Good Data and Bad Data: The Welfare Effects of Price Discrimination

Maryam Farboodi[*] Nima Haghpanah[†] Ali Shourideh[‡]

February 14, 2025[§]

Abstract

We ask when additional data collection by a monopolist to engage in price discrimination monotonically increases or decreases weighted surplus. To answer this question, we develop a model to study endogenous market segmentation subject to residual uncertainty. We give a complete characterization of when data collection is good or bad for surplus, which consists of a reduction of the problem to one with only two demand curves, and a condition for the two-demand-curves case that highlights three distinct effects of information on welfare. These results provide insights into when data collection and usage for price discrimination should be allowed.

[*]MIT Sloan, NBER and CEPR. Email: farboodi@mit.edu.

[†]Penn State. Email: nima.haghpanah@gmail.com.

[‡]Tepper School of Business, Carnegie Mellon University. Email: ashourid@andrew.cmu.edu.

[§]We thank Alessandro Bonatti, Roberto Corrao, Stephen Morris, Christopher Palmer, Jonathan Parker, Maryam Saeedi, and Alexander Wolitzky for their thoughts and comments.

1 Introduction

The rise of big data technologies, allowing firms to collect detailed consumer data to estimate their willingness to pay, has reignited the longstanding debate on the welfare implications of price discrimination. The prevalence of these practices has raised concerns among policy-makers about big tech's exploitation of consumer data. Quoting a letter that followed a senate hearing on May 2, 2024:

> As more consumers shop online, large tech platforms have access to vast stores of personal data [. . .] that can be exploited by corporations to set prices based on the time of day, location, or even the electronic device used by a consumer.[1]

The Federal Trade Commission also recently issued an order seeking information from several companies about their "segmentation solutions" that categorize consumers based on location, demographics, and credit history and set different prices for the same good or service.[2]

A significant difficulty in regulating data collection practices is that it is close to impossible to perfectly monitor and control how firms use consumer data, which makes highly targeted regulation impractical. Often the relevant question is if data collection should be permitted, without knowing how much information the firm already has nor how much additional information it might be able to collect. It is clear, however, that if additional consumer data collection is guaranteed to be beneficial (harmful) *regardless* of the firm's existing or additional information, the it should (should not) be permitted. As a first step towards informed policy design, we seek to characterize when such guarantees can be provided.

To answer this question, we develop a model to study endogenous market segmentation by a monopolist subject to residual uncertainty. There is a given set of consumer types each represented by a downward-sloping demand curve specifying the distribution of consumers' valuations of that type. The seller has access to some information structure that maps types to signal realizations, allowing her to segment the market and charge a profit-maximizing price for each segment. Types in our setting represent everything that is possibly knowable, e.g., a complete profile of consumer characteristics, and a segmentation reflects what the seller actually knows, e.g., perhaps only consumer locations. The heterogeneity of values of consumers of the same type, represented by the downward-sloping demand curve, is residual uncertainty that cannot be resolved even if the seller perfectly observes types. This residual uncertainty reflects the practical limitations, legal and technological, that sellers face in perfectly predicting individual willingness to pay.

Our model bridges the classical and modern approaches to price discrimination. The classical approach to this problem, pioneered by Pigou (1920)'s foundational work, compares only the two extreme cases where the seller is either fully informed or fully uninformed of the

[1]See the hearing's follow-up letter to Amazon for details.
[2]See the order for details.

types. The modern approach to this problem, pioneered by the seminal work by Bergemann et al. (2015), studies endogenous market segmentation in a setting where the seller's has access to signals about consumer *values*. In their setting, all consumer values can be perfectly learned, which may overstate the seller's ability to extract surplus in scenarios with constraints on market segmentation. Our framework merges these two approaches and allows us to study endogenous market segmentation that is subject to residual uncertainty.

Within this framework, we examine how additional information impacts welfare through a pair of opposing properties: "monotonically bad" and "monotonically good" information. Information is "monotonically bad" if every refinement of any segmentation reduces weighted surplus — a convex combination of consumer and producer surplus. Conversely, information is "monotonically good" if weighted surplus is higher for any refinement. We refer to these two properties as surplus-monotonicity properties.

Our main result characterizes each surplus-monotonicity property and has two parts. The first part of the result reduces the problem, for an unrestricted set of demand curves, to one with only two demand curves. It says that surplus-monotonicity holds if and only if three conditions are satisfied. First, the demand curves cannot be too far apart in the sense that the optimal monopoly price of each of them is in the interior of any other's domain of prices. Second, the set of all demand curves is decomposable into at most two basis demand curves. Third, the two basis demand curves satisfy the surplus-monotonicity condition themselves. Given the reduction in the first part of the result, the second part then identifies a closed-form expression that characterizes when a given pair of demand curves satisfies surplus-monotonicity.

As an application of our result, we study a family of demand curves constructed from additive and multiplicative shifts of a single initial demand, and show that the surplus-monotonicity characterization simplifies considerably. This simplified characterization implies that if the initial demand curve has a log-concave density function, information is monotonically bad not only for consumer but also for total surplus. This result confirms and generalizes results from the classical approach to price discrimination (Cowan, 2007; Aguirre et al., 2010), and in particular the seminal work of Pigou (1920) showing that price discrimination reduces total and consumer surplus if demand curves are linear. We also use this characterization to illustrate how information can have opposing effects on total and consumer surplus, increasing the former but decreasing the latter.

In another application, we relate our results to the finding of Bergemann et al. (2015) that for a family of unit-demand curves, information is neither monotonically good nor bad. Our analysis does not apply directly to their setting because it relies on the assumption that each demand curve is downward sloping with a concave revenue function, which is violated by the unit-demand curves considered in their setting. However, we show that as we approach unit-demand curves in a way that satisfies our assumption, then regardless of how we do so, information is neither monotonically good nor bad, showing the robustness of the findings of Bergemann et al. (2015) that study the limit case.

We further study the implications and the interpretation of the formula that characterizes surplus-monotonicity for two demand curves. As an implication, we study two demand curves with constant elasticity of substitution. We show information is monotonically bad for total surplus exactly when the two elasticities are not too far apart.

The interpretation of the formula that characterizes surplus-monotonicity for two demand curves is that it aggregates and highlights three distinct effects of information. To understand what these three effects are, consider an example. Suppose there are two types, type 1 and type 2, half the consumers have type 1, and the other half have type 2. Suppose without loss of generality that, if the seller could perfectly price discriminate based on the type, she would choose a higher price for type 2 consumers. Without any information, the seller offers a uniform price p to all consumers.

Now consider a segmentation with two segments that partially separates the two types in a way that the first segment contains $\frac{2}{3}$ of type 1 consumers and $\frac{1}{3}$ of type 2 consumers, and the second segment contains the remaining $\frac{1}{3}$ of type 1 consumers and $\frac{2}{3}$ of type 2 consumers. The seller lowers the price to some $p_1 < p$ for the first segment, and increases the price to some $p_2 > p$ for the second segment. Table 1 shows, for each type and each price, the fraction of consumers of that type that face that price, before and after information is provided.

		p_1	p	p_2
Before	Type 1	0	1	0
	Type 2	0	1	0
After	Type 1	$\frac{2}{3}$	0	$\frac{1}{3}$
	Type 2	$\frac{1}{3}$	0	$\frac{2}{3}$

Table 1: The fraction of consumers of each type facing each price before and after information is provided.

The three effects of information are reflected in the table. The first effect is the *within-type price change* effect: For each type, information disperses prices, with some consumers facing a price drop and some consumers facing a price increase. The second effect is the *cross-types price change* effect: The price drop applies asymmetrically, to more type 1 consumers than type 2 ones. The third effect is the *price curvature* effect: The size of the price drop might not be equal to the size of the price increase. The larger the price drop is compared to the price increase, the larger the positive effect of providing information becomes.

The power of the expression that characterizes surplus-monotonicity for two demand curves is that it gives a single formula to aggregate the overall effect of information. But it can also be used to give easy-to-check sufficient conditions for surplus-monotonicity: If each of the three effects is positive (negative), then the overall effect is positive (negative) and information is monotonically good (bad). We relate properties of the demand curves to the sign of each effect: The within-type price change effect is positive (negative) when surplus is a convex (concave) function of price, the cross-type price change effect is positive (negative) when the more elastic demand curve has a higher (lower) level then the less elastic one, and the price curvature effect is positive (negative) when the more elastic demand curve has a more (less) steep marginal revenue curve. We construct a parametric example where each of the effects is positive for some parameters, implying information is monotonically good, and negative for some parameters, implying information is monotonically bad.

The rest of the paper is organized as follows. Section 1.1 reviews the literature. Section 2 describes our model. Section 3 defines our surplus-monotonicity properties and states the main result that gives the reduction to two demand curves and the characterization for two demand curves. Section 4 proves and interprets the two-demands part of the main result. Section 5 explains the reduction part of the main result, providing details of the methodology.

1.1 Related Literature

First and foremost, our work relates to the large literature on monopolistic third-degree price discrimination.[3] As discussed earlier, the classical approach to this problem compares uniform pricing to full segmentation of a given set of demand curves. These papers typically either focus on studying total surplus (Varian, 1985; Aguirre et al., 2010; Cowan, 2016) or consumer surplus (Cowan, 2012; Aguirre and Cowan, 2015).[4] Our general approach allows us to characterize surplus-monotonicity for any weighted combination of consumer and producer surplus, including as special cases consumer and total surplus. Additionally, our framework allows us to study intermediate forms of price discrimination where the types are partially separated. Most importantly, whereas the papers in the literature give sufficient conditions for (weighted) surplus to rise or fall, we give a complete characterization of our surplus-monotonicity conditions.

A growing literature focuses on the modern approach to monopolistic third-degree price discrimination, studying all possible segmentations, pioneered by Bergemann et al. (2015).[5] A common focus is to identify all possible pairs of consumer and producer surplus (Bergemann et al., 2015; Kartik and Zhong, 2023). Yang (2022) studies how a profit-maximizing intermediary sells segmentations to a producer for price discrimination. Haghpanah and Siegel (2022), Haghpanah and Siegel (2023), Asseyer (2024), and Bergemann et al. (2024) study a

[3]Oligopolistic third-degree price discrimination is studied in a literature that is more distant to ours, such as Holmes (1989) and Elliott et al. (2021).

[4]Other objectives have also been considered. Several papers focus directly on total output (Robinson, 1969; Schmalensee, 1981) because it is shown (for example, in Varian, 1985) that a necessary condition for price discrimination to increase total surplus is that it increases total output. The effect of price discrimination on producer surplus is clear: It increases profits. Bergemann et al. (2022) quantify *how much* price discrimination increases a seller's profit.

[5]A literature that is more remote to our work, starting from Roesler and Szentes (2017), studies all information structures when the *buyer* receives information about her type.

multi-product seller, combining second- and third-degree price discrimination. Our conceptual contribution to this literature is to introduce a framework in which segmentations are constrained by some residual uncertainty.

Methodologically, our work builds on the recent literature that uses duality in Bayesian persuasion (Kolotilin, 2018; Dworczak and Martini, 2019; Dworczak and Kolotilin, 2024; Kolotilin et al., 2024).[6] In particular, we convert our surplus-monotonicity properties to a class of Bayesian-persuasion problems that seek to identify when no-information maximizes or minimizes weighted surplus *for all* prior distributions over the given set of demand curves. Because the seller's profit-maximization problem in our setting is identified by its first-order condition, we take the strong-duality results of Kolotilin (2018) and Kolotilin et al. (2024) off-the-shelf to solve these Bayesian persuasion problem. We then show that for no-information to be a solution to the entire class of Bayesian-persuasion problems, our separability condition must hold. The main technical result in Kolotilin et al. (2024) gives a sufficient optimality condition, the "twist" condition, that is reminiscent of, but different than, our separability condition. Our characterization identifies an additional condition on two demands capturing the three effects of information, which has no counter-part in Kolotilin et al. (2024), that together give necessary *and sufficient* conditions for our surplus-monotonicity properties.

2 Model

A single seller sells a good, whose constant marginal cost of production is normalized to zero, to a unit mass of unit-demand consumers. The set of (consumer) types Θ is a compact subset of a linear normed vector space (which trivially holds if Θ is finite). The types are distributed according to a full-support prior distribution $\mu_0 \in \Delta\Theta$. There is a demand curve associated with each type $\theta \in \Theta$ where $D(p, \theta)$ specifies the quantity demanded by type θ consumers when the good is sold at price p. This demand can be thought of as representing the measure of type θ consumers whose willingness to pay for the good is at least p.[7] We denote the family of demand curves by $\mathcal{D} = \{D(p, \theta)\}_{\theta \in \Theta}$. The primitive of our model is a tuple $(\mathcal{D}, \mu_0)$, that is, the family of demand curves and the prior distribution over them.

We make the following assumption on the demand curves throughout the paper.

Assumption 1. *For all types $\theta \in \Theta$, the demand curve $D(\cdot, \theta) : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies the following properties:*

[6] See also Immorlica et al. (2022), Smolin and Yamashita (2022), Saeedi and Shourideh (2024), and Saeedi et al. (2024).

[7] An alternative interpretation of our model is that the seller faces a single consumer with a random type drawn from μ. The consumer has multi-unit demands with non-linear utility for quantity and her utility-maximization problem for a given price induces a demand curve for each type. One can also combine the two interpretations and think about each demand as representing a population of consumers with multi-unit demands.

1. *There exists an interval* $I(\theta) = \left[\underline{p}(\theta), \overline{p}(\theta)\right]$, $0 \leq \underline{p}(\theta) < \overline{p}(\theta) \leq \infty$*, such that* $D(p,\theta)$ *is differentiable and strictly decreasing in* p, $D_p(p,\theta) < 0$*, for values* $p \in I(\theta)$.

2. *For values of* $p < \underline{p}(\theta)$, $D(p,\theta) = D(\underline{p}(\theta),\theta)$ *and for values of* $p > \overline{p}(\theta)$, $D(p,\theta) = 0$.

3. *For values of* $p \in I(\theta)$*, the revenue function* $R(p,\theta) = pD(p,\theta)$ *is strictly concave.*[8]

4. *There exists* $p^*(\theta) \in I(\theta)$ *such that* $R_p(p^*(\theta),\theta) = 0$.

The above assumption implies that marginal revenue $R_p(p,\theta)$ is well defined and decreasing for all values of p, and also that $p^*(\theta)$ is the optimal price set by the seller when the seller faces only type θ consumers.

The demand curve of a population of unit-demand consumers all with the same willingness-to-pay θ, used in Bergemann et al. (2015), is represented by a step function which violates Assumption 1. Later we show how our analysis confirms their results by studying demand curves that pointwise approach step functions while maintaining Assumption 1.

A market $\mu \in \Delta\Theta$ is a probability distribution over types. A segmentation $\sigma \in \Delta\Delta\Theta$ is a Bayes-plausible distribution over markets, that is, $\mathbb{E}_\sigma[\mu] = \mu_0$, and $S(\mu_0)$ denotes the set of all segmentations. The interpretation is that the seller has access to some information structure that reveals a signal about the type of the buyer. The seller then forms a posterior μ in the support of σ and chooses a profit-maximizing price for that posterior. The seller uses a pricing rule $p^* : \Delta\Theta \to \mathbb{R}_+$ that specifies an optimal price for every possible posterior, breaking ties if necessary,

$$p^*(\mu) \in \arg\max_{p \in \mathbb{R}_+} \int_\Theta R(p,\theta)\,\mathrm{d}\mu(\theta), \forall \mu \in \Delta\Theta.$$

We study the weighted surplus that each segmentation induces. For this, let

$$CS(p,\theta) = \int_p^{\overline{p}(\theta)} D(z,\theta)\,\mathrm{d}z$$

denote the surplus of type θ consumers from facing price p. The α-surplus of price p for type θ consumers is a weighted average of consumer surplus and producer surplus with weights $\alpha \in (0,1]$ and $1-\alpha$,

$$V^\alpha(p,\theta) := \alpha \cdot CS(p,\theta) + (1-\alpha) \cdot R(p,\theta).$$

The special case of $\alpha = 1$ corresponds to consumer surplus and the special case of $\alpha = \frac{1}{2}$

[8] This assumption allows us to replace the seller's profit-maximization problem by its first-order condition. Kolotilin et al. (2024) show that the first-order approach remains valid under a slight relaxation of strict concavity, roughly assuming concavity up to a normalization, so our results go through with this weaker assumption. We state our assumption with strict concavity because it is more straightforward and is commonly used in the literature.

corresponds to total surplus.[9] The α-surplus of a segmentation is the expectation of the α-surplus of all price and type pairs in the segmentation, given by

$$V^\alpha(\sigma) := \int_{\Delta\Theta} \int_{\Theta} V^\alpha(p^*(\mu), \theta)\, \mathrm{d}\mu(\theta)\, \mathrm{d}\sigma(\mu).$$

Our main goal in the paper is to characterize families of demands $\mathcal{D}$ for which changes in information, i.e., a refinement of the market segmentation σ, either always decreases or always increases the surplus, $V^\alpha(\sigma)$, as formalized next.

3 Main Result: Surplus-Monotonicity Characterization

Our main result characterizes when allowing the monopolist to use consumer data, to refine segmentation and price discriminate accordingly, monotonically affects (weighted) surplus. In this section we state this main result and briefly explain what it says. The rest of the paper interprets the theorem, proves it, and provides several examples.

Consider two segmentations σ and σ'. We say σ a mean-preserving spread of σ', and write

$$\sigma \underset{\text{MPS}}{\succeq} \sigma',$$

if there exists a joint distribution over pairs of markets $\nu \in \Delta(\Delta\Theta \times \Delta\Theta)$ that induces marginals σ and σ', i.e., $\nu(\cdot, \Delta\Theta) = \sigma(\cdot)$ and $\nu(\Delta\Theta, \cdot) = \sigma'(\cdot)$, and random markets (μ, μ') drawn from ν satisfy $\mathbb{E}[\mu|\mu'] = \mu'$ almost surely. We call σ a "refinement" of σ' because it can be obtained by splitting each market μ' in segmentation σ' into possibly multiple markets in a way that satisfies Bayes-plausibility. If σ is a mean-preserving spread of σ', then we can garble the signals of the information structure that leads to σ to obtain σ' (as shown by Blackwell, 1953), so the information structure that corresponds to σ Blackwell-dominates that of σ'. The *full-information segmentation*, in which each market μ in the support of the segmentation has only a single type in its support, is finer than any segmentation, and any segmentation is finer than the *no-information segmentation* that assigns probability 1 to the prior market μ_0.

Given this definition of refinement, we define our main notion of surplus-monotonicity:

Definition 1 (Surplus-monotonicity in information.). *Consider a given $(\mathcal{D}, \mu_0)$.*

1. *Information is monotonically bad for α-surplus, "α-IMB holds," if*

$$V^\alpha(\sigma) \leq V^\alpha(\sigma'), \qquad \forall \sigma, \sigma' \in S(\mu_0) \text{ such that } \sigma \underset{\text{MPS}}{\succeq} \sigma'.$$

[9] We assume $\alpha > 0$ because our properties of interest are trivial if the entire weight is on producer surplus (we explain why after defining the properties).

