degenerate at 0 or v_K, respectively, and everything in between is achievable with an intermediate distribution. When we examine how the distribution of values impacts consumer surplus, however, the effect is more subtle. In particular, if the distribution is degenerate, the seller extracts the full surplus. Hence, the problem is not trivial.

The solution to (B.2) with a continuum of values when $Q = 1$ and $c = 0$ is given by Condorelli and Szentes (2020). The optimal distribution is a Pareto distribution with shape 1 and scale parameter $\frac{1}{e}$, truncated on the interval $[\frac{1}{e}, 1]$. With $Q = 1$, the assumption that $c = 0$ is without loss, but for our purposes, it is convenient to "undo" this normalization.

Proposition B.4 (Consumer-Optimal Distribution—Single Good)
The solution to (B.2) *when* $Q = 1$ *is*

$$D^x(v) = \begin{cases} 1 & \text{if } v < \rho; \\ \frac{\rho - c}{v - c} & \text{if } v \geq \rho, \end{cases} \tag{B.3}$$

for some $\rho > c$.

The distribution is constructed as to keep profit constant for every price in $[\rho, v_K]$, and (given this indifference) the seller sells the product at the lowest optimal price ρ. The proof of this result then shows that, given this complete indifference, any other distribution which induces price ρ is first order stochastically dominated by (B.3), and hence gives lower consumer surplus.

We know that the piecewise generalized Pareto distributions also maintain constant profits across the relevant regions, and by Proposition B.2 they also first order stochastically dominate any other distribution with the same or lower prices. Thus, we can apply the same proof for the multi-unit case on these piecewise Pareto distributions.

Proposition B.5 (Consumer-Optimal Distribution)
If (x, ρ) is a solution to the consumer surplus maximization problem (B.2)*, then the optimal demand function is given by D_ρ for some ρ.*

For a given cost vector c, the solution (x, ρ) includes a particular price vector $\rho \in V^Q$ that maximizes the consumer surplus in the market x. Figure B.3 plots the optimal demand function for two products, $Q = 2$, and the cost vector $c = (0, 0.5)$. The resulting solution is compared with the corresponding single product solutions for $c = 0$ and $c = 0.5$. In general, the optimal values of ρ_i are non-elementary expressions of c.

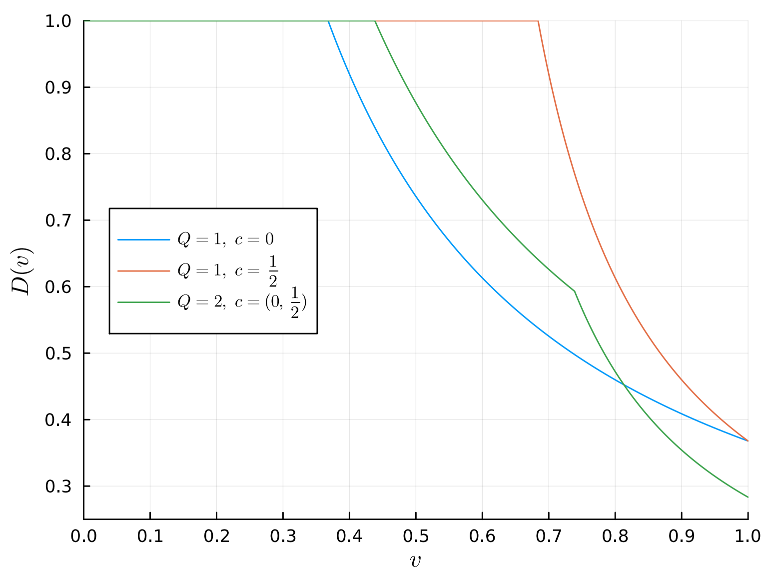

Figure B.3: Consumer-Optimal demands.

We can provide a sharper characterization of the optimal solution than offered by Proposition B.2 in the case of binary products, $Q = 2$. When there are only 2 product for sale, it is without loss of generality to normalize $c_1 = 0$, so that the model is parametrized only by c_2, which we refer to c in the remainder of this Subsection. It is useful to consider the limit as the set of values becomes more refined. For this purpose, we define:

$$\Delta \triangleq \max_k \{v_{k+1} - v_k\}$$

as we let the number of values become arbitrarily large, $K \to \infty$.