2. *Information is monotonically good for α-surplus, "α-IMG holds," if*

$$V^{\alpha}(\sigma) \geq V^{\alpha}(\sigma'), \qquad \forall \sigma, \sigma' \in S(\mu_0) \textit{ such that } \sigma \underset{MPS}{\succeq} \sigma'.$$

When distinguishing the two properties or α is unimportant, we refer to them as "surplus-monotonicity" properties.

The two surplus-monotonicity properties are related to the question of which segmentation maximizes the weighted surplus over all segmentations, that is, which segmentation solves

$$\max_{\sigma \in S(\mu_0)} V^{\alpha}(\sigma).$$

Under α-IMB (α-IMG), the no-information segmentation (respectively, the full-information segmentation) solves the above problem because it is less (more) fine than any other segmentation. But α-IMB (α-IMG) is a stronger property than optimality of the no-information (full-information) segmentation because it requires that any refinement of *any* segmentation leads to a lower (higher) weighted surplus.

As stated in the introduction, our motivation for studying these properties is to shed light on situations where a regulator cannot fully monitor or control a seller's information. The seller might have already collected some information and the regulator might not even know how much additional information the seller might be able to collect. The regulator only gets to decide whether or not to allow the seller to collect and use additional information for price discrimination. A cautious regulator, who wants to provide information only if it is guaranteed to improve its objective, provides information whenever information is monotonically good. A less cautious regulator, who wants to provide information only if it is *plausible* that doing so improves its objective, provides information whenever information is not monotonically bad.

Our main result completely characterizes when the surplus-monotonicity properties hold.[10] The result starts by a simplifying step: It shows that the surplus-monotonicity properties are prior-free, allowing us to drop the prior distribution through the rest of the statement. Formally, prior-freeness means that for any two distributions μ, μ', α-IMG (α-IMB) holds for $(\mathcal{D}, \mu)$ if and only if α-IMG (α-IMB) holds for $(\mathcal{D}, \mu')$.

The main content of the result are statements (i) and (ii). Statement (i) gives a reduction of the problem from any number of demand curves to *binary* families of demand curves. It gives conditions under which we only need to verify surplus-monotonicity for the binary family $\{D(\cdot, \theta)\}_{\theta \in \{\theta_1, \theta_2\}}$ that consists of the two demand curve with the lowest and the highest

[10] Blackwell (1953) implies that information is monotonically good for producer surplus, which is the reason we focus on $\alpha > 0$ throughout the paper.

optimal monopoly price, $p^*(\theta_1) = \min_\theta p^*(\theta)$ and $p^*(\theta_2) = \max_\theta p^*(\theta)$ (ties can be broken arbitrarily). Statement (ii) provides the characterization for binary families.

A condition used in the characterization is the *partial inclusion* condition, which means $\underline{p}(\theta') \leq p^*(\theta) \leq \bar{p}(\theta')$ for all $\theta, \theta' \in \Theta$. We call this the partial inclusion condition because it says that at the optimal price for type θ, some but not all of the demand of type θ' consumers will be served.

We now state the main result and then go over the two main statements of the theorem in more detail.

Theorem 1 (Surplus-Monotonicity). *The surplus-monotonicity properties are prior-free, and we can therefore refer to them as properties of $\mathcal{D}$. Let $\theta_1, \theta_2 \in \Theta$ be the two types with the lowest and the highest optimal monopoly prices in $\mathcal{D}$. The following two statements hold:*

(i) α-IMB (α-IMG) holds for $\mathcal{D}$ if and only if

(A) there is partial inclusion, and

(B) there exist two function $f_1, f_2 : \Theta \to \mathbb{R}_+$ such that

$$D(p, \theta) = f_1(\theta) D(p, \theta_1) + f_2(\theta) D(p, \theta_2)$$

for all θ and $p \in [p^*(\theta_1), p^*(\theta_2)]$, and

(C) α-IMB (α-IMG) holds for the binary family $\{D(\cdot, \theta)\}_{\theta \in \{\theta_1, \theta_2\}}$.

(ii) α-IMB (α-IMG) holds for the binary family $\{D(\cdot, \theta)\}_{\theta \in \{\theta_1, \theta_2\}}$ if and only if there is partial inclusion and

$$V^\alpha(p, \theta_2) - V^\alpha(p, \theta_1) + \frac{-\frac{R_p(p,\theta_1)}{R_p(p,\theta_2)} V_p^\alpha(p, \theta_2) + V_p^\alpha(p, \theta_1)}{-\frac{R_p(p,\theta_1)}{R_p(p,\theta_2)} R_{pp}(p, \theta_2) + R_{pp}(p, \theta_1)} (R_p(p, \theta_1) - R_p(p, \theta_2)) \quad (1)$$

is decreasing (increasing) on $[p^*(\theta_1), p^*(\theta_2)]$.

Let us now explain in more detail what the theorem says. Consider statement (i). Condition (A) of the statement, partial inclusion, says that the monopoly price $p^*(\theta)$ for any type θ cannot be less than the lowest price $\underline{p}(\theta')$ in the support of another type θ' or higher than the largest price $\bar{p}(\theta')$ of type θ'. Roughly speaking, this partial-inclusion condition means that the demand curves cannot be too far from each other. For example, with linear demands $D(p, \theta) = \theta - p$ where and $\underline{p}(\theta) = 0, \bar{p}(\theta) = \theta$, the optimal monopoly price for each type is $p^*(\theta) = \frac{\theta}{2}$, and the partial-inclusion condition requires that

$$\frac{\theta}{2} \leq \theta',$$

for every two types θ, θ'. That is, the set of all possible values of θ must be in some interval $[c, 2c]$ for some constant $c \geq 0$. Notice that this partial-inclusion condition must hold regardless of what α is or which one of the surplus-monotonicity properties we are characterizing.

Condition (B) of statement (i) says that all demand curves in the family must be decomposable into a linear combination of at most two "base" demand curves $D(\cdot, \theta_1), D(\cdot, \theta_2)$ with possibly varying weights, where these base demand curves are those with the lowest and the highest optimal monopoly price. The only possible source of heterogeneity among the demand curves is the pair of weights $f_1(\theta)$ and $f_2(\theta)$, so heterogeneity must be reducible to a two-dimensional sufficient statistic. Notice that this condition must also hold regardless of what the weight is or which one of the surplus-monotonicity properties we are characterizing.

Condition (C) of statement (i) puts additional constraints on the base demand curves that we can use for the decomposition. The binary family of demand curves that consists only of the two base demand curves $D(\cdot, \theta_1), D(\cdot, \theta_2)$ must itself satisfy the corresponding surplus-monotonicity property. This condition therefore reduces the problem of characterizing the surplus-monotonicity properties for an arbitrary class of distributions $\mathcal{D}$ to characterizing them with binary distributions.

Statement (ii) of the theorem characterizes the surplus-monotonicity properties for a binary family of distributions in terms of partial inclusion and monotonicity of an expression that depends on α. The direction of the monotonicity condition depends on which one of the surplus-monotonicity properties we are characterizing.

The entirety of Section 4 is dedicated to understanding the monotonicity condition in Equation (1), but for now, note that the expression can be written in a way that is a bit less compact but clarifies how α-IMB (α-IMG) becomes easier to satisfy as α gets larger (smaller). In particular, substituting the definition of V^α, the expression from Proposition 1 is equivalent to

$$\alpha \left[CS(p, \theta_2) - CS(p, \theta_1) + \frac{-\frac{R_p(p,\theta_1)}{R_p(p,\theta_2)} CS_p(p, \theta_2) + CS_p(p, \theta_1)}{-\frac{R_p(p,\theta_1)}{R_p(p,\theta_2)} R_{pp}(p, \theta_2) + R_{pp}(p, \theta_1)} \left(R_p(p, \theta_2) - R_p(p, \theta_1) \right) \right]$$
$$+ (1 - \alpha) \left[R(p, \theta_2) - R(p, \theta_1) \right].$$

The above expression is a convex combination of two functions with weights α and $1 - \alpha$. The second function $R(p, \theta_2) - R(p, \theta_1)$ is increasing over $(p^*(\theta_1), p^*(\theta_2))$ because $R(p, \theta_2)$ is increasing for any price below the optimal monopoly price $p^*(\theta_2)$ for type θ_2 and $R(p, \theta_1)$ is decreasing for any price above the optimal monopoly price $p^*(\theta_1)$ for type θ_1. Now suppose the convex combination is decreasing, so α-IMB holds, and consider a larger weight $\alpha' \geq \alpha$. Because the convex combination is decreasing but the second term is increasing, the first term must be decreasing. So if we replace α with the larger α', we are increasing the weight of the decreasing function and decreasing the weight of the increasing function, so the resulting combination must be decreasing as well, implying information is monotonically bad for α'.

A similar argument applies for the case of monotonically good information.[11] We summarize this monotonicity property in the following corollary.[12]

Corollary 1. *If α-IMB (α-IMG) holds, then α'-IMB (α'-IMG) holds for any $\alpha' \geq \alpha$ ($\alpha' \leq \alpha$).*

Before diving deeper into each of the two main statements of Theorem 1, in the next two subsections we provide two applications of this result. The first application enables us to relate our result to the literature that only compares the no-information segmentation to the full-information segmentation (Pigou, 1920; Cowan, 2007; Aguirre et al., 2010). The second application enables us to relate to seminal work of Bergemann et al. (2015).

3.1 Demand Curve Family $\{aD + b\}$

Consider the family of demand curves consisting of additive and multiplicative shifts of some base demand curve. This family has two features that makes it particularly amenable to applying Theorem 1. First, it satisfies the separability condition in statement (i) of Theorem 1 because we can write any demand curve in the family as a linear combination (with positive weights) of the two demands in the family that have the lowest and the highest monopoly price. The characterization of the surplus-monotonicity properties therefore reduces to verifying whether the monotonicity condition of the statement (ii) of Theorem 1 holds for the binary family of demand curves that contains only those demands in the family that have the lowest and the highest monopoly price. Second, this monotonicity condition in statement (ii) of Theorem 1 significantly simplifies.

Corollary 2. *Consider the family of demand curves*

$$\mathcal{D} = \Big\{a(\theta)D(p) + b(\theta)\Big\}_{\theta}$$

for some demand curve D. Then α-IMB (α-IMG) holds if and only if

$$(2\alpha - 1)p + \alpha(\frac{pD'(p)}{R''(p)}) \tag{2}$$

is increasing (decreasing) over $[\min_{\theta} p^*(\theta), \max_{\theta} p^*(\theta)]$ and there is partial inclusion.[13]

First, note that as α approaches 0, the expression in Equation (2) approaches $-p$, which is a decreasing function. This indicates that information is monotonically good for small enough values of α, consistent with information being monotonically good for the seller.

[11] Suppose the convex combination is increasing, so α-IMG holds, and consider $\alpha' \leq \alpha$. Replacing α with α' increases the weight of the increasing function in the convex combination, so the resulting combination with weights $\alpha', 1 - \alpha'$ is also increasing.

[12] This result can also be directly shown using Blackwell (1953)'s characterization without using ours. We call it a corollary as a "sanity check" of our characterization and because of its usefulness later.

[13] We denote the first and second derivative of $R(p)$ by $R'(p)$ and $R''(p)$, and similarly $D'(p)$ is the derivative of $D(p)$.

Second, consider the case of total surplus, $\alpha = \frac{1}{2}$. Information is monotonically bad (good) for total surplus if and only if

$$\frac{pD'(p)}{R''(p)}$$

is increasing (decreasing). This condition can be written in terms of log-concavity of the density function of the distribution of values that generates demand D. That is, letting, $f(p) = -D'(p) \geq 0$, the condition is that $p^2 f(p)$ is log-concave (log-convex).[14] For instance, for $D = 1 - p^c$ for $c > 0$, $p^2 f(p)$ is log-concave and information is monotonically bad for total surplus. A sufficient condition for log-concavity of $p^2 f(p)$ is log-concavity of the density function, in which case it follows from Corollary 1 that information is monotonically bad for consumer surplus as well.

It is useful to compare Corollary 2 to existing results. Cowan (2007) and Aguirre et al. (2010) also study the family of demand curves of the form $a(D + b)$ and show that if $\frac{pD'(p)}{R''(p)}$ is increasing, then total surplus is lower under full segmentation than under uniform pricing.[15] Our Corollary 2 generalizes their result in several ways: It shows information is *monotonically* bad under this condition, applies to any weighted combination of consumer and producer surplus, and most importantly gives a tight characterization. In fact, whereas these papers only give sufficient conditions for total surplus to decrease, we use our result to give an example below that covers all possible cases for how information might (monotonically) affect total and consumer surplus.

If information affects total and consumer surplus monotonically, there are three possible cases for what the effects might be: Information is good for both, bad for both, or good for total surplus but bad for consumer surplus (information cannot be good for consumer surplus but bad for total surplus). In the example below, each case is covered for some parameters. Detailed calculations are in Online Appendix B.1.

Example 1 (Impact of Information on Total and Consumer Surplus: Three Cases). *Suppose the family of demands $\mathcal{D}$ satisfies partial inclusion and the density function is given by*

$$f(p) = \frac{c_1(c_2 + c_3 p)^{c_4}}{p^2}.$$

The revenue curve is strictly concave whenever $c_4 \neq 0$ and c_3 and c_4 have the same sign. We show, as summarized in Figure 1, that information is monotonically good for total surplus but bad for consumer surplus if $c_4 < -1$, good for both if $-1 \leq c_4 \leq 0$, and bad for both if $c_4 > 0$.

[14] Because $-D'(p) = f(p)$, we have $R_{pp}(p) = -2f(p) - pf'(p)$. Now notice that $\frac{pD'(p)}{R''(p)}$ is increasing (decreasing) whenever $\frac{R''(p)}{pD'(p)} = \frac{2}{p} + \frac{f'(p)}{f(p)} = \frac{d}{dp}(\log p^2 f(p))$ is decreasing (increasing), that is, when $p^2 f(p)$ is log-concave (log-convex).

[15] Aguirre et al. (2010) give several examples where this condition holds, including exponential, probits and logits, and demand functions derived from log-normal distributions.

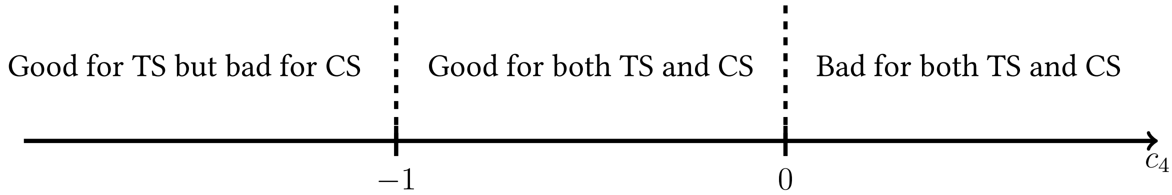

Figure 1: The three possible cases for how information affects total surplus (TS) and consumer surplus (CS) in Example 1.

A convenient feature of Example 1 is that the expression in Equation (2) becomes linear,

$$\Big(\alpha(2+\frac{1}{c_4})-1\Big)p+\frac{\alpha c_2}{c_3 c_4}.$$

Thus, information is monotonically bad (good) whenever the multiplier of p in this expression is positive (negative), that is,

$$\alpha(2+\frac{1}{c_4})-1 \geq (\leq)\ 0.$$

When $-1 \leq c_4 \leq 0$, the multiplier is negative for all α and therefore information is monotonically good regardless of α. Otherwise, that is when c_4 is either above 0 or below -1, there exists a threshold $\widehat{\alpha} \in (0,1)$ such that information is monotonically bad when α is above the threshold, and is monotonically good when α is below the threshold. When $c_4 < -1$, this threshold is between $\frac{1}{2}$ and 1, so information is good for total surplus but bad for consumer surplus. When $c_4 > 0$, this threshold is less than $\frac{1}{2}$, so information is bad for both total and consumer surplus.

A special case of Example 1 is $D(p) = d_1 + d_2 p^{d_3}$.[16] In this case, $\widehat{\alpha} = \frac{d_3+1}{2d_3+3} < \frac{1}{2}$. Therefore information is monotonically bad for total and consumer surplus. Linear demand belongs to this family and arises when $d_3 = 1$. In this case, $\widehat{\alpha} = 0.4$: Information increases α-surplus only for $\alpha < 0.4$, i.e. when producer surplus is over-weighted relative to consumer surplus. This finding confirms and generalizes the foundational result of Pigou (1920) that price discrimination reduces total and consumer surplus if demand curves are linear.

3.2 Approaching Demand Curve Family of Step Functions

Because the demand curves in our setting must be downward-sloping over a non-empty interval, they cannot be step functions, considered in Bergemann et al. (2015). Here we study what happens when we *approach* step functions within our framework. We use this analysis to discuss how our findings align with those of Bergemann et al. (2015).

[16]For D to be decreasing and have a strictly concave revenue curve, we need $d_3 > -1$ and d_2 and d_3 have opposite signs. To see how $D(p) = d_1 + d_2 p^{d_3}$ has a density function of the form $f(p) = \frac{c_1(c_2+c_3 p)^{c_4}}{p^2}$, notice that the density associated with D is $-d_2 d_3 p^{d_3-1}$, so we can let $c_1 = -d_2 d_3, c_2 = 0, c_3 = 1, c_4 = d_3 + 1$.

To do so, let us start with an example.

Example 2 (Demand Curves Approaching Unit Demand). *Consider two demand curves* $D(p, \theta_i) = (\theta_i - p)^c$ *with supports* $[0, \theta_i]$ *for* $i \in \{1, 2\}$*, where* $0 \leq \theta_1 < \theta_2$ *and* $c \in (0, 1]$*.* 1*-IMB holds, i.e. information is monotonically bad for consumer surplus, if and only if there is partial inclusion.*

Details are in Online Appendix B.2. As c approaches 0, each demand curve converges to a step function that is positive and constant at any price $p < \theta_i$ but jumps to zero at $p > \theta_i$. This limit step function represents the demand from a population of consumers all with the same value θ_i for the product. Bergemann et al. (2015) consider these demand curves and show that as long as the prior probability of step function θ_2 is high enough, then there exists some segmentation that increases consumer surplus relative to no segmentation. This result might appear in contrast with our finding that information is monotonically bad regardless of the prior distribution.

The key observation is that surplus-monotonicity of information *requires partial inclusion.* Partial inclusion is violated for small c, which means information is *not* monotonically bad when the demand curves become close enough to step functions. This is because the optimal monopoly price for each demand curve is $\frac{\theta_i}{1+c}$, so the partial-inclusion condition becomes $\frac{\theta_2}{1+c} \leq \theta_1$, which is violated for small c because $\theta_1 < \theta_2$.