Proposition B.6 (Prices with Two Goods)
When $Q = 2$, there exists $\bar{c}$ such that in the consumer-optimal distribution of values, the seller bundles the products ($\rho_1 = \rho_2$) if and only if $c \leq \bar{c}$. Furthermore,

$$\lim_{\Delta \to 0} \bar{c} \to \frac{v_K}{1+e}.$$

Proof. We first compute $U(x)$ for extremal markets $x = D_p$ in the limit where $V = [0,1]$. The consumer surplus from selling unit i to consumers with valuations $v \in [\rho_j, \rho_{j+1})$ is:

$$A_j \int_{\rho_j}^{\rho_{j+1}} \frac{v - \rho_i}{(v - c_j)^2}\, dv = A_j \left(\log \frac{\rho_{j+1} - c_j}{\rho_j - c_j} - \frac{(\rho_i - c_j)(\rho_{j+1} - \rho_j)}{(\rho_{j+1} - c_j)(\rho_j - c_j)} \right).$$

The total consumer surplus is the above expression summed over all i and $j \geq i$, plus the consumer surplus of the mass point at the end, $\frac{A_Q}{1-c_Q} \sum_i (1 - \rho_i)$. Summing everything together, we get:

$$U(x,p) = \sum_{i=1}^{n} \left[iA_i \log \frac{\rho_{i+1} - c_i}{\rho_i - c_i} + \frac{A_Q}{1 - c_Q}(1 - \rho_i) - \sum_{j \geq i} A_j \frac{(\rho_i - c_j)(\rho_{j+1} - \rho_j)}{(\rho_{j+1} - c_j)(\rho_j - c_j)} \right].$$

We claim that the non-logarithmic terms cancel out. To see this, fix an i, then take the difference between the middle term and the $j = Q$ summand of the last term:

$$\frac{A_Q}{1 - c_Q} \left[1 - \rho_i - \frac{(\rho_i - c_Q)(1 - \rho_Q)}{\rho_Q - c_Q} \right] = \frac{A_Q}{\rho_Q - c_Q}(\rho_Q - \rho_i) = \frac{A_{n-1}}{\rho_Q - c_{n-1}}(\rho_Q - \rho_i).$$

Repeat this recursively for remaining j, and at the last step the sum becomes 0, giving us

$$U(x,p) = \sum_{i=1}^{Q} \left[iA_i \log \frac{\rho_{i+1} - c_i}{\rho_i - c_i}. \right] \tag{B.4}$$

(B.4) reduces finding the optimal p to a constrained optimization problem, subject to $\rho_1 \leq \rho_2$. The KKT conditions are:

$$\mu = \log \frac{\rho_2}{\rho_1} + 2\left(1 - \frac{c}{\rho_2}\right) \log \frac{1 - c}{\rho_2 - c} - 1 = \frac{\rho_1}{\rho_2} - 2\left(\frac{\rho_1 c}{\rho_2^2}\right) \log \frac{1 - c}{\rho_2 - c} \tag{FOC}$$

$$\mu(\rho_2 - \rho_1) = 0. \tag{CS}$$

We can check when $\rho_1 = \rho_2$ is a solution. Plugging in $\rho_1 = \rho_2 = \rho$ yields $\rho = \frac{1}{e}(1 - c) + c$.

Dual feasibility then requires that:

$$\mu = \frac{2(1-c)}{1+(e-1)c} - 1 \geq 0 \iff c \leq \frac{1}{1+e}$$

as desired. ∎

This result can be compared with Haghpanah and Hartline (2021), which discusses conditions under which bundling is optimal for the seller. Figure B.4 plots the prices the seller charges against the consumer-optimal distribution as a function of c, in the limit where $V = [0, 1]$. We can see that when the units have a similar cost, it is optimal to force the seller to bundle the units together. But, as c increases above $\bar{c} = \frac{1}{1+e}$, the price of unit 1 falls while the price for unit 2 rises.

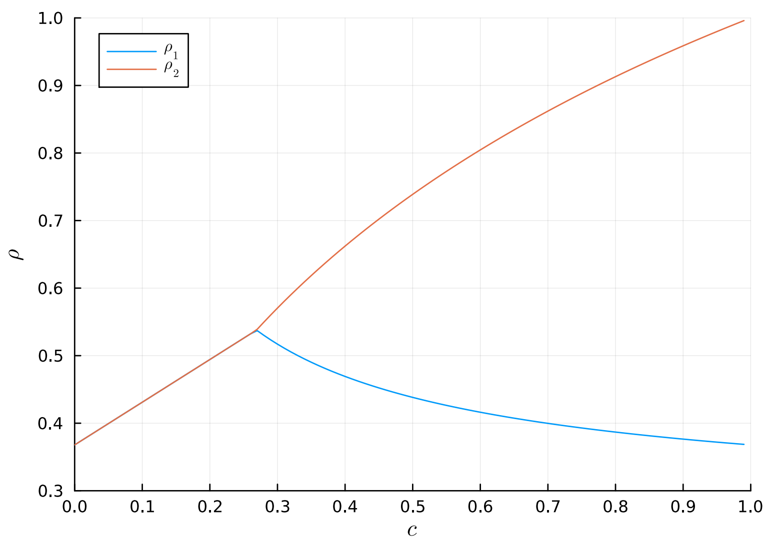

Figure B.4: Prices under consumer-optimal demand when $Q = 2$.