But what if we approach step functions some other way? Based on the above discussion, it might appear that we might be able to "fix" the partial-inclusion violation by extending the supports of the demand curves and adding a vanishingly small downward sloping demand curve, if necessary, to each. However, doing so necessarily violates the concavity of the revenue curves over the extended supports, and more generally the validity of the first-order approach upon which our analysis rests. In other words, to approach step functions while maintaining concavity of revenue, the partial inclusion condition must be violated, and therefore neither surplus-monotonicity property holds, consistent with the findings of Bergemann et al. (2015) in the limit case.[17]

To see why, suppose we approach two step functions with steps at $\theta_1 < \theta_2$ in a way that respects partial inclusion, as shown in Figure 2, (a). Because the optimal price for the second step function approaches θ_2, to ensure partial inclusion, the support of the first demand curve must reach θ_2 as well. Now consider the revenue curve associated with a market μ. As shown in Figure 2, (b), this revenue curve has two local maxima and a local minimum. In particular, not only is this revenue curve not concave, but also it violates the weaker aggregate single-

[17]Even though the findings of Bergemann et al. (2015) are not stated in terms of surplus-monotonicity, they imply that both surplus-monotonicity properties are violated. In particular, any segmentation in which some market has price θ_2 can be refined in a way that increases consumer surplus and therefore weighted surplus. And any segmentation in which some market has price θ_1 can be refined in a way that reduces consumer surplus while keeping producer surplus fixed, reducing weighted surplus for any weight.

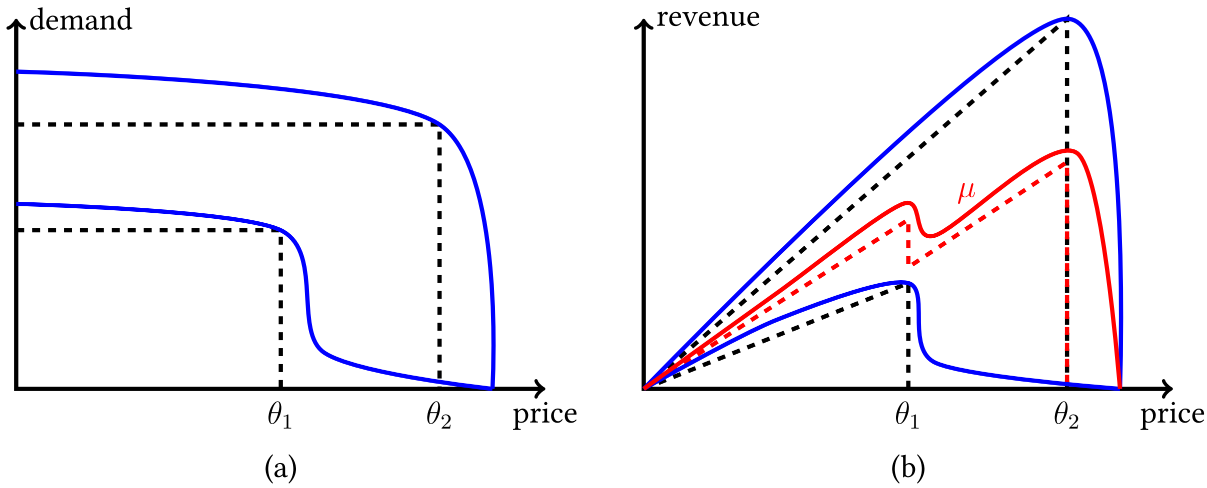

Figure 2: Approaching step functions with partial inclusion violates concavity of revenue.

crossing property that is needed to ensure the validity of the first-order approach.[18] This discussion is summarized in the following corollary.

Corollary 3. *Consider a family of demand curves $\mathcal{D}^\epsilon$ that uniformly converges to a family of unit-demand curves as ϵ goes to zero while satisfying our Assumption 1 for every $\epsilon > 0$. Then, for small enough ϵ, the partial-inclusion condition is violated and therefore information is neither monotonically good nor bad.*

To summarize, approaching step functions within our framework, i.e., with concave revenue curves, necessarily leads to a violation of the partial-inclusion condition, meaning that neither surplus-monotonicity condition holds. This is in line with the findings of Bergemann et al. (2015) in the limit case and shows the robustness of their findings.

4 Two Demand Curves

In this section and the next one we visit the two statements of Theorem 1 and sketch the proof of each statement.

This section considers two demand curves, establishing and interpreting statement (ii) of Theorem 1. In order to focus on the monotonicity expression of the statement, which allows us to provide intuition about the effects of price discrimination, we assume that there is partial inclusion. In Section 5 we show that partial inclusion is in fact necessary for both of the surplus-monotonicity properties. Assuming partial inclusion, the following is a restatement of statement (ii) of Theorem 1 that is repeated here for convenience.

[18]This condition, defined in Kolotilin et al. (2024) says that if marginal revenue is zero at some price, it must be positive for lower prices and negative for higher prices.

Proposition 1 (Surplus-Monotonicity with Two Demand Curves). *Suppose there is partial inclusion. α-IMB (α-IMG) holds for a binary family $\mathcal{D} = \{D(\cdot, \theta)\}_{\theta \in \{\theta_1, \theta_2\}}$, where without loss of generality $p^*(\theta_1) \leq p^*(\theta_2)$, if and only if*

$$V^\alpha(p, \theta_2) - V^\alpha(p, \theta_1) + \frac{-\frac{R_p(p,\theta_1)}{R_p(p,\theta_2)} V_p^\alpha(p, \theta_2) + V_p^\alpha(p, \theta_1)}{-\frac{R_p(p,\theta_1)}{R_p(p,\theta_2)} R_{pp}(p, \theta_2) + R_{pp}(p, \theta_1)} (R_p(p, \theta_1) - R_p(p, \theta_2)) \tag{3}$$

is decreasing (increasing) on $(p^*(\theta_1), p^*(\theta_2))$.

Notice that Proposition 1 says that if $p^*(\theta_1) = p^*(\theta_2)$ then both surplus-monotonicity properties hold regardless of what α is because any function is monotone over a degenerate interval. This makes sense because if the two demands have the same optimal price, then the seller will optimally choose that same price for any market, so segmentation has no effect on price and therefore on consumer and producer surplus, and both surplus-monotonicity properties trivially hold.

Before explaining Proposition 1 in more detail, Let us use it to get tight conditions for surplus-monotonicity of constant-elasticity demand curves, widely used in macroeconomics and empirical work for demand estimation.

Example 3 (CES demand: α-IMB). *Consider two demand curves $D(p, \theta_i) = (c + p)^{-\theta_i}$ for $i \in \{1, 2\}$ and $\theta_1 > \theta_2 > 1$ and some constant $c > 0$. Then $\frac{1}{2}$-IMB holds, i.e., information is monotonically bad for total surplus, if and only if $\theta_1 \leq \theta_2 + \frac{1}{2}$. Under this condition, α-IMB holds for any $\alpha \geq \frac{1}{2}$.*

These demand curves have elasticities θ_1 and θ_2.[19] Example 3 shows that providing information to the monopolist monotonically decreases total surplus exactly when the two demand elasticities are not too far apart. In other words, unless consumers have widely different elasticities with respect to price, price discrimination is bad for total (and consumer) surplus. Detailed calculations are in Online Appendix B.3.

As far as we are aware, none of the existing results in the literature applies to this example. In particular, this example violates the conditions of Aguirre et al. (2010) that give the most general conditions for price discrimination to decrease total surplus. Namely, their condition that the demand curve with a lower monopoly price has a higher curvature is "more concave" than the demand curve with a higher monopoly price, is violated in this example. Our example therefore shows that this ranking of curvatures is not necessary for price discrimination to lower total surplus.

[19]More precisely, these are the elasticities before costs are normalized to zero. In particular, the demand curves in our example are the results of normalizing costs to zero of a setting where the demand curves are $p^{-\theta_i}$ and the marginal cost is $c > 0$. Demand curve $p^{-\theta_i}$ has a constant elasticity of demand θ_i.

4.1 Proof Sketch

It is useful to begin by describing how Proposition 1 is proved. The proof uses the concavification result of Kamenica and Gentzkow (2011). To state the result, let us simply denote by μ a market in which the probability of type θ_2 is μ. Consider the weighted surplus function $W^\alpha : [0,1] \to R$ that specifies the expected α-surplus of a market μ,

$$W^\alpha(\mu) = \mathbb{E}_{\theta\sim\mu}[V^\alpha(p^*(\mu), \theta)],$$

where $p^*(\mu) \in [p^*(\theta_1), p^*(\theta_2)]$ is the profit-maximizing price for the market and is uniquely identified by the first-order condition

$$\mathbb{E}[R_p(p^*(\mu), \theta)] = 0.$$

The concavification result of Kamenica and Gentzkow (2011) implies that information is monotonically bad for α-surplus if and only if W^α is concave. Indeed, if W^α is concave, then splitting any market μ' in the support of a segmentation into multiple markets with the same mean would only (weakly) decrease weighted surplus. And if W^α is not concave, that is, if $W^\alpha(\mu')$ is below the concavification of W^α for some market μ', then we can take any segmentation with market μ' in its support and split μ' into two markets in a way that increases weighted surplus. A similar argument shows that Information is monotonically good for α-surplus if and only if W^α is convex.

To see how the concavity (convexity) of W^α relates to the monotonicity condition in Proposition 1, consider the derivative of W^α with respect to μ,

$$W^\alpha_\mu(\mu) = -V^\alpha(p^*(\mu), \theta_1) + V^\alpha(p^*(\mu), \theta_2) + p^*_\mu(\mu)\mathbb{E}[V^\alpha_p(p^*(\mu), \theta)].$$

The first term is the effect of changing the composition of types, holding the price fixed, and the second term is the effect of changing the price, holding the composition of types fixed. The weighted surplus function W^α is concave if and only if its derivative is decreasing. The expression above contains $p^*(\mu)$ and $p^*_\mu(\mu)$, both of which are defined implicitly given the seller's profit-maximization condition. In the proof of Proposition 1 we eliminate this implicit dependence by observing that because $p^*(\mu)$ is increasing in μ, W^α_μ is a monotone function of μ if and only if $W^\alpha_\mu(\mu(p))$ is a monotone function of p, where $\mu(p)$ is the inverse of the price function, that is, $\mu(p^*(\mu)) = \mu$. Indeed, the expression in Equation (3) is exactly the above expression evaluated at $\mu(p)$.

4.2 Interpretation: The Three Effects of Information

To interpret the monotonicity condition, we identify three effects that providing information has on weighted surplus, and show that Proposition 1 combines these three effects into one formula.

For our purposes here to develop intuition, we continue with keeping p (and its derivatives) implicitly defined. Let us derive the expression for the second derivative of W^α, which is

$$\begin{aligned} W^\alpha_{\mu\mu}(\mu) = (p^*_\mu(\mu))^2\,\mathbb{E}\Big[V^\alpha_{pp}(p^*(\mu),\theta)\Big] \\ +2p^*_\mu(\mu) \quad \Big[V^\alpha_p(p^*(\mu),\theta_2) - V^\alpha_p(p^*(\mu),\theta_1)\Big] \\ +p^*_{\mu\mu}(\mu)\,\mathbb{E}\Big[V^\alpha_p(p^*(\mu),\theta)\Big]. \end{aligned} \tag{4}$$

The above expression highlights that information affects weighted surplus in three ways. To understand these effects, consider what happens if we take a market μ with optimal price p and split it into two markets $\mu_1 = \mu - \delta$ and $\mu_2 = \mu + \delta$ with optimal prices $p_1 < p < p_2$, each with probability a half. Table 2 shows, for each type and each price, the fraction of consumers of that type that face that price, before and after information is provided.[20]

		p_1	p	p_2
before	θ_1	0	1	0
	θ_2	0	1	0
after	θ_1	$\frac{1}{2}(1+\frac{\delta}{1-\mu})$	0	$\frac{1}{2}(1-\frac{\delta}{1-\mu})$
	θ_2	$\frac{1}{2}(1-\frac{\delta}{\mu})$	0	$\frac{1}{2}(1+\frac{\delta}{\mu})$

Table 2: The fraction of consumers of each type facing each price before and after information is provided.

The three effects of information are reflected in Table 2, each represented in one of the terms in Equation (4).

1. The *within-type price change* effect. Information disperses prices for each type: Some consumers face a price drop and some consumers face a price increase. The sign of this effect depends on how V^α is affected by a price dispersion, which depends on the curvature of V^α for each group. If both functions are convex, then this effect is positive, and if they are both concave, the effect is negative. This effect corresponds to the first term in Equation (4).

2. The *cross-types price change* effect. Relatively more type θ_1 consumers face a price drop than do type θ_2 consumers. In fact, more than half of type θ_1 consumers but less than

[20] Consider type θ_1 consumers. After information is provided, those in μ_1 face price p_1 and those in μ_2 face price p_2, and their measures are $\frac{1}{2}(1-\mu+\delta)$ and $\frac{1}{2}(1-\mu-\delta)$. So the fraction of type θ_1 consumers that face price p_1 is $\frac{\frac{1}{2}(1-\mu+\delta)}{\frac{1}{2}(1-\mu+\delta)+\frac{1}{2}(1-\mu-\delta)} = \frac{1}{2}(1+\frac{\delta}{1-\mu})$.

half of type θ_2 consumers face a price drop. In the extreme case of full information, all type θ_1 consumers face a price drop and all type θ_2 consumers face a price increase. The sign of this effect depends on how the marginal effect of changing the price compares across the two types. This effect corresponds to the second term in Equation (4).

3. The *price curvature* effect. The size of price drop $p - p_1$ might not be equal to the size of the price increase $p_2 - p$. The comparison depends on the curvature of the price function $p^*(\mu)$. This effect corresponds to the third term in Equation (4). Notice that the seller's profit-maximization condition means that the expected marginal revenue is zero and therefore

$$\mathbb{E}\Big[V_p^\alpha(p^*(\mu),\theta)\Big] = \alpha\mathbb{E}\Big[CS_p(p^*(\mu),\theta)\Big] \leq 0,$$

and so the sign of this effect only depends on the curvature of the price function. This effect is positive if $p^*(\mu)$ is concave, in which case the price drop is larger than the price increase, benefiting consumers overall. Similarly, this effect is negative if the price function is convex, so the price increase is larger than the price drop.

4.3 Sufficient Conditions

The overall effect of information depends on the *aggregation* of three effects identified above, and each of the three effects might be positive or negative. However, a sufficient condition for information to have a positive (negative) effect is that each of the three effects have the positive (negative) sign. The following corollary formalizes this discussion by directly considering each term in Equation (4).

Corollary 4. *α-IMG (respectively α-IMB) holds if there is partial inclusion and the following hold over the range of prices $(p^*(\theta_1), p^*(\theta_2))$.*

1. *The within-type price change effect is positive (negative): $V^\alpha(p,\theta)$ is convex (concave) for each type $\theta \in \{\theta_1, \theta_2\}$.*

2. *The cross-types price change effect is positive (negative): $V_p^\alpha(p,\theta_2) \geq (\leq) V_p^\alpha(p,\theta_1)$.*

3. *The price curvature effect is positive (negative): $p^*(\mu)$ is concave (convex). A sufficient condition for this is $R_{ppp}(p,\theta) \leq (\geq)\ 0$ and $R_{pp}(p,\theta_2) \leq (\geq) R_{pp}(p,\theta_1)$.*

We next visit IMG and IMB in turn, explaining the sufficient conditions that Corollary 4 gives in each case and relating them to features of demand curves. For these discussions, let us refer to the demand curve with a lower monopoly price as the "more elastic" demand curve. Indeed, over the interval of prices $[p^*(\theta_1), p^*(\theta_2)]$, the elasticity of the demand curve θ_1 is less

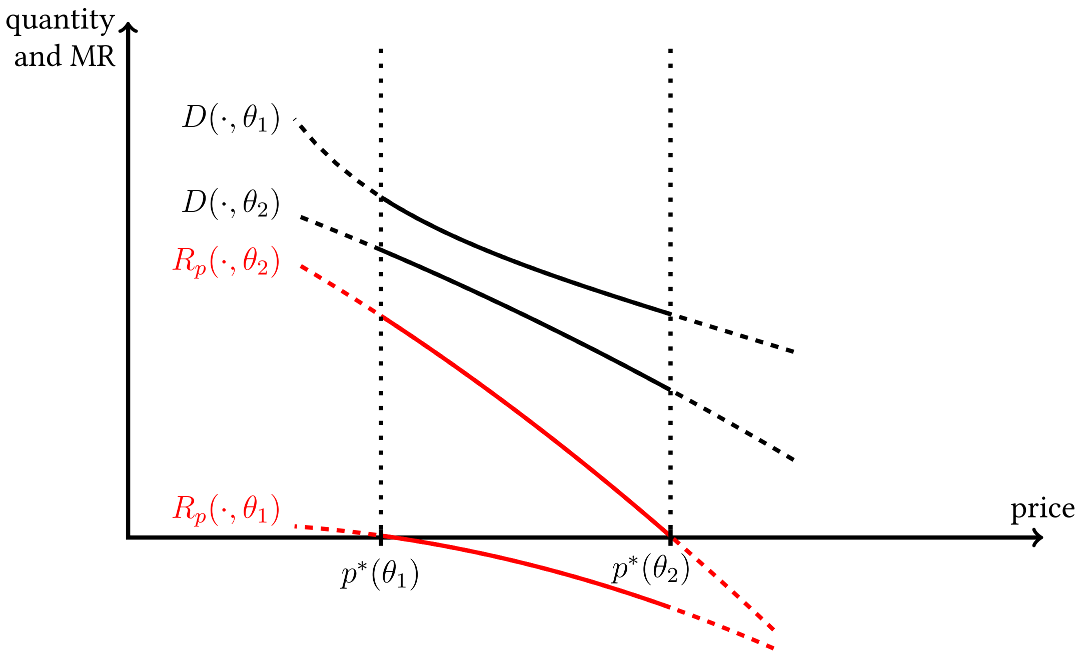

Figure 3: Information is monotonically good when both marginal revenue curves are concave and the more elastic demand θ_1 has a higher level and a less steep marginal revenue curve.

than 1 (because its marginal revenue is negative), but the elasticity of the demand curve θ_1 is more than 1 (because its marginal revenue is positive).[21]

4.3.1 Sufficient Conditions for Monotonically Good Information

Let us start by IMG, summarizing what Corollary 4 implies and then explaining why the implication follows. Corollary 4 implies that assuming $R_{ppp}(p, \theta) \leq 0$, that is, the marginal revenue curves are concave, information is monotonically good for consumer surplus (and therefore any α) if the more elastic demand curve, θ_1, has

1. a higher level, $D(p, \theta_1) \geq D(p, \theta_2)$, and,
2. a less concave revenue curve, $R_{pp}(p, \theta_2) \leq R_{pp}(p, \theta_1)$, or equivalently a less steep marginal revenue curve.

These conditions are shown in Figure 3.

Let us first explain why these conditions imply information is monotonically good and then give and example of they are satisfied. For this, let us discuss each of the three effects of information in turn. The fact that the within-type and cross-group price change effects are positive are straightforward to see. The within-type price change effect is positive be-

[21]If $R_p(p, \theta) \leq 0$ then $D(p, \theta) + pD_p(p, \theta) \geq 0$ so the elasticity is $-\frac{pD_p(p,\theta)}{D(p,\theta)} \leq 1$. A similar argument implies if $R_p(p, \theta) \leq 0$ then $-\frac{pD_p(p,\theta)}{D(p,\theta)} \geq 1$. Additionally, if the elasticities of two demand curves are ranked so that one demand curve is more elastic than the other over the entire range of prices, then the more elastic demand curve has a lower monopoly price.

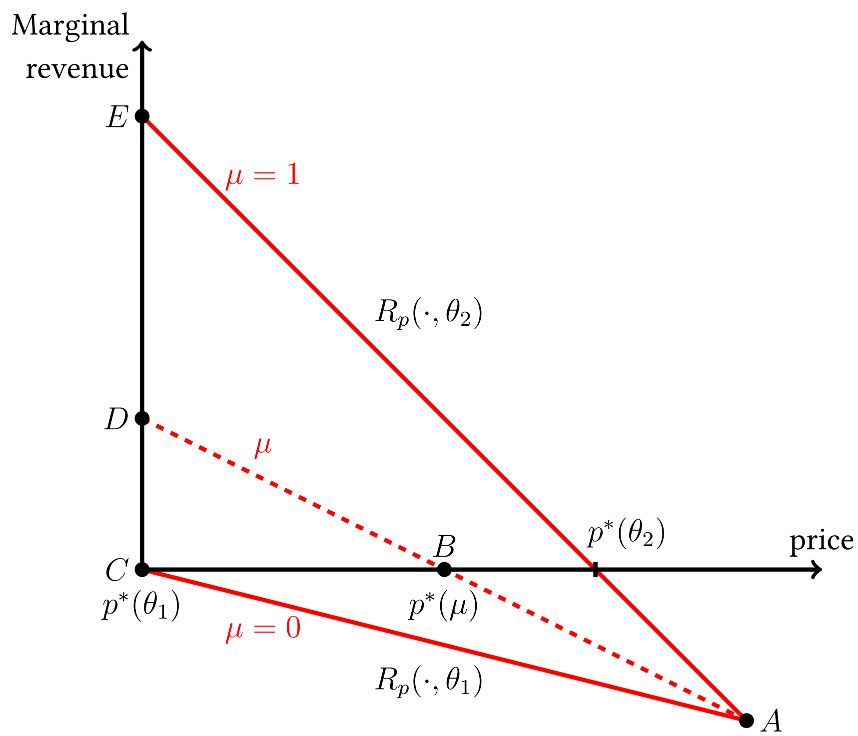

Figure 4: The case where the marginal revenue curve of θ_2 is more steep than that of θ_1, and both curves are linear.

cause $V^1(p, \theta)$ measures consumer surplus which is a convex function of price. The cross-group price change effect is positive because of the ranking of the demand levels, $D(p, \theta_1) \geq D(p, \theta_2)$. The intuition is that if the demand level of type θ_1 consumers is higher than the demand level of type θ_2 consumers, then the benefit to type θ_1 consumers from decreasing the price is larger than the harm to type θ_2 consumers from increasing the price, so the cross-group price change effect is positive.

The price change effect requires a more detailed analysis compared to the other two effects. To understand this effect, let us study each of the two conditions that together imply concavity of the price function, $R_{ppp}(p, \theta) \leq 0$ and $R_{pp}(p, \theta_2) \leq R_{pp}(p, \theta_1)$, in isolation.

To isolate the condition $R_{pp}(p, \theta_2) \leq R_{pp}(p, \theta_1)$, suppose $R_{ppp}(p, \theta) = 0$, that is, marginal revenue curves are linear. This is shown in Figure 4, which zooms in on the relevant range of prices $[p^*(\theta_1), p^*(\theta_2)]$. Because the marginal revenue curve of type θ_2 is higher and is more steep than the marginal revenue curve of type θ_1, they meet at some point A below the horizontal axis. The marginal revenue curve associated with market μ is the line that connects point A to point D which is obtained by taking a convex combination of points E and C with weights μ and $1 - \mu$, where E is the intersection of the marginal revenue curve $R_p(\cdot, \theta_2)$ and the vertical axis, and C is the intersection of the marginal revenue curve $R_p(\cdot, \theta_1)$ and the vertical axis. The intersection of this line AD with the horizontal axis is point B which specifies the optimal price $p^*(\mu)$ for the market μ.

To see that $p^*(\mu)$ is concave, notice that the slope s of the marginal revenue curve asso-

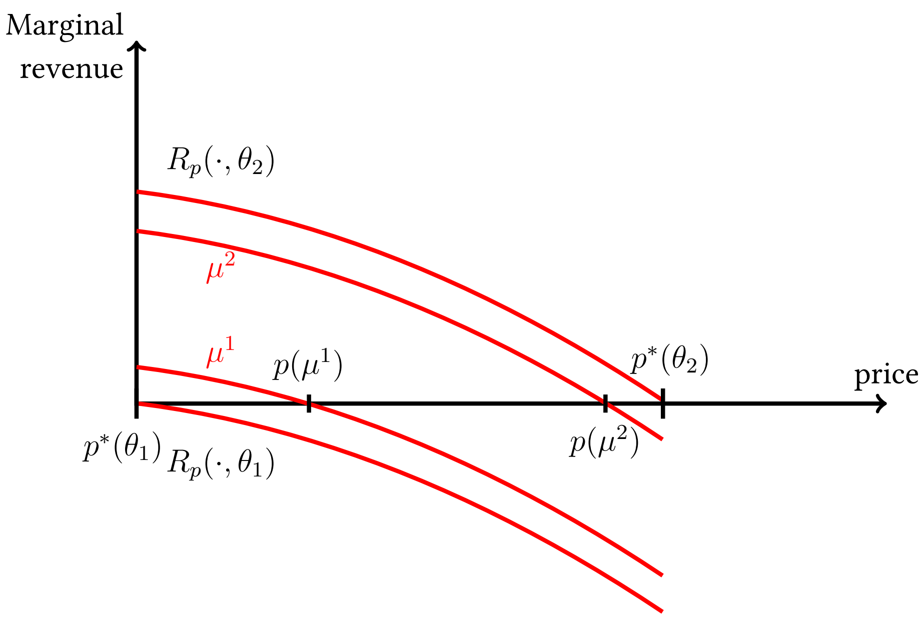

Figure 5: The case where the marginal revenue curves are concave and parallel to each other.

ciated with μ is the negative of the ratio of the length of the line segment CD to the line segment BC,

$$s = -\frac{|CD|}{|BC|}.$$

Let s_1, s_2 be the constant slopes of the two marginal revenue curves $R_p(\cdot, \theta_1)$ and $R_p(\cdot, \theta_2)$, so $s = (1-\mu)s_1 + \mu s_2 = s_1 + \mu(s_2 - s_1)$. We therefore have

$$p^*(\mu) - p^*(\theta_1) = |BC| = -\frac{|CD|}{s} = \frac{\mu|CE|}{-s_1 + \mu(s_1 - s_2)}.$$

Because $|CE|, -s_1$, and $s_1 - s_2$ are all positive constants that do not depend on μ, $p^*(\mu)$ is concave.

To isolate the effect of the condition $R_{ppp}(p, \theta) \leq 0$ on the price curvature, suppose $R_{pp}(p, \theta_2) = R_{pp}(p, \theta_1)$, which implies that the marginal revenue curves are parallel to each other, $R_p(p, \theta_2) = R_p(p, \theta_1) + c$. This is shown in Figure 5. Intuitively, when μ is small, such as at $\mu = \mu^1$ in the figure, the marginal revenue curve intersects the horizontal axis at low prices where the curve is relatively flat, so if we shift the curve upwards, then a large price increase is required to bring us back to the horizontal axis. On the other hand, when μ is large, such as at $\mu = \mu^2$ in the figure, the slope of the marginal revenue curve where it intersects the horizontal axis is large, so a small price increase is sufficient to offset an upwards shift.

We now give an example where the sufficient conditions discussed in this section (shown in Figure 3) are satisfied, implying information is monotonically good.

Example 4 (Sufficient Condition for α-IMG). *Consider two demand curves* $D(p, \theta_i) = a_i -$

$p + \frac{c_i}{p}$ for $i \in \{1, 2\}$ and $a_i, c_i \geq 0$ with supports $[\delta, a_i]$ for small enough $\delta > 0$.[22] *Without loss of generality, assume $a_1 \leq a_2$. Then α-IMG holds for all α if*

$$c_1 - c_2 \geq (a_2 - a_1)\frac{a_2}{2}.$$

In this example, the marginal revenue curves are linear and parallel to each other, and therefore information is monotonically good for total surplus if the level of the demand curve with a lower optimal price is sufficiently higher. It then follows from Corollary 1 that α-IMG holds for all α. Detailed calculations are in Online Appendix B.4.

4.3.2 Sufficient Conditions for Monotonically Bad Information

Having discussed what Corollary 4 says about monotonically good information, let us now discuss what it says about monotonically bad information. We focus on the case where $\alpha \geq \frac{1}{2}$.[23]

Corollary 4 implies that assuming $R_{ppp}(p, \theta) \geq 0$, that is, the marginal revenue curves are convex, and assuming an additional condition that each $V^{1/2}$ is concave, $\frac{1}{2}$-IMB (and therefore α-IMB for any $\alpha \geq \frac{1}{2}$) holds if the more elastic demand curve, θ_1, has

1. a higher derivative, $D_p(p, \theta_1) \geq D_p(p, \theta_2)$, and,
2. a more convex revenue curve, $R_{pp}(p, \theta_2) \geq R_{pp}(p, \theta_1)$, or equivalently a more steep marginal revenue curve.

These conditions, other than the additional concavity condition, are shown in Figure 6. This additional condition does not have a parallel in the case of monotonically good information. Notice that $D_p(p, \theta_1) \leq D_p(p, \theta_2)$ is about how the slopes of the demand curves, and not their levels, rank.

Let us explain why the conditions shown in Figure 6 imply information is monotonically bad for total surplus using from Corollary 4. Instead of reviewing the three effects of information in detail as we did for monotonically good information, we only highlight what the differences are.

An important difference is the additional condition that each $V^{1/2}$ is concave to ensure that the within-type price change effect is negative. Because $V^{1/2}$ measures total surplus, its second derivative is

$$V^1_{pp}(p, \theta) = \frac{1}{2}\frac{d^2}{dp^2}\left(\int_p^{\bar{p}(\theta)} D(x, \theta)\, \mathrm{d}x + R(p, \theta)\right) = \frac{D_p(p, \theta) + pD_{pp}(p, \theta)}{2},$$

[22] A small enough positive δ guarantees that p is bounded away from 0 and each demand is bounded.

[23] The asymmetry between the discussion of this section and the previous section that studied *all* α is because information can never be bad for producer surplus. We can give sufficient conditions for any $\alpha > 0$, but choose to focus on the economically relevant case where $\alpha \geq \frac{1}{2}$.

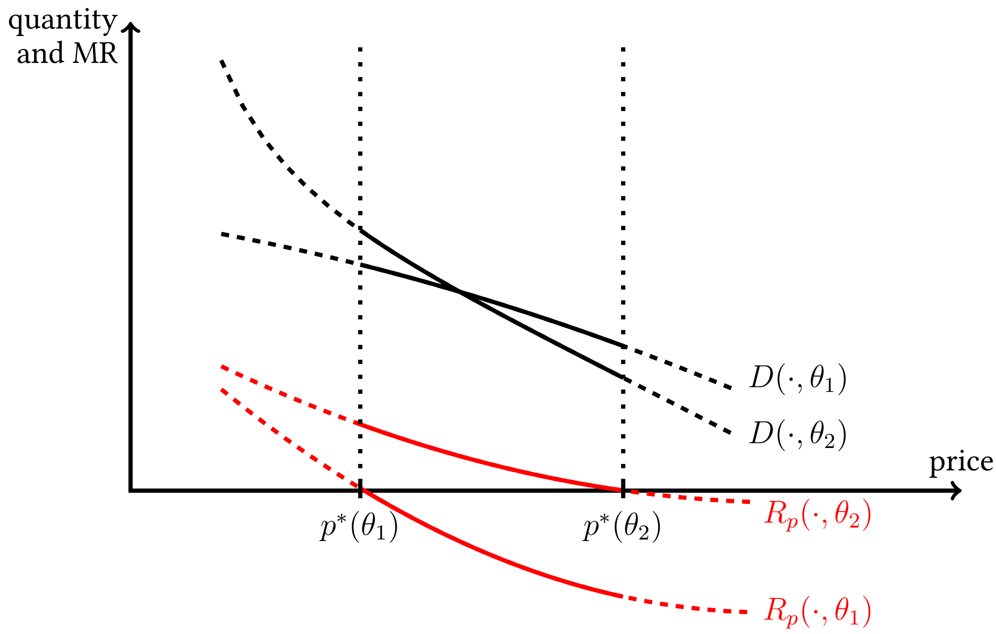

Figure 6: Information is monotonically good for total surplus when over the interval $[p^*(\theta_1), p^*(\theta_2)]$, both marginal revenue curves are convex and the more elastic demand θ_1 has a higher derivative and a more steep marginal revenue curve.

so for the within-type price effect to be negative, we want the above expression to be negative. In words, we want the sum of the consumer surplus, which is convex, and revenue, which is concave, to be concave. We can therefore think about this condition as saying that the revenue curve must be "concave enough" so that even adding the consumer surplus function to it results in a concave function. The marginal revenue curves therefore need to be "decreasing fast enough."

The conditions for the other two effects are natural analogues of the IMG case. Because the objective is now the total surplus, the derivative of the value function is

$$V_p^1(p,\theta) = \frac{d}{dp}\left(\int_p^{\bar{p}(\theta)} D(x,\theta)\,\mathrm{d}x + R(p,\theta)\right) = pD_p(p,\theta),$$

so the condition $V_p^1(p,\theta_2) \leq V_p^1(p,\theta_1)$ becomes $D_p^1(p,\theta_2) \leq D_p^1(p,\theta_1)$. The two conditions $R_{pp}(p,\theta_2) \geq R_{pp}(p,\theta_1)$ and $R_{ppp}(p,\theta) \geq 0$ imply that the price function is convex the same way their negation implies that the price function is concave.

We now give an example where the sufficient conditions discussed in this section (shown in Figure 6) are satisfied, implying information is monotonically good.

Example 4 Continued (Sufficient Condition for α-IMB). *Recall the case of two demand curves $D(p,\theta_i) = a_i - p + \frac{c_i}{p}$ for $i \in \{1,2\}$ where $a_1 \leq a_2$. α-IMB holds for all $\alpha \geq \frac{1}{2}$ if*

$$c_1 \leq c_2 \leq \frac{a_1^2}{4}.$$

In this example, the marginal revenue curves are linear and parallel to each other. Therefore, $\frac{1}{2}$-IMB holds if the the demand curve with a lower optimal price has a higher derivative, $c_1 \leq c_2$, and each $V^{1/2}$ is concave, $c_2 \leq \frac{a_1^2}{4}$. Detailed calculations are in Online Appendix B.4.

5 Reduction from Many Demand Curves to Two

In this section we prove the rest of Theorem 1 that shows how to reduce a general family of demand curves to two demand curves. Combined with the characterization of surplus-monotonicity for two demand curves discussed in Section 4, we obtain a complete characterization of surplus-monotonicity for any number of demand curves.

The reduction has three main steps. We first show how to transform the problem to an optimization problem and establish prior-freeness. We then show why partial inclusion is necessary for surplus-montonicity. Finally, we show how the partial inclusion assumption allows us to replace the seller's profit-maximization with its first-order condition and use strong duality to solve the optimization problem.

5.1 Transformation to an Optimization Problem

Recall that α-IMB for $(\mathcal{D}, \mu_0)$ is a stronger property than saying that the no-information segmentation maximizes V^α for a given prior distribution μ_0. We observe below, however, that α-IMB for $(\mathcal{D}, \mu_0)$ is *equivalent* to saying that the no-information segmentation maximizes V^α *for all* prior distributions over $\mathcal{D}$. This equivalence follows from the standard observation that both of these two properties are equivalent to the concavity of the value function $W^\alpha : \Delta\Theta \to R$ that specifies the expected α-surplus of each market μ,

$$W^\alpha(\mu) = \int_\Theta V^\alpha(p^*(\mu), \theta)\, \mathrm{d}\mu(\theta).$$

Because concavity of W^α is not affected by which prior distribution we are considering, a consequence of this equivalence is that α-IMB and α-IMG are prior-independent properties.

Lemma 1. *The following three statements are equivalent for any $(\mathcal{D}, \mu_0)$.*

1. *α-IMB holds for $(\mathcal{D}, \mu_0)$.*

2. *The value function W^α is concave.*

3. *No-information is "V^α universally optimal", that is, for every market μ, the no-information segmentation that assigns probability 1 to μ solves*

$$\max_{\sigma \in S(\mu)} V^\alpha(\sigma).$$

Furthermore, the following three statements are equivalent for any $(\mathcal{D}, \mu_0)$.

1′. α*-IMG holds for* $(\mathcal{D}, \mu_0)$.

2′. The value function W^α *is convex.*

3′. No-information is "$-V^\alpha$ *universally optimal", that is, for every market* μ*, the no-information segmentation that assigns probability 1 to* μ *solves*

$$\max_{\sigma \in S(\mu)} -V^\alpha(\sigma).$$

Notice that the only difference between statements 3 and 3′ of the lemma is the objectives that are considered in the maximization problem. In statement 3, this objective is V^α and in statement 3′, this objective is $-V^\alpha$. The two statements are identical otherwise, and in particular they both seek the optimality of the no-information segmentation for the corresponding objective. This allows us to give a unified proofs for the two surplus-monotonicity properties by considering a problem

$$\max_{\sigma \in S(\mu)} U(\sigma), \tag{5}$$

for all μ, and then letting $U = V^\alpha$ to characterize α-IMB and letting $U = -V^\alpha$ to characterize α-IMG.[24]

5.2 Necessity of Partial Inclusion

Let us now sketch why partial inclusion is necessary for both of the surplus-monotonicity properties. A preliminary observation is that we can focus on binary demands: If family $\mathcal{D}$ contains two demands $D(\cdot, \theta_1)$ and $D(\cdot, \theta_2)$, once we show that α-IMB (α-IMG) does not hold for the binary family $\mathcal{D}' = \{D(\cdot, \theta_1), D(\cdot, \theta_2)\}$, then we can construct a segmentation and a refinement of it for the family $\mathcal{D}$ that shows that α-IMB (α-IMG) does not hold for $\mathcal{D}$ either. Given this preliminary observation, we sketch the arguments for two demand curves. Furthermore, because surplus-monotonicity properties are prior-free, to show their violation we can freely choose the prior distribution.

Suppose there is full exclusion, that is, $p^*(\theta_2) > \bar{p}(\theta_1)$, as shown in Figure 7 (a). We first argue why α-IMB is violated. The idea is similar to ones in Bergemann et al. (2015) and Pram (2021). If there is full exclusion, there is scope for improving α-surplus by a segmentation that removes some fraction of the excluded consumers and puts them in a new segment in which

[24] One might wonder if statement 3′ can be replaced by a statement that says *full information* maximizes V^α for all priors. Perhaps surprisingly, this statement is not equivalent to α-IMG and is strictly weaker than it. Roughly speaking, for the full information segmentation to maximize V^α for all priors, all that is required is that the value function V^α is pointwise below a hyperplane that connects the points $(\mu, V^\alpha(\mu))$ for markets μ that assign probability 1 to each given state, which is implied by convexity of V^α but does not imply it.

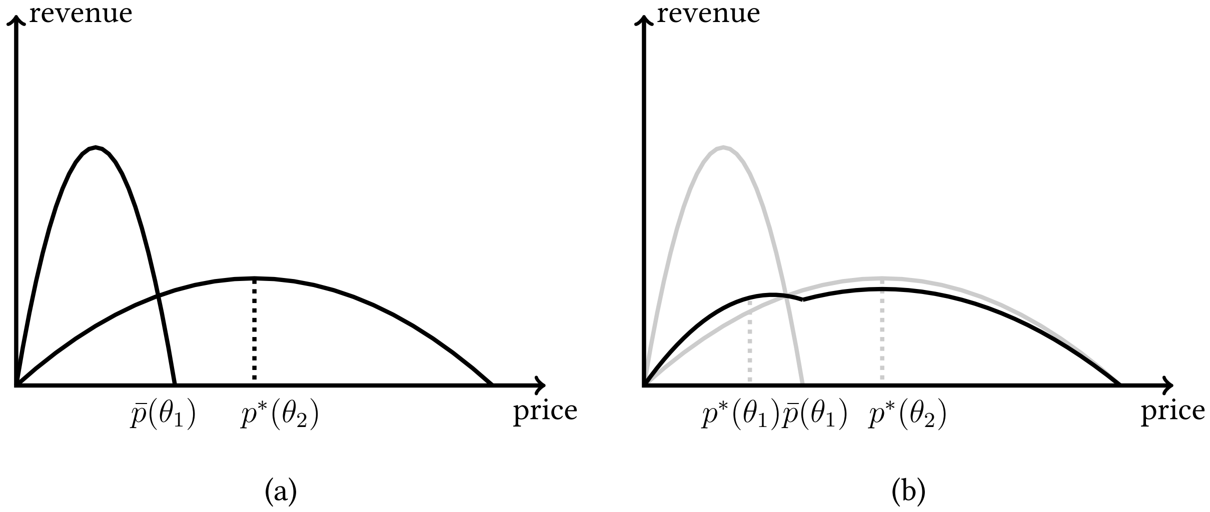

Figure 7: (a) The case of full exclusion, $\bar{p}(\theta_1) < p^*(\theta_2)$. (b) At some threshold $\widehat{\mu}$, there will be two optimal prices.

the price is lower, benefiting those consumers. In particular, suppose the prior probability of type θ_1 is small enough, so that the monopolist optimally excludes type θ_1 and charge price $p^*(\theta_2)$ uniformly to both types. Now consider a segmentation of that completely separates the two types. The first market has only θ_1 in its support and the price is $p^*(\theta_1)$. The second market has only θ_2 in its support and the price is $p^*(\theta_2)$. In the prior market, the monopolist charges a uniform price of $p^*(\theta_2)$, but in the segmented markets the monopolist lowers the price for type θ_1 and keeps the price the same in type θ_2, which leads to an increase in α-surplus for any α.

To prove that α-IMG is violated we use a different construction. In this case, we argue that there is a threshold $\widehat{\mu}$ on the probability of type θ such that at $\widehat{\mu}$ the monopolist is indifferent between price $p^*(\theta_2)$ or a lower price p that is at most $\bar{p}(\theta_1) < p^*(\theta_2)$. This threshold $\widehat{\mu}$ leads to a market with a demand curve shown in Figure 7 (b). Suppose the probability of type θ is just below $\widehat{\mu}$, so the seller charges the lower price. Now consider a segmentation into two segments, one in which the probability of type θ_2 is just above $\widehat{\mu}$, and a second market in which the probability of type θ_2 is just below $\widehat{\mu}$. In the first market, the price jumps discontinuously up to $p^*(\theta_2)$. In the second market, the price goes down. But importantly, the price would go down continuously. As a result, for small enough changes, the negative discontinuous effect of the higher price dominates the positive continuous effect and therefore the segmentation decreases α-surplus.

Let us also briefly sketch the argument if there is full inclusion, that is, $p^*(\theta_1) < \underline{p}(\theta_2)$, as shown in Figure 8. We show that the revenue curve of type θ_2 has a "kink" at price $\underline{p}(\theta_2)$, and as a result, there is an open set (μ_1, μ_2) such that price $\underline{p}(\theta_2)$ is optimal for any market in that interval. This means that we can perturb any such market by a small amount without changing the optimal price. This enables us to construct segmentations in which all prices are

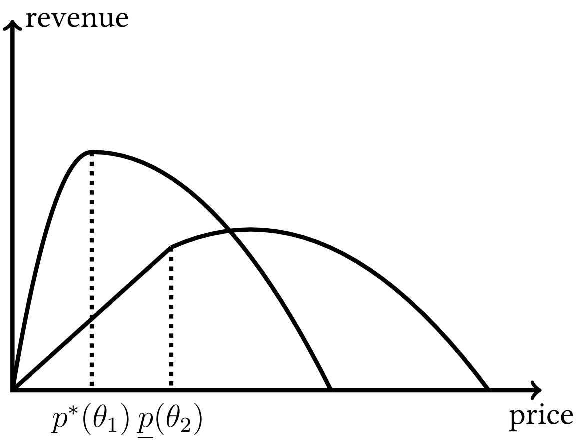

Figure 8: The case of full inclusion, $p^*(\theta_1) < \underline{p}(\theta_2)$.

at most $\underline{p}(\theta_2)$, with some price strictly less than $\underline{p}(\theta_2)$, and to similarly construct segmentations in which all prices are at least $\underline{p}(\theta_2)$, with some price strictly more than $\underline{p}(\theta_2)$.[25]

5.3 Applying Duality

The optimization problem derived in Section 5.1 nests another optimization problem in which the seller chooses an optimal price for each market. We next show how the partial-inclusion assumption allows us to replace the seller's optimal pricing problem with a first-order condition, as formalized below, and then apply duality.

Lemma 2. *Suppose there is partial inclusion. For any market μ, the optimal price $p^*(\mu)$ is unique and solves the following first-order condition*

$$\int_\Theta R_p(p^*(\mu), \theta)\,\mathrm{d}\mu(\theta) = 0.$$

We can therefore write surplus-monotonicity as follows. Surplus-monotonicity holds if and only if for every market μ, the no-information segmentation that assigns probability 1 to μ solves the following problem,

$$\max_{\sigma \in S(\mu), p^*} \int_{\Delta\Theta} \int_\Theta U(p^*(\mu'), \theta)\,\mathrm{d}\mu'(\theta)\,\mathrm{d}\sigma(\mu')$$
$$\int_\Theta R_p(p^*(\mu), \theta)\,\mathrm{d}\mu'(\theta) = 0, \forall \mu' \in \mathrm{Supp}(\sigma),$$

where $U = V^\alpha$ to characterize α-IMB and $U = -V^\alpha$ to characterize α-IMG.

To solve this problem, we apply the duality framework of Kolotilin (2018), Dworczak and Martini (2019), Dworczak and Kolotilin (2024), and Kolotilin et al. (2024). For this, it is convenient to reformulate the problem as one of choosing a joint distribution G over types and

[25] Our construction is related to Haghpanah and Siegel (2023) who also consider markets that can be perturbed without changing the optimal mechanism to construct Pareto improving segmentations.

prices subject to the constraint that the marginal of G over the types agrees with μ and the obedience constraint that conditioned on a given price recommendation, the expected marginal revenue of the seller must be zero.

$$\max_{G \in \Delta(I \times \Theta)} \int_{I \times \Theta} U(p, \theta) \, \mathrm{d}G(p, \theta) \tag{P}$$
$$\text{s.t. } G(I \times \Theta') = \mu(\Theta'), \forall \Theta' \subset \Theta$$
$$\int_{I' \times \Theta} R_p(p, \theta) \, \mathrm{d}G(p, \theta) = 0, \forall I' \subset I$$

Our starting point in solving the problem above is the following complementary slackness conditions that follow from the strong duality results of Kolotilin (2018) and Kolotilin et al. (2024), as we explain in Appendix A.5.

Proposition 2. *A probability measure G is an optimal solution to the problem (P) if and only if there exists continuous functions $\lambda(\theta), \zeta(p)$ such that*

$$\lambda(\theta) + \zeta(p) R_p(p, \theta) = U(p, \theta), G\text{-almost surely} \tag{6}$$
$$\lambda(\theta) + \zeta(p) R_p(p, \theta) \geq U(p, \theta), \forall (p, \theta) \in I \times \Theta \tag{7}$$

We can now use this result to establish the reduction in Theorem 1. We show that for the no-information segmentation to be optimal for every prior, it is necessary and sufficient that the separability condition holds and the two demand curves with the lowest and the highest optimal monopoly prices satisfy surplus-monotonicity.[26]

To use this result, consider a prior distribution μ with optimal price p. The no-information segmentation corresponds to a distribution G that assigns probability 1 to price p (and induces marginal μ over types). So suppose that such a distribution G is optimal for problem (P). This observation pins down the function λ, because given Equation (6) for every θ we must have

$$\lambda(\theta) = U(p, \theta) - \zeta(p) R_p(p, \theta). \tag{8}$$

Now consider an arbitrary price $p' \in I$. Equation (7), applied to the pair (p', θ), implies that

$$\lambda(\theta) + \zeta(p') R_p(p', \theta) \geq U(p', \theta).$$

[26]The observation that the separability condition is necessary for surplus-monotonicity is consistent with the finding in Corollary 1 of Kolotilin et al. (2024) that no-information is suboptimal for generic utility functions that the receiver might have in their Bayesian persuasion setting. A difference is that the value functions of the receiver and the sender in their setting are unrelated whereas in our setting the value function U and the seller's objective R are both pinned down given the demand curves. We use this property to give a complete characterization of our surplus-monotonicity properties.

Substitute the definition of $\lambda(\theta)$ from Equation (8) into the above inequality,

$$U(p,\theta) - \zeta(p)R_p(p,\theta) \geq U(p',\theta) - \zeta(p')R_p(p',\theta),$$

which we can write as

$$p \in \arg\max_{p' \in I} U(p',\theta) - \zeta(p')R_p(p',\theta). \tag{9}$$

We conclude that the no-information segmentation is optimal for a prior μ with optimal price p if and only if there exists a function ζ such that for every type $\theta \in \Theta$, Equation (9) holds.

Recall that surplus-monotonicity holds if the no-information segmentation is optimal for *every* prior μ. Notice that any price in I is optimal for *some* prior μ. Surplus-monotonicity therefore holds if and only if for every p in I, there exists ζ such that Equation (9) holds. This function ζ may depend on p. To make this dependence explicit, let us use $\zeta(p,\cdot)$ as a function that is indexed by p. Using this notation, surplus-monotonicity holds if and only if there exists a single function ζ such that for every $p \in I$ and every type θ,

$$p \in \arg\max_{p' \in I} U(p',\theta) - \zeta(p,p')R_p(p',\theta). \tag{10}$$

The "sufficiency" direction of Theorem 1 is straightforward. Suppose that the family of demand curves $\mathcal{D}$ satisfies the separability condition of the theorem,

$$D(p,\theta) = f_1(\theta)D(p,\theta_1) + f_2(\theta)D(p,\theta_2),$$

for all θ and $p \in I = [p^*(\theta_1), p^*(\theta_2)]$, and that the binary family $\{D(\cdot,\theta)\}_{\theta\in\{\theta_1,\theta_2\}}$ satisfies surplus-monotonicity, which means there exists a function ζ such that for all $p \in I$,

$$p \in \arg\max_{p' \in I} U(p',\theta_1) - \zeta(p,p')R_p(p',\theta_1), \tag{11}$$

$$p \in \arg\max_{p' \in I} U(p',\theta_2) - \zeta(p,p')R_p(p',\theta_2). \tag{12}$$

We want to show that surplus-monotonicity is satisfied for $\mathcal{D}$.

The argument relies on two key observations. First, because each demand is a linear combination of the two base demands, each value, revenue, and marginal revenue function can also be written as the same linear combination of the corresponding objects for the base demand curves,

$$U(p,\theta) = f_1(\theta)U(p,\theta_1) + f_2(\theta)U(p,\theta_2),$$

$$R_p(p,\theta) = f_1(\theta)R_p(p,\theta_1) + f_2(\theta)R_p(p,\theta_2).$$

This step critically uses the property that no matter what α is and which surplus-monotonicity property we are characterizing, U is a linear combination of demand and revenue. Second, because p maximizes each of the two expressions in Equation (11) and Equation (12), it also maximizes their linear combination with positive weights,

$$
\begin{aligned}
p &\in \arg\max_{p' \in I} f_1(\theta)\Big(U(p', \theta_1) - \zeta(p, p') R_p(p', \theta_1)\Big) + f_2(\theta)\Big(U(p', \theta_2) - \zeta(p, p') R_p(p', \theta_2)\Big) \\
&= \arg\max_{p' \in I} U(p', \theta) - \zeta(p, p') R_p(p', \theta)
\end{aligned}
$$

as claimed, completing the proof.

Let us now sketch the proof of the "necessity" part of the Theorem 1. Again, one part of the argument is straightforward. If surplus-monotonicity holds for $\mathcal{D}$, then Equation (10) is satisfied for every type, and in particular it must hold for the two types θ_1, θ_2 that have the lowest and the highest optimal monopoly price in the family, so surplus-monotonicity must hold for the binary family that contains only the demand curves associated with θ_1, θ_2 as well.

To argue that the separability condition of the theorem must be satisfied, we argue first that *each* pair θ, θ' of types in Θ pin down the function ζ over the interval $(p^*(\theta), p^*(\theta'))$ of prices that are in between their two optimal monopoly prices. We then argue that if the separability condition of the theorem is violated, the ζ functions that are pinned down with different pairs do not agree, and therefore a single function ζ does not exist. Appendix A.6 provides detail of the proof.

References

Aguirre, Iñaki and Simon G Cowan, "Monopoly price discrimination with constant elasticity demand," *Economic Theory Bulletin*, 2015, *3*, 329–340. 5

Aguirre, Inaki, Simon Cowan, and John Vickers, "Monopoly price discrimination and demand curvature," *American Economic Review*, 2010, *100* (4), 1601–1615. 3, 5, 12, 13, 17, 62

Asseyer, Andreas, "Information Orders in Screening Problems," *Available at SSRN 4192468*, 2024. 5

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris, "The limits of price discrimination," *American Economic Review*, 2015, *105* (3), 921–957. 3, 5, 7, 12, 14, 15, 16, 27

_ , Francisco Castro, and Gabriel Weintraub, "Third-degree price discrimination versus uniform pricing," *Games and Economic Behavior*, 2022, *131*, 275–291. 5

_ , Tibor Heumann, and Michael C Wang, "A unified approach to second and third degree price discrimination," *arXiv preprint arXiv:2401.12366*, 2024. 5

Blackwell, David, "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 1953, pp. 265–272. 8, 9, 12

Cowan, Simon, "The welfare effects of third-degree price discrimination with nonlinear demand functions," *The RAND Journal of Economics*, 2007, *38* (2), 419–428. 3, 12, 13

_ , "Third-degree price discrimination and consumer surplus," *The Journal of Industrial Economics*, 2012, *60* (2), 333–345. 5

_ , "Welfare-increasing third-degree price discrimination," *The RAND Journal of Economics*, 2016, *47* (2), 326–340. 5

Dworczak, Piotr and Anton Kolotilin, "The persuasion duality," *Theoretical Economics*, 2024, *19* (4), 1701–1755. 6, 29

_ and Giorgio Martini, "The simple economics of optimal persuasion," *Journal of Political Economy*, 2019, *127* (5), 1993–2048. 6, 29

Elliott, Matthew, Andrea Galeotti, Andrew Koh, and Wenhao Li, "Market segmentation through information," *Available at SSRN 3432315*, 2021. 5

Haghpanah, Nima and Ron Siegel, "The limits of multiproduct price discrimination," *American Economic Review: Insights*, 2022, *4* (4), 443–458. 5

_ and _ , "Pareto-improving segmentation of multiproduct markets," *Journal of Political Economy*, 2023, *131* (6), 1546–1575. 5, 29

Holmes, Thomas J, "The effects of third-degree price discrimination in oligopoly," *The American Economic Review*, 1989, *79* (1), 244–250. 5

Immorlica, Nicole, Andre M Sztutman, and Robert M Townsend, "Optimal Credit Scores Under Adverse Selection," in "Proceedings of the 23rd ACM Conference on Economics and Computation" 2022, pp. 737–738. 6

Kamenica, Emir and Matthew Gentzkow, "Bayesian persuasion," *American Economic Review*, 2011, *101* (6), 2590–2615. 18, 37

Kartik, Navin and Weijie Zhong, "Lemonade from lemons: Information design and adverse selection," *arXiv preprint arXiv:2305.02994*, 2023. 5

Kolotilin, Anton, "Optimal information disclosure: A linear programming approach," *Theoretical Economics*, 2018, *13* (2), 607–635. 6, 29, 30

_, Roberto Corrao, and Alexander Wolitzky, "Persuasion and Matching: Optimal Productive Transport," *Journal or Political Ecconomy*, 2024. 6, 7, 16, 29, 30, 42

Pigou, Arthur C., *The economics of welfare*, London: Macmillan, 1920. 2, 3, 12, 14

Pram, Kym, "Disclosure, welfare and adverse selection," *Journal of Economic Theory*, 2021, *197*, 105327. 27

Robinson, Joan, *The economics of imperfect competition*, Springer, 1969. 5

Roesler, Anne-Katrin and Balázs Szentes, "Buyer-optimal learning and monopoly pricing," *American Economic Review*, 2017, *107* (7), 2072–2080. 5

Saeedi, Maryam and Ali Shourideh, "Optimal Rating Design under Moral Hazard," *arXiv preprint arXiv:2008.09529*, 2024. 6

_, Yikang Shen, and Ali Shourideh, "Getting the Agent to Wait," *arXiv preprint arXiv*, 2024, *2407*. 6

Schmalensee, Richard, "Output and welfare implications of monopolistic third-degree price discrimination," *The American Economic Review*, 1981, *71* (1), 242–247. 5

Smolin, Alex and Takuro Yamashita, "Information Design in Smooth Games," *arXiv e-prints*, 2022, pp. arXiv–2202. 6

Varian, Hal R, "Price discrimination and social welfare," *The American Economic Review*, 1985, *75* (4), 870–875. 5

Yang, Kai Hao, "Selling consumer data for profit: Optimal market-segmentation design and its consequences," *American Economic Review*, 2022, *112* (4), 1364–1393. 5

Appendix

A Proofs

A.1 Proof of Corollary 2

Proof of Corollary 2. Let us without loss of generality parameterize the family as $\mathcal{D} = \{a(\theta)(D(p)+b(\theta))\}_\theta$. The first step is to show that $\mathcal{D}$ can be separated as in the theorem.

Let θ_1, θ_2 be the demands in the family with the lowest and highest b, that is, $b(\theta_1) = \underline{b}$, $b(\theta_2) = \bar{b}$) (choosing arbitrarily if there are multiple candidates).

For θ define

$$f_1(\theta) = \frac{a(\theta)(\bar{b} - b(\theta))}{a(\theta_1)(\bar{b} - \underline{b})}$$
$$f_2(\theta) = \frac{a(\theta)(b(\theta) - \underline{b})}{a(\theta_2)(\bar{b} - \underline{b})}.$$

We need to check that for all $\theta = (a, b)$,

$$f_1(\theta)D(p, \theta_1) + f_2(\theta)D(p, \theta_2) = D(p, \theta),$$

That is,

$$f_1(\theta)a(\theta_1) + f_2(\theta)a(\theta_2) = a,$$
$$f_1(\theta)a(\theta_1)b(\theta_1) + f_2(\theta)a(\theta_2)b(\theta_2) = ab.$$

These two equations hold because

$$f_1(\theta)a(\theta_1) + f_2(\theta)a(\theta_2) = \frac{a(\theta)}{\bar{b} - \underline{b}}((\bar{b} - b(\theta)) + (b(\theta) - \underline{b})) = a.$$
$$f_1(\theta)a(\theta_1)b(\theta_1) + f_2(\theta)a(\theta_2)b(\theta_2) = \frac{a(\theta)}{\bar{b} - \underline{b}}((\bar{b} - b(\theta))\underline{b} + (b(\theta) - \underline{b})\bar{b}) = ab.$$

The second step is to characterize the surplus-monotonicity properties for the binary family $\{D(p, \theta_1), D(p, \theta_2)\}$. In order to save on notation, suppose without loss of generality that $a(\theta_2) = b(\theta_2) = 1$, that is, $D(p, \theta_2) = D(p)$, and let $a = a(\theta_1), b = b(\theta_1)$. We want to show that the family $\{a(D(p) + b), D(p)\}$, satisfies α-IMB (α-IMG) if and only if there is partial inclusion and

$$(2\alpha - 1)p + \alpha(\frac{pD'(p)}{R''(p)})$$

is increasing (decreasing) over $[(R')^{-1}(-b), (R')^{-1}(0)]$.

We have

$$\begin{aligned}
D_1 &= a(D_2 + b) \\
R_1 &= aR_2 + apb \\
R_1' &= aR_2' + ab \\
R_1'' &= aR_2'' \\
V_1 &= aV_2 + ab(\alpha\bar{p} + (1-2\alpha)p) \\
V_1' &= aV_2' + ab(1-2\alpha).
\end{aligned}$$

So

$$\begin{aligned}
&V_2(p) - V_1(p) + \frac{-\frac{R_1'(p)}{R_2'(p)}V_2' + V_1'}{-\frac{R_1'(p)}{R_2'(p)}R_2'' + R_1''}(R_1'(p) - R_2'(p)) \\
=&V_2(p) - V_1(p) + \frac{-\frac{aR_2'(p)+ab}{R_2'(p)}V_2' + aV_2' + ab(1-2\alpha)}{-\frac{aR_2'+ab}{R_2'(p)}R_2'' + aR_2''}(R_2'(p)(a-1) + ab) \\
=&V_2(p)(1-a) - ab(\alpha\bar{p} + (1-2\alpha)p) + \frac{V_2'(p) - (1-2\alpha)R_2'(p)}{R_2''(p)}(R_2'(p)(a-1) + ab)
\end{aligned}$$

The derivative of the expression is

$$\begin{aligned}
&V_2'(p)(1-a) - ab(1-2\alpha) + \frac{V_2'(p) - (1-2\alpha)R_2'(p)}{R_2''(p)}R_2''(p)(a-1) \\
&+ (\frac{V_2'(p) - (1-2\alpha)R_2'(p)}{R_2''(p)})'(R_2'(p)(a-1) + ab) \\
=& -(1-2\alpha)R_2'(p)(a-1) - ab(1-2\alpha) + (\frac{V_2'(p) - (1-2\alpha)R_2'(p)}{R_2''(p)})'(R_2'(p)(a-1) + ab) \\
=&(R_2'(p)(a-1) + ab)(-(1-2\alpha) + (\frac{V_2'(p) - (1-2\alpha)R_2'(p)}{R_2''(p)})').
\end{aligned}$$

Notice that $R_2'(p)(a-1) + ab = R_1'(p) - R_2'(p) \leq 0$,

so we want to show that

$$2\alpha - 1 + (\frac{V_2'(p) - (1-2\alpha)R_2'(p)}{R_2''(p)})' \geq 0.$$

Notice also that $V_2'(p) - (1-2\alpha)R_2'(p) = \alpha p D'(p)$. So we want to show that

$$2\alpha - 1 + (\frac{\alpha p D_2'(p)}{R_2''(p)})' \geq 0.$$

In other words we want to show that

$$(2\alpha - 1)p + \alpha(\frac{pD_2'(p)}{R_2''(p)})$$

to be increasing in p over $(p^*(\theta_1), p^*(\theta_2)) = [(R_2^{-1})'(-b), (R_2^{-1})'(0)]$, as claimed.

To see that log-concavity of f is sufficient when $\alpha = 1/2$, note that it is sufficient for

$$\frac{pD'(p)}{R_2''(p)}$$

to be increasing everywhere, or for

$$\frac{R_2''(p)}{pD'(p)} = \frac{2D'(p) + pD''(p)}{pD'(p)} = \frac{2}{p} + \frac{-D''(p)}{-D'(p)} = \frac{2}{p} + (log - D'(p))' = \frac{2}{p} + (\log f(p))'$$

to be decreasing. It is sufficient for $(\log f(p))'$ to be decreasing, that is, f is log-concave. □

A.2 Proof of Proposition 1

Proof. Suppose there is partial inclusion. To save on notation, let $\mu = \mu(\theta_2) \in [0, 1]$ be the probability of θ_2, which means $1 - \mu$ is the probability of θ_1. By the concavification result of Kamenica and Gentzkow (2011), α-IMB (α-IMG) holds if and only if value as a function of μ,

$$W^\alpha(\mu) := (1 - \mu)V^\alpha(p^*(\mu), \theta_1) + \mu V^\alpha(p^*(\mu), \theta_2),$$

is concave where $p^*(\mu)$ is the profit-maximizing price for the seller. As argued in Section 4, the partial inclusion condition implies that this price is uniquely defined and $p^*(\mu) \in [p^*(\theta_1), p^*(\theta_2)]$ is the profit-maximizing price for the seller that solves the seller's first-order condition

$$(1 - \mu)R_p(p^*(\mu), \theta_1) + \mu R_p(p^*(\mu), \theta_2) = 0.$$

The profit-maximizing price is uniquely defined because

$$(1 - \mu)R_p(p, \theta_1) + \mu R_p(p, \theta_2)$$

is decreasing in p and takes a positive value at $p = p^*(\theta_1)$ and a negative value at $p = p^*(\theta_2)$.

We identify three consequences of the first-order condition for future use. First, re-arranging the first-order condition, we have

$$\frac{\mu}{1 - \mu} = \frac{-R_p(p^*(\mu), \theta_1)}{R_p(p^*(\mu), \theta_2)}. \tag{13}$$

Second, taking the derivative of the first-order condition with respect to μ, we have

$$-R_p(p^*(\mu), \theta_1) + R_p(p^*(\mu), \theta_2) + \Big((1-\mu)R_{pp}(p^*(\mu), \theta_1) + \mu R_{pp}(p^*(\mu), \theta_2)\Big)p^*_\mu(\mu) = 0,$$

which, after re-arranging, means

$$p^*_\mu(\mu) = \frac{R_p(p^*(\mu), \theta_1) - R_p(p^*(\mu), \theta_2)}{(1-\mu)R_{pp}(p^*(\mu), \theta_1) + \mu R_{pp}(p^*(\mu), \theta_2)}. \tag{14}$$

Third, $p^*(\mu)$ is increasing in μ. To see this, consider the above equation. Notice that because $p^*(\theta_1) \le p^*(\mu) \le p^*(\theta_2)$ and each revenue curve is concave, the first term in the numerator, $R_p(p^*(\mu), \theta_1)$, is negative and the second term, $R_p(p^*(\mu), \theta_2)$, is positive, so the numerator is negative.[27] Further, because both revenue curves are strictly concave, the denominator is negative. We conclude that $p^*_\mu(\mu) > 0$.

Now $W^\alpha(\mu)$ is concave if and only if its derivative $W^\alpha_\mu(\mu)$ is decreasing. From the definition of $W^\alpha(\mu)$, we can write its derivative $W^\alpha_\mu(\mu)$ as

$$\begin{aligned} W^\alpha_\mu(\mu) =& V^\alpha(p^*(\mu), \theta_2) - V^\alpha(p^*(\mu), \theta_1) + \Big((1-\mu)V^\alpha_p(p^*(\mu), \theta_1) + \mu V^\alpha_p(p^*(\mu), \theta_2)\Big)p^*_\mu(\mu) \\ =& V^\alpha(p^*(\mu), \theta_2) - V^\alpha(p^*(\mu), \theta_1) \\ &+ \Big((1-\mu)V^\alpha_p(p^*(\mu), \theta_1) + \mu V^\alpha_p(p^*(\mu), \theta_2)\Big)\frac{R_p(p^*(\mu), \theta_1) - R_p(p^*(\mu), \theta_2)}{(1-\mu)R_{pp}(p^*(\mu), \theta_1) + \mu R_{pp}(p^*(\mu), \theta_2)} \end{aligned}$$

where the second equation followed from substituting p^*_μ using Equation (14). Re-arranging terms, we have

$$\begin{aligned} W^\alpha_\mu(\mu) =& V^\alpha(p^*(\mu), \theta_2) - V^\alpha(p^*(\mu), \theta_1) \\ &+ \frac{(1-\mu)V^\alpha_p(p^*(\mu), \theta_1) + \mu V^\alpha_p(p^*(\mu), \theta_2)}{(1-\mu)R_p(p^*(\mu), \theta_1) + \mu R_p(p^*(\mu), \theta_2)}\Big(R_p(p^*(\mu), \theta_1) - R_p(p^*(\mu), \theta_2)\Big). \end{aligned}$$

Dividing the numerator and the denominator of the fraction by $1-\mu$, we have

$$\begin{aligned} W^\alpha_\mu(\mu) =& V^\alpha(p^*(\mu), \theta_2) - V^\alpha(p^*(\mu), \theta_1) \\ &+ \frac{V^\alpha_p(p^*(\mu), \theta_1) + \frac{\mu}{1-\mu}V^\alpha_p(p^*(\mu), \theta_2)}{R_{pp}(p^*(\mu), \theta_1) + \frac{\mu}{1-\mu}R_{pp}(p^*(\mu), \theta_2)}\Big(R_p(p^*(\mu), \theta_1) - R_p(p^*(\mu), \theta_2)\Big) \\ =& V^\alpha(p^*(\mu), \theta_2) - V^\alpha(p^*(\mu), \theta_1) \\ &+ \frac{V^\alpha_p(p^*(\mu), \theta_1) + \frac{-R_p(p^*(\mu),\theta_1)}{R_p(p^*(\mu),\theta_2)}V^\alpha_p(p^*(\mu), \theta_2)}{R_p(p^*(\mu), \theta_1) + \frac{-R_p(p^*(\mu),\theta_1)}{R_p(p^*(\mu),\theta_2)}R_p(p^*(\mu), \theta_2)}\Big(R_p(p^*(\mu), \theta_1) - R_p(p^*(\mu), \theta_2)\Big), \end{aligned}$$

[27] To be precise, at the two extremes one term might be zero but then the other term is non-zero. At $p^*(\theta_1)$ the first term is zero but the second term is positive, so the expression is negative. Similarly at $p^*(\theta_2)$ the second term is zero but the first term is negative, so the expression is negative.

where the second equality followed from substituting Equation (13).

The above expression is complex because it contains a term, $p^*(\mu)$, that is implicitly defined via the seller's first-order condition. The observation that $p^*(\mu)$ is increasing in μ is useful because then we know $W^\alpha_\mu(\mu)$ is decreasing if and only if $W^\alpha_\mu(\mu(p))$ is decreasing, where μ is the inverse of p, that is, $p(\mu(p)) = p$. So we want

$$
\begin{aligned}
&V^\alpha(p,\theta_2) - V^\alpha(p,\theta_1) + \frac{V^\alpha_p(p,\theta_1) + \frac{-R_p(p,\theta_1)}{R_p(p,\theta_2)} V^\alpha_p(p,\theta_2)}{R_p(p,\theta_1) + \frac{-R_p(p,\theta_1)}{R_p(p,\theta_2)} R_p(p,\theta_2)} \Big(R_p(p,\theta_1) - R_p(p^*(\mu),\theta_2) \Big) \\
=&V^\alpha(p,\theta_2) - V^\alpha(p,\theta_1) + \frac{-R_p(p,\theta_1)V^\alpha_p(p,\theta_2) + R_p(p,\theta_2)V^\alpha_p(p,\theta_1)}{-R_p(p,\theta_1)R_{pp}(p,\theta_2) + R_p(p,\theta_2)R_{pp}(p,\theta_1)} \Big(R_p(p,\theta_1) - R_p(p,\theta_2) \Big)
\end{aligned}
$$

to be decreasing in p, where the equality followed from multiplying the numerator and denominator of the fraction by $R_p(p,\theta_2)$. □

A.3 Proof of Corollary 4

Proof. Consider each of the terms in Equation (4).

The first term corresponds to the within-type price change effect. It is

$$
(p^*_\mu(\mu))^2 \Big[(1-\mu) V^\alpha_{pp}(p^*(\mu),\theta_1) + \mu V^\alpha_{pp}(p^*(\mu),\theta_2) \Big],
$$

which is positive (negative) whenever

$$
(1-\mu) V^\alpha_{pp}(p^*(\mu),\theta_1) + \mu V^\alpha_{pp}(p^*(\mu),\theta_2) \tag{15}
$$

is positive (negative). Rearranging the first-order condition of the seller's profit-maximization problem, we have

$$
\frac{\mu}{1-\mu} = \frac{-R_p(p^*(\mu),\theta_1)}{R_p(p^*(\mu),\theta_2)}.
$$

Therefore, the expression in Equation (15) is positive (negative) whenever

$$
-R_p(p,\theta_2) V^\alpha_{pp}(p,\theta_1) + R_p(p,\theta_1) V^\alpha_{pp}(p,\theta_2)
$$

is positive (negative), as claimed.

The second term corresponds to the cross-types price change effect. It is

$$
2p^*_\mu(\mu) \Big[V^\alpha_p(p^*(\mu),\theta_2) - V^\alpha_p(p^*(\mu),\theta_1) \Big],
$$

which is positive (negative) whenever

$$V_p^\alpha(p^*(\mu), \theta_2) - V_p^\alpha(p^*(\mu), \theta_1)$$

is positive (negative) because $p^*(\mu)$ is increasing in μ.

The third term corresponds to the price curvature effect. It is

$$p^*_{\mu\mu}(\mu)\Big[(1-\mu)V_p^\alpha(p^*(\mu), \theta_1) + \mu V_p^\alpha(p^*(\mu), \theta_2)\Big].$$

Notice that from the definition of V^α,

$$\begin{aligned}(1-\mu)V_p^\alpha(p^*(\mu), \theta_1) + \mu V_p^\alpha(p^*(\mu), \theta_2) &= \alpha\Big((1-\mu)CS_p(p^*(\mu), \theta_1) + \mu CS_p(p^*(\mu), \theta_2)\Big) \\ &\quad + (1-\alpha)\Big((1-\mu)R_p(p^*(\mu), \theta_1) + \mu R_p(p^*(\mu), \theta_2)\Big) \\ &= \alpha\Big((1-\mu)CS_p(p^*(\mu), \theta_1) + \mu CS_p(p^*(\mu), \theta_2)\Big), \\ &\leq 0,\end{aligned}$$

where the second equality follows because of the first-order condition of the seller's profit-maximization problem, and the inequality follows because for each θ,

$$CS_p(p, \theta) = \frac{d}{dp}\left(\int_p^{\bar{p}(\theta)} D(x, \theta)\,\mathrm{d}x\right) = -D(p, \theta) \leq 0.$$

So the third term is positive (negative) whenever $p^*_{\mu\mu}(\mu)$ is negative (positive).

To complete the proof, we relate the curvature of p to the two sufficient conditions in the third bullet of the corollary. For this, let us take two derivatives of the seller's profit-maximization condition,

$$(1-\mu)R_p(p^*(\mu), \theta_1) + \mu R_p(p^*(\mu), \theta_2) = 0.$$

The first derivative implies

$$-R_p(p^*(\mu), \theta_1) + R_p(p^*(\mu), \theta_2) + p^*_\mu(\mu)\Big[(1-\mu)R_{pp}(p^*(\mu), \theta_1) + \mu R_{pp}(p^*(\mu), \theta_2)\Big] = 0.$$

The second derivative implies

$$
\begin{aligned}
(p^*_\mu(\mu))^2 \quad & \Big[(1-\mu)R_{ppp}(p^*(\mu),\theta_1) + \mu R_{ppp}(p^*(\mu),\theta_2)\Big] \\
+2p^*_\mu(\mu) \quad & \Big[R_{pp}(p^*(\mu),\theta_2) - R_{pp}(p^*(\mu),\theta_1)\Big] \\
+p^*_{\mu\mu}(\mu) \quad & \Big[(1-\mu)R_{pp}(p^*(\mu),\theta_1) + \mu R_{pp}(p^*(\mu),\theta_2)\Big] = 0.
\end{aligned} \tag{16}
$$

Note that $p^*_{\mu\mu}(\mu)$ is multiplied by a negative term because both revenue curves are concave. So $p^*_{\mu\mu}(\mu)$ is positive (negative) if the sum of the other two terms are positive (negative). The first term is positive (negative) whenever

$$(1-\mu)R_{ppp}(p^*(\mu),\theta_1) + \mu R_{ppp}(p^*(\mu),\theta_2)$$

is positive (negative), which is the case if $R_{ppp}(p,\theta)$ is positive (negative) for all p and both θ_1 and θ_2. The second term is positive (negative) whenever

$$R_{pp}(p^*(\mu),\theta_2) - R_{pp}(p^*(\mu),\theta_1)$$

is positive (negative), completing the proof. □

A.4 Proof of Lemma 2

Proof. First notice that the optimal price $p^*(\mu)$ for any market μ must be at least $\min_\theta p^*(\theta)$ and at most $\max_\theta p^*(\theta)$. Otherwise, if the price is less than $\min_\theta p^*(\theta)$, we can increase the price to $\min_\theta p^*(\theta)$ and increase the revenue from every demand in the support of μ because each demand by assumption has a concave revenue curve. A similar argument shows the optimal price is at most $\max_\theta p^*(\theta)$. So we only need to consider prices in the interval $[\min_\theta p^*(\theta), \max_\theta p^*(\theta)]$.

Because of partial inclusion, it must be that $\max_\theta p^*(\theta) \le \bar{p}(\theta')$ and $\min_\theta p^*(\theta) \ge \underline{p}(\theta')$ for every θ'. Because $R(\cdot,\theta')$ is strictly concave and differentiable over the range $[\underline{p}(\theta'), \bar{p}(\theta')]$, it is also concave and differentiable over the smaller range $[\min_\theta p^*(\theta), \max_\theta p^*(\theta)]$. As a result, the revenue function associated with any market μ

$$\int_\Theta R(\cdot,\theta)\,\mathrm{d}\mu(\theta)$$

is also strictly concave and differentiable over the range of prices $[\min_\theta p^*(\theta), \max_\theta p^*(\theta)]$. Because any optimal price for market μ is in the interval $[\min_\theta p^*(\theta), \max_\theta p^*(\theta)]$, a price p is optimal for market μ if and only if it satisfies the first-order condition

$$\int_{\Theta} R_p(p,\theta)\,\mathrm{d}\mu(\theta) = 0.$$

As a result, the optimal price $p^*(\mu) \in [\min_\theta p^*(\theta), \max_\theta p^*(\theta)]$ is unique and is characterized by

$$\int_{\Theta} R_p(p^*(x),\theta)\,\mathrm{d}\mu(\theta) = 0,$$

as claimed. □

A.5 Proof of Proposition 2

The strong duality result of Kolotilin et al. (2024) relates problem (P) to the following dual problem.

$$\begin{aligned} &\min_{\lambda,\zeta} \int_{\Theta} \lambda(\theta)\,\mathrm{d}\mu(\theta) \\ &\text{s.t. } \lambda(\theta) + \zeta(p)R_p(p,\theta) \geq U(p,\theta), \forall (p,\theta) \in I \times \Theta. \end{aligned} \tag{D}$$

Lemma 3 (Kolotilin et al., 2024)**.** *Optimal solutions G and λ, ζ to the primal problem* (P) *and the dual problem* (D) *exist and their values are equal,*

$$\int_{I\times\Theta} U(p,\theta)\,\mathrm{d}G(p,\theta) = \int_{\Theta} \lambda(\theta)\,\mathrm{d}\mu(\theta). \tag{17}$$

We now use this strong duality result to establish Proposition 2. By Lemma 3, a feasible G is an optimal solution the primal problem (P) if and only if there exists a feasible solution λ, ζ to the dual problem (D) such that Equation (17) holds. The feasibility of λ, ζ is exactly Equation (7) from Proposition 2. Furthermore, for any feasible G, we have

$$\int_{I\times\Theta} \zeta(p)R_p(p,\theta)\,\mathrm{d}G(p,\theta) = 0,$$

so

$$\int_{I\times\Theta} U(p,\theta)\,\mathrm{d}G(p,\theta) = \int_{I\times\Theta} [U(p,\theta) - \zeta(p)R_p(p,\theta)]\,\mathrm{d}G(p,\theta),$$

and

$$\int_{\Theta} \lambda(\theta)\,\mathrm{d}\mu(\theta) = \int_{I\times\Theta} \lambda(\theta)\,\mathrm{d}G(p,\theta).$$

So Equation (17) is equivalent to

$$\int_{I\times\Theta}[U(p,\theta)-\zeta(p)R_p(p,\theta)]\,\mathrm{d}G(p,\theta)=\int_{I\times\Theta}\lambda(\theta)\,\mathrm{d}G(p,\theta),$$

which is Equation (6) of Proposition 2.

A.6 Proof of Theorem 1

We here complete the proof of Theorem 1.

A.6.1 Necessity of Partial Inclusion

A simplifying step is that we can focus on two demands. In particular, for a family of demands $\mathcal{D}$, and any two types θ_1,θ_2 in it, if we can construct two segmentations $\sigma_1' \succeq_{\text{MPS}} \sigma_2'$ of the binary family $\mathcal{D}'=\{D(p,\theta_1),D(p,\theta_2)\}$ that shows that α-IMB (α-IMG) does not hold, then α-IMB (α-IMG) does not hold for $\mathcal{D}$ either. For this, we can consider any segmentation of $\mathcal{D}$ that is a convex combination of σ_1' and any other segmentation, i.e., $\sigma_1=\beta\sigma_1'+(1-\beta)\sigma_3$, which exists for small enough β, and then further refine σ_1 to obtain $\sigma_2=\beta\sigma_2'+(1-\beta)\sigma_3$.

Suppose first that there is full exclusion, that is, $\bar{p}(\theta_1)<p^*(\theta_2)$. We want to show that both α-IMB and α-IMG must be violated. It is sufficient to prove this claim for a certain (not every) prior market μ on θ_1,θ_2. This is because for any prior market μ' we can first construct a segmentation that contains market μ in its support, and then show that segmenting μ' further does not increase or decrease weighted surplus (depending on which one of IMG or IMB we are proving).

Consider the set $p^*(\mu)$ of optimal prices in a market in which θ has probability $1-\mu$ and θ_2 has probability μ,

$$p^*(\mu)=\arg\max_p(1-\mu)R(p,\theta_1)+\mu R(p,\theta_2).$$

Notice that $p^*(\mu)$ never includes a price in $(\bar{p}(\theta_1),p^*(\theta_2))$ because any such price, which excludes θ_1 entirely, leads to a lower revenue than $p^*(\theta_2)$. Moreover, $p^*(\mu)$ never includes a price less than $p^*(\theta_1)$ because any such price is also lower than $p^*(\theta_2)$ and increasing it increases revenue in both types. Therefore, $p^*(\mu)$ may only contain prices in $[p^*(\theta_1),\bar{p}(\theta_1)]\cup\{p^*(\theta_2)\}$. At $\mu=0$, the only optimal price is $p^*(\theta_1)$, and at $\mu=1$, the only optimal price is $p^*(\theta_2)$. Because for each p, revenue changes linearly in μ, there exists a threshold $\widehat{\mu}\in(0,1)$ such that $p^*(\theta_2)\in p^*(\mu)$ if and only if $\mu\geq\widehat{\mu}$, and $p^*(\theta_2)$ is the unique optimal price for all $\mu>\widehat{\mu}$.

To see that α-IMB does not hold, consider any segmentation of the prior market that contains a segment in which the probability of θ_2 in the prior market is $\mu'>\widehat{\mu}$ (such a segmentation exists because the prior market has full support over θ_1,θ_2). The unique optimal price for μ' is $p^*(\theta_2)$. Consider segmenting μ' further into two segments $\mu_1<\widehat{\mu}<\mu'$ and $\mu_2>\mu'$.

Because any optimal price for market μ_1 is strictly less than $p^*(\theta_2)$ and the unique optimal price for market μ_2 is $p^*(\theta_2)$, this segmentation increases consumer surplus. Because any segmentation weakly increases producer surplus, it also strictly increases weighted surplus.

To see that α-IMG does not hold, consider any segmentation of the prior market that contains a segment in which the probability of θ_2 in the prior market is $\mu' = \widehat{\mu} - \epsilon$ for some small ϵ. Consider segmenting μ' into two segments $\mu_1 = \mu' - 2\epsilon$ and $\mu_2 = \mu' + 2\epsilon = \widehat{\mu} + \epsilon$, with probability $\frac{1}{2}$ each. As ϵ goes to zero, the seller's revenue approaches the optimal revenue in market μ' continuously by the Maximum theorem. But we argue that the consumer surplus decreases discontinuously, that it, it decreases by at least some $\delta > 0$, and therefore weighted surplus decreases for small enough ϵ.

To see the discontinuity in consumer surplus, consider optimal prices for markets μ_1 and μ_2. Because $\mu_1 < \widehat{\mu}$, optimal prices for μ_1 are those that maximize revenue over $[p^*(\theta_1), \bar{p}(\theta_1)]$. Because the revenue curve is strictly concave over this range, the optimal price and therefore consumer surplus changes continuously. But for market μ_2 the only optimal price is $p^*(\theta_2)$ and the consumer surplus of this market is bounded away from that of μ', so the consumer surplus and total surplus increase for small enough ϵ.

Now suppose there is full inclusion, $p^*(\theta_1) < \underline{p}(\theta_2)$. We argue that there exist μ_1, μ_2 such that $\mu_1 < \mu_2$ and price $\underline{p}(\theta_2)$ is the unique optimal price for any market in (μ_1, μ_2). For this, we first show that note that the revenue curve associated with θ_2 has a kink at price $\underline{p}(\theta_2)$. The revenue of type θ_2 at any price $p < \underline{p}(\theta_2)$ is

$$R(p, \theta_2) = pD(\underline{p}(\theta_2), \theta_2)$$

so the left derivative of $R(p, \theta_2)$ at $p = \underline{p}(\theta)$ is

$$R_p(p, \theta_2) = D(\underline{p}(\theta), \theta_2).$$

The revenue of type θ_2 at any price $p > \underline{p}(\theta_2)$ is

$$R(p, \theta_2) = pD(p, \theta_2)$$

so marginal revenue is

$$R_p(p, \theta_2) = D(p, \theta_2) + pD_p(p, \theta_2).$$

As p converges to $\underline{p}(\theta_2)$ from above, marginal revenue converges to

$$\lim_{p \to^+ \underline{p}(\theta_2)} R_p(p, \theta_2) = D(\underline{p}(\theta_2), \theta_2) + pD_p(\underline{p}(\theta_2), \theta_2) < D(\underline{p}(\theta_2), \theta_2).$$

So the right derivative of the revenue curve at $\underline{p}(\theta_2)$ is strictly less than its left derivative. Let

$\delta^- > \delta^+$ be the left and the right derivatives of the revenue curve of type θ_2 at $\underline{p}(\theta_2)$. Price $\underline{p}(\theta_2)$ is optimal in any market μ such that

$$(1-\mu)R_p(\underline{p}(\theta_2),\theta_1)+\mu\delta^- > 0, \text{ and } (1-\mu)R_p(\underline{p}(\theta_2),\theta_1)+\mu\delta^+ < 0,$$

that is

$$\mu \in (\mu_1,\mu_2) := \left(\frac{-R_p(\underline{p}(\theta_2),\theta_1)}{\delta^+ - R_p(\underline{p}(\theta_2),\theta_1)}, \frac{-R_p(\underline{p}(\theta_2),\theta_1)}{\delta^- - R_p(\underline{p}(\theta_2),\theta_1)}\right).$$

Now consider IMB. consider any segmentation of the prior market that contains a segment in which the probability of θ_2 in the prior market is $\mu' \in (\mu_1,\mu_2)$. Consider segmenting μ' further into two segments μ_1' and μ_2' such that $\mu_1' < \mu_1 < \mu'$ and $\mu' < \mu_2' < \mu_2$. Any optimal price for μ_1' is less than $\underline{p}(\theta_2)$ and the optimal price for μ_2' is $\underline{p}(\theta_2)$. This segmentation therefore increases consumer surplus. Because any segmentation weakly increases producer surplus, it also increases weighted surplus.

Finally, consider IMG. For this, we examine the value function $W^\alpha(\mu)$ around μ_2 and show that the value function is locally concave, which means that providing a small amount of information reduces weighted surplus. For this, consider the left derivative of $W^\alpha(\mu)$ at $\mu=\mu_2$. Recall that price $\underline{p}(\theta_2)$ is optimal for all markets in (μ_1,μ_2), which means that the left derivative of $W^\alpha(\mu)$ at $\mu<\mu_2$ is

$$\begin{aligned} W^\alpha_\mu(\mu) &= V^\alpha(p^*(\mu),\theta_2) - V^\alpha(p^*(\mu),\theta_1) + \Big((1-\mu)V^\alpha_p(p^*(\mu),\theta_1) + \mu V^\alpha_p(p^*(\mu),\theta_2)\Big)p^*_\mu(\mu), \\ &= V^\alpha(p^*(\mu),\theta_2) - V^\alpha(p^*(\mu),\theta_1), \end{aligned}$$

where the equality follows because $p^*_\mu(\mu) = 0$ for μ in (μ_1,μ_2) as $p^*(\mu) = \underline{p}(\theta_2)$. As μ converges to μ_2 from below, this derivative converges to

$$V^\alpha(\underline{p}(\theta_2)),\theta_2) - V^\alpha(\underline{p}(\theta_2),\theta_1).$$

Now consider what happens as μ converges to μ_2 from above. The derivative of W^α is

$$V^\alpha(p^*(\mu),\theta_2) - V^\alpha(p^*(\mu),\theta_1) + \Big((1-\mu)V^\alpha_p(p^*(\mu),\theta_1) + \mu V^\alpha_p(p^*(\mu),\theta_2)\Big)p^*_\mu(\mu).$$

Given the first-order condition of the seller's problem, we can write this derivative as

$$V^\alpha(p^*(\mu),\theta_2) - V^\alpha(p^*(\mu),\theta_1) + \alpha\Big((1-\mu)CS_p(p^*(\mu),\theta_1) + \mu CS_p(p^*(\mu),\theta_2)\Big)p^*_\mu(\mu).$$

which, as μ goes to μ_2 from above, converges to

$$V^{\alpha}(\underline{p}(\theta_2),\theta_2) - V^{\alpha}(\underline{p}(\theta_2),\theta_1) + \alpha\Big((1-\mu)CS_p(\underline{p}(\theta_2),\theta_1) + \mu CS_p(\underline{p}(\theta_2),\theta_2)\Big)p^{*}_{\mu}(\mu_2)$$
$$< V^{\alpha}(\underline{p}(\theta_2),\theta_2) - V^{\alpha}(\underline{p}(\theta_2),\theta_1),$$

where the inequality follows because $p^*(\mu)$ is strictly increasing at $\mu \geq \mu_2$, and $CS_p < 0$. We conclude that W^{α} is not concave, which means that IMG does not hold.

A.6.2 Statement (i)

We start by an implication of the separability property that will be later used in the proof.

Lemma 4. *If the separability condition of Theorem 1 is violated, then there exists some $\theta, p \in I$, and a, b such that $R_p(p,\theta) \neq 0$ and*

$$\begin{aligned} R_p(p,\theta) &= aR_p(p,\theta_1) + bR_p(p,\theta_2) \\ R_{pp}(p,\theta) &= aR_{pp}(p,\theta_1) + bR_{pp}(p,\theta_2) \\ U_p(p,\theta) &\neq aU_p(p,\theta_1) + bU_p(p,\theta_2). \end{aligned}$$

Proof. We will show that there is an open interval of prices $I' \subset (p^*(\theta_1), p^*(\theta_2))$ such that the above three equations hold for all $p \in I'$. Because marginal revenue is decreasing in p, there exists such a price that furthermore satisfies $R_p(p,\theta) \neq 0$.

For each p there exists a pair a, b of constants for which the first two equalities hold, that is, writing a, b as functions of p to make this dependence explicit,

$$R_p(p,\theta) = a(p)R_p(p,\theta_1) + b(p)R_p(p,\theta_2) \tag{18}$$
$$R_{pp}(p,\theta) = a(p)R_{pp}(p,\theta_1) + b(p)R_{pp}(p,\theta_2), \tag{19}$$

which is obtained by solving the system of two equations and two variables,

$$\begin{aligned} a(p) &= \frac{R_{pp}(p,\theta)R_p(p,\theta_2) - R_p(p,\theta)R_{pp}(p,\theta_2)}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)} \\ b(p) &= \frac{-R_{pp}(p,\theta)R_p(p,\theta_1) + R_p(p,\theta)R_{pp}(p,\theta_1)}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)}. \end{aligned}$$

This solution is well-defined because $R_{pp} < 0$ and $R_p(p,\theta_1) < 0 < R_p(p,\theta_2)$.

Similarly, let us define functions c, d as follows

$$\begin{aligned} R(p,\theta) &= c(p)R(p,\theta_1) + d(p)R(p,\theta_2) \\ R_p(p,\theta) &= c(p)R_p(p,\theta_1) + d(p)R_p(p,\theta_2) \end{aligned}$$

Using a similar argument to above, the pair $c(p), d(p)$ is also unique for each p. Violation of the separability condition implies that c, d are not constant functions.

Taking a derivative of the above, we have

$$\begin{aligned} R_p(p,\theta) =& c'(p)R(p,\theta_1) + d'(p)R(p,\theta_2) + c(p)R_p(p,\theta_1) + d(p)R_p(p,\theta_2) \\ =& c(p)R_p(p,\theta_1) + d(p)R_p(p,\theta_2), \end{aligned} \tag{20}$$

which means

$$c'(p)R(p,\theta_1) = -d'(p)R(p,\theta_2),$$

and

$$\begin{aligned} R_{pp}(p,\theta) =& c'(p)R_p(p,\theta_1) + d'(p)R_p(p,\theta_2) + c(p)R_{pp}(p,\theta_1) + d(p)R_{pp}(p,\theta_2) \\ =& (\frac{R_p(p,\theta_1)}{R(p,\theta_1)} - \frac{R_p(p,\theta_2)}{R(p,\theta_2)})c'(p)R(p,\theta_1) + c(p)R_{pp}(p,\theta_1) + d(p)R_{pp}(p,\theta_2) \\ =& \Delta(p) + c(p)R_{pp}(p,\theta_1) + d(p)R_{pp}(p,\theta_2), \end{aligned} \tag{21}$$

where $\Delta(p)$ is defined as

$$\Delta(p) := (\frac{R_p(p,\theta_1)}{R(p,\theta_1)} - \frac{R_p(p,\theta_2)}{R(p,\theta_2)})c'(p)R(p,\theta_1).$$

Let I' be an interval for which $c'(p) \neq 0$, which means that $\Delta(p) \neq 0$ for the same interval.

Combining Equation (18) and Equation (19) with Equation (20) and Equation (21), we have

$$\begin{aligned} a(p) &= \frac{R_{pp}(p,\theta)R_p(p,\theta_2) - R_p(p,\theta)R_{pp}(p,\theta_2)}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)} \\ &= c(p) + \frac{\Delta(p)R_p(p,\theta_2)}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)} \end{aligned} \tag{22}$$

$$\begin{aligned} b(p) &= \frac{-R_{pp}(p,\theta)R_p(p,\theta_1) + R_p(p,\theta)R_{pp}(p,\theta_1)}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)} \\ &= d(p) - \frac{\Delta(p)R_p(p,\theta_1)}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)}. \end{aligned} \tag{23}$$

Now let us evaluate

$$U_p(p,\theta) - a(p)U_p(p,\theta_1) - b(p)U_p(p,\theta_2) \tag{24}$$

using the above two equations and show that it is not equal to zero. From the definition of U, when characterizing α-IMG the above expression is

$$\begin{aligned}U_p(p,\theta)-aU_p(p,\theta_1)-bU_p(p,\theta_2) &= \alpha\Big(CS_p(p,\theta)-a(p)CS_p(p,\theta_1)-b(p)CS_p(p,\theta_2)\Big)\\ &\quad+(1-\alpha)\Big(R_p(p,\theta)-a(p)R_p(p,\theta_1)-b(p)R_p(p,\theta_2)\Big)\\ &= -\alpha\Big(D(p,\theta)-a(p)D(p,\theta_1)-b(p)D(p,\theta_2)\Big)\\ &= \frac{-\alpha}{p}\Big(R(p,\theta)-a(p)R(p,\theta_1)-b(p)R(p,\theta_2)\Big).\end{aligned}$$

When characterizing α-IMB, we want to show that the negative of the above expression is non-zero. In either case, because $\alpha > 0$, to show that Equation (24) is not zero, it is sufficient to show that

$$R(p,\theta)-a(p)R(p,\theta_1)-b(p)R(p,\theta_2)\neq 0.$$

For this, use Equation (22) and Equation (23) to write

$$\begin{aligned}&R(p,\theta)-a(p)R(p,\theta_1)-b(p)R(p,\theta_2)\\ =&R(p,\theta_1)-c(p)R(p,\theta_1)-d(p)R(p,\theta_2)-\Delta(p)\frac{R_p(p,\theta_2)R(p,\theta_1)-R_p(p,\theta_1)R(p,\theta_2)}{R_p(p,\theta_2)R_{pp}(p,\theta_1)-R_p(p,\theta_1)R_{pp}(p,\theta_2)}\\ =&-\Delta(p)\frac{R_p(p,\theta_2)R(p,\theta_1)-R_p(p,\theta_1)R(p,\theta_2)}{R_p(p,\theta_2)R_{pp}(p,\theta_1)-R_p(p,\theta_1)R_{pp}(p,\theta_2)}.\end{aligned}$$

For any $p \in I'$, this expression is non-zero because $\Delta(p) \neq 0$, which proves the claim. □

Given Lemma 4, we now complete the proof of Statement (i) of Theorem 1.

Proof of Statement (i). We have already shown in Appendix A.6.1 that partial inclusion is necessary for the surplus-monotonicity properties. We therefore assume partial inclusion here and establish the necessity and sufficiency of the reduction.

The case where $\min_\theta p^*(\theta) = \max_\theta p^*(\theta)$ is straightforward. In this case, both surplus-monotonicity properties as well as the conditions of the statement hold trivially. So suppose for the rest of the proof that $\min_\theta p^*(\theta) < \max_\theta p^*(\theta)$.

Necessity. As argued in Section 5.3, using Lemma 1 and the strong duality result, Proposition 2, surplus-monotonicity holds if and only if there exists a continuous function ζ such that for every $p \in I$ and $\theta \in \Theta$, we have

$$p \in \arg\max_{p' \in I} U(p',\theta)-\zeta(p,p')R_p(p',\theta), \tag{25}$$

where $W = V^\alpha$ characterizes α-IMB and $W = -V^\alpha$ characterizes α-IMG.

First, notice that if Equation (25) is satisfied for all $\theta \in \Theta$, then it must be satisfied for θ_1, θ_2 that have the lowest and the highest monopoly price in Θ, and therefore surplus-monotonicity holds for the binary family that consists only of θ_1, θ_2.

Now consider any pair θ, θ' of types for which $p^*(\theta) < p^*(\theta')$ and any $p \in (p^*(\theta), p^*(\theta'))$. Equation (25) implies

$$U(p', \theta) - \zeta(p, p')R_p(p', \theta) \leq U(p, \theta) - \zeta(p, p)R_p(p, \theta), \forall p' \in (p^*(\theta), p^*(\theta')).$$

Because $R_p(p', \theta) < 0$, we can divide the above inequality by $R_p(p', \theta)$ and write it as

$$\zeta(p, p') \leq \frac{U(p', \theta) - U(p, \theta) + \zeta(p, p)R_p(p, \theta)}{R_p(p', \theta)}.$$

A similar argument repeated for θ', but with the difference that $R_p(p', \theta') > 0$, implies

$$\zeta(p, p') \geq \frac{U(p', \theta') - U(p, \theta') + \zeta(p, p)R_p(p, \theta')}{R_p(p', \theta')}.$$

Therefore, we must have that for all $p, p' \in (p^*(\theta), p^*(\theta'))$:

$$\frac{U(p', \theta) - U(p, \theta) + \zeta(p, p)R_p(p, \theta)}{R_p(p', \theta)} \geq \frac{U(p', \theta') - U(p, \theta') + \zeta(p, p)R_p(p, \theta')}{R_p(p', \theta')}.$$

Note that evaluated at $p' = p$, both sides of the above are equal to $\zeta(p, p)$. Because both sides are continuously differentiable, it has to be that they are tangent at $p' = p$. Therefore,

$$\frac{U_p(p, \theta)}{R_p(p, \theta)} - \zeta(p, p)\frac{R_{pp}(p, \theta)}{R_p(p, \theta)} = \frac{U_p(p, \theta')}{R_p(p, \theta')} - \zeta(p, p)\frac{R_{pp}(p, \theta')}{R_p(p, \theta')}$$

which pins down $\zeta(p, p)$,

$$\zeta(p, p) = \frac{\frac{U_p(p,\theta)}{R_p(p,\theta)} - \frac{U_p(p,\theta')}{R_p(p,\theta')}}{\frac{R_{pp}(p,\theta)}{R_p(p,\theta)} - \frac{R_{pp}(p,\theta')}{R_p(p,\theta')}} = \frac{U_p(p, \theta)R_p(p, \theta') - U_p(p, \theta')R_p(p, \theta)}{R_{pp}(p, \theta)R_p(p, \theta') - R_{pp}(p, \theta')R_p(p, \theta)}. \tag{26}$$

Now suppose the separability condition of the theorem is violated, so there exists some θ whose demand curve cannot be written as a linear combination of $D(\cdot, \theta_1)$ and $D(\cdot, \theta_2)$ over the interval I. Lemma 4 implies that that there is price $p \in I$ and a, b such that $R_p(p, \theta) \neq 0$ and

$$R_p(p, \theta) = aR_p(p, \theta_1) + bR_p(p, \theta_2)$$
$$R_{pp}(p, \theta) = aR_{pp}(p, \theta_1) + bR_{pp}(p, \theta_2)$$
$$U_p(p, \theta) \neq aU_p(p, \theta_1) + bU_p(p, \theta_2).$$

Because $R_p(p, \theta_1) < 0 < R_p(p, \theta_2)$ and $R_p(p, \theta) \neq 0$, the marginal revenue of θ has a different sign with either θ_1, θ_2. Suppose $0 < R_p(p, \theta)$ (the other case is similar). Then our discussion above pins down $\zeta(p, p)$ in two different ways

$$\zeta(p,p) = \frac{U_p(p,\theta_1)R_p(p,\theta_2) - U_p(p,\theta_2)R_p(p,\theta_1)}{R_{pp}(p,\theta_1)R_p(p,\theta_2) - R_{pp}(p,\theta_2)R_p(p,\theta_1)}$$

and

$$\zeta(p,p) = \frac{U_p(p,\theta_1)R_p(p,\theta) - U_p(p,\theta)R_p(p,\theta_1)}{R_{pp}(p,\theta_1)R_p(p,\theta) - R_{pp}(p,\theta)R_p(p,\theta_1)}.$$

But these two expressions cannot be equal by Lemma 4, implying that ζ satisfying Equation (25) does not exist. To see this, let us use Lemma 4 and the above two equalities to write

$$\begin{aligned}
\zeta(p,p) &= \frac{U_p(p,\theta_1)R_p(p,\theta) - U_p(p,\theta)R_p(p,\theta_1)}{R_{pp}(p,\theta_1)R_p(p,\theta) - R_{pp}(p,\theta)R_p(p,\theta_1)} \\
&\neq \frac{U_p(p,\theta_1)\Big(aR_p(p,\theta_1) + bR_p(p,\theta_2)\Big) - \Big(aU_p(p,\theta_1) + bU_p(p,\theta_2)\Big)R_p(p,\theta_1)}{R_{pp}(p,\theta_1)\Big(aR_p(p,\theta_1) + bR_p(p,\theta_2)\Big) - \Big(aR_{pp}(p,\theta_1) + bR_{pp}(p,\theta_2)\Big)R_p(p,\theta_1)} \\
&= \frac{b\Big(U_p(p,\theta_1)R_p(p,\theta_2) - U_p(p,\theta_2)R_p(p,\theta_1)\Big)}{b\Big(R_{pp}(p,\theta_1)R_p(p,\theta_2) - R_{pp}(p,\theta_2)R_p(p,\theta_1)\Big)} \\
&= \zeta(p,p),
\end{aligned}$$

which is a contradiction.

Sufficiency. We now prove the sufficiency part of the reduction. Suppose a family of demand curves $\mathcal{D}$ can be decomposed into two demands that satisfy α-IMG (α-IMB),

$$D(p,\theta) = f_1(\theta)D(p,\theta_1) + f_2(\theta)D(p,\theta_2), \forall p, \theta.$$

We want to show that $\mathcal{D}$ satisfies α-IMG (α-IMB). That is, there exists a function ζ such that for every price $p \in I$, we have

$$p \in \arg\max_{p' \in I} U(p',\theta) - \zeta(p,p')R_p(p',\theta),$$

for $W = V^\alpha$ ($W = -V^\alpha$).

Notice that because each demand is a linear combination of the two base demands, each value, revenue, and marginal revenue function can also be written using a linear combination of the corresponding objects for the base demand curves. Formally,

$$U(p,\theta) = f_1(\theta)U(p,\theta_1) + f_2(\theta)U(p,\theta_2),$$
$$R(p,\theta) = f_1(\theta)R(p,\theta_1) + f_2(\theta)R(p,\theta_2),$$
$$R_p(p,\theta) = f_1(\theta)R_p(p,\theta_1) + f_2(\theta)R_p(p,\theta_2).$$

Because the family $\{D(p,\theta_1), D(p,\theta_2)\}$ satisfies α-IMG (α-IMB), for each $p \in I$ we have

$$p \in \arg\max_{p' \in I} U(p',\theta_1) - \zeta(p,p')R_p(p',\theta_1),$$
$$p \in \arg\max_{p' \in I} U(p',\theta_2) - \zeta(p,p')R_p(p',\theta_2).$$

Because p maximizes each of the above two expressions, it also maximizes their linear combination,

$$\begin{aligned} p &\in \arg\max_{p' \in I} f_1(\theta)\Big(U(p',\theta_1) - \zeta(p,p')R_p(p',\theta_1)\Big) + f_2(\theta)\Big(U(p',\theta_2) - \zeta(p,p')R_p(p',\theta_2)\Big) \\ &= \arg\max_{p' \in I} U(p',\theta) - \zeta(p,p')R_p(p',\theta) \end{aligned}$$

as claimed, completing the proof.

A.6.3 Statement (ii)

We have already shown in Appendix A.6.1 that partial inclusion is necessary for the surplus-monotonicity properties. Proposition 1 establishes the remainder of the statement assuming partial inclusion.

Online Appendix

B Derivation for Examples

B.1 Example 1

Recall that

$$f(p)=\frac{c_1(c_2+c_3p)^{c_4}}{p^2}.$$

This density function is positive if $c_1\geq 0$ and $c_2+c_3p\geq 0$ for all $p\in[\underline{p},\bar{p}]$.

Let us calculate the expression Equation (2). We have

$$\log p^2 f(p)=\log c_1+c_4\log(c_2+c_3p).$$

Because $D'(p)=-f(p)$ and $R''(p)=2D'(p)+pD''(p)=-2f(p)-pf'(p)$, we have

$$\frac{pD'(p)}{R''(p)}=\frac{pf(p)}{2f(p)+pf'(p)}=(\frac{2}{p}+\frac{f'(p)}{f(p)})^{-1}=(\frac{d}{dp}\log p^2f(p))^{-1}=\frac{c_2+c_3p}{c_3c_4}=\frac{c_2}{c_3c_4}+\frac{p}{c_4}.$$

Therefore, the expression in Equation (2) is

$$(2\alpha-1)p+\alpha(\frac{pD'(p)}{R''(p)})=p(2\alpha-1+\frac{\alpha}{c_4})+\frac{\alpha c_2}{c_3c_4},$$

which is increasing (decreasing) whenever the multiplier of p is positive (negative), that is, when

$$\alpha(2+\frac{1}{c_4})\geq(\leq)\ 1.$$

Now consider three cases for c_4. When $2+\frac{1}{c_4}\leq 1$, that is, when $-1\leq c_4\leq 0$, then $\alpha(2+\frac{1}{c_4})\leq 1$ for all α and therefore information is monotonically good regardless of α. When $2+\frac{1}{c_4}\geq 1$, that is, when $c_4\leq -1$ or $c_4\geq 0$, α-IMB (α-IMG) holds whenever

$$\alpha\geq(\leq)\ \widehat{\alpha}:=\frac{1}{2+\frac{1}{c_4}}=\frac{c_4}{2c_4+1}. \tag{27}$$

This threshold $\widehat{\alpha}\in(0,1]$ is such that information is monotonically bad for all α above the threshold, is monotonically good for all α below the threshold, and has no effect on $\widehat{\alpha}$-surplus. When $c_4\leq -1$, $\widehat{\alpha}$ ranges from $\frac{1}{2}$ to 1. This means that information is monotonically good for total surplus but bad for consumer surplus. When $c_4\geq 0$, $\widehat{\alpha}$ ranges from 0 to $\frac{1}{2}$. This means that information is monotonically good for both total surplus and consumer surplus.

Let us specialize our analysis to the case where $D(p) = d_1 + d_2 p^{d_3}$, where $d_3 > -1$ and d_2 and d_3 have opposite signs (to ensure downward-sloping demand and concave revenue). For this demand, $p^2 f(p)$ is log-concave, and therefore information is monotonically bad for total surplus. We can use our analysis of Example 1 to pin down the threshold $\widehat{\alpha}$. In this case, $c_1 = -d_2 d_3, c_2 = 0, c_3 = 1, c_4 = d_3 + 1$. Parameter c_4 is therefore positive because $d_3 > -1$. Then, using Equation (27), we have

$$\widehat{\alpha} = \frac{d_3 + 1}{2d_3 + 3},$$

which ranges from 0 to $\frac{1}{2}$. The special case of the linear demand is where $d_3 = 1$, and therefore $\widehat{\alpha} = 0.4$. For linear demands, information has no effect on 0.4-surplus, and has a positive effect when $\alpha \leq 0.4$ and a negative effect when $\alpha \geq 0.4$.

Concavity of revenue. We next verify that this example indeed has a concave revenue function. For this, let us calculate the derivative of the density,

$$f'(p) = \frac{c_1 c_3 c_4 (c_2 + c_3 p)^{c_4 - 1}}{p^2} - \frac{2c_1 (c_2 + c_3 p)^{c_4}}{p^3}.$$

Therefore,

$$\begin{aligned} R''(p) = -2f - pf' &= \frac{c_1 (c_2 + c_3 p)^{c_4 - 1}}{p^2} \Big(-2(c_2 + c_3 p) - p c_3 c_4 + 2(c_2 + c_3 p) \Big) \\ &= \frac{-c_1 c_3 c_4 (c_2 + c_3 p)^{c_4 - 1}}{p}, \end{aligned}$$

which is negative because $c_1 \geq 0$, $c_2 + c_3 p \geq 0$ for all $p \in [\underline{p}, \bar{p}]$, and c_3, c_4 have the same signs.

The special case where $D(p) = d_1 + d_2 p^{d_3}$. Notice that the density function is $f(p) = -d_2 d_3 p^{d_3 - 1}$. So for the density function to be positive (the demand to be downward-sloping), we need $d_2 d_3 \leq 0$. The marginal revenue of this demand curve is

$$R''(p) = 2D'(p) + pD''(p) = 2d_2 d_3 p^{d_3 - 1} + d_2 d_3 (d_3 - 1) p^{d_3 - 1} = d_2 d_3 (d_3 + 1) p^{d_3 - 1}$$

which, because $d_2 d_3 \leq 0$, is negative whenever $d_3 + 1 \geq 0$, that is, $d_3 \geq -1$.

B.2 Example 2

We show that 1-IMB holds for two demand curves $D(p, \theta_i) = (\theta_i - p)^c$ for $i \in \{1, 2\}$ and $\theta_i \geq 0, c \in (0, 1]$ with supports $[0, \theta_i]$, if and only if there is partial inclusion. The necessity of partial inclusion follows directly from Theorem 1. So suppose there is partial inclusion. In this example,

$$R(p,\theta_i) = p(\theta_i - p)^c$$
$$R_p(p,\theta_i) = (\theta_i - p)^c - pc(\theta_i - p)^{c-1} = (a_i - p)^{c-1}\big(a_i - p(1+c)\big)$$

so the optimal monopoly price for each demand is $p^*(\theta_i) = \frac{\theta_i}{1+c}$. Assuming without loss of generality that $\theta_1 \geq \theta_2$, the partial inclusion condition is $\frac{\theta_2}{1+c} \leq \theta_1$.

To establish 1-IMB, we want to show that

$$CS(p,\theta_2) - CS(p,\theta_1) + \frac{R_p(p,\theta_2)CS_p(p,\theta_1) - R_p(p,\theta_1)CS_p(p,\theta_2)}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)}(R_p(p,\theta_1) - R_p(p,\theta_2))$$

is decreasing. To simplify notation, let us define

$$h(p) := \frac{R_p(p,\theta_2)CS_p(p,\theta_1) - R_p(p,\theta_1)CS_p(p,\theta_2)}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)}$$

and write the expression as

$$CS(p,\theta_2) - CS(p,\theta_1) + h(p)(R_p(p,\theta_1) - R_p(p,\theta_2)).$$

Because $CS_p(p,\theta) = -D(p,\theta)$, the derivative of this expression, that we want to show is negative, is

$$D(p,\theta_1) - D(p,\theta_2) + h(p)(R_{pp}(p,\theta_1) - R_{pp}(p,\theta_2)) + h_p(p)(R_p(p,\theta_1) - R_p(p,\theta_2)) =$$
$$(D(p,\theta_1) - D(p,\theta_2))(1 + h_p(p)) +$$
$$(D_p(p,\theta_1) - D_p(p,\theta_2))(ph_p(p) + 2h(p)) +$$
$$(D_{pp}(p,\theta_1) - D_{pp}(p,\theta_2))ph(p),$$

where the equality followed from writing R_p and R_{pp} in terms of the demand curves and their derivatives, $R_p(p,\theta) = D(p,\theta) + pD_p(p,\theta)$ and $R_{pp}(p,\theta) = 2D_p(p,\theta) + pD_{pp}(p,\theta)$. The sign of the last term is negative. For this, notice that

$$D_{pp}(p,\theta_1) - D_{pp}(p,\theta_2) = c(c-1)\Big((\theta_1 - p)^{c-2} - (\theta_2 - p)^{c-2}\Big) \leq 0$$

because $\theta_1 \leq \theta_2$ and $c - 1 \leq 0$ (and therefore $c - 2 \leq 0$), and $h(p) \geq 0$ given that $R_p(p,\theta_1) \leq 0 \leq R_p(p,\theta_2)$. Therefore, it is sufficient to show that

$$(D(p,\theta_1) - D(p,\theta_2))(1 + h_p(p)) + (D_p(p,\theta_1) - D_p(p,\theta_2))(ph_p(p) + 2h(p)) \tag{28}$$

is negative. To do so, we use the following two claims which we prove in the following two subsections.

Claim 1. $2h(p) - p \le 0$.

Claim 2. $1 + h_p(p) \ge 0$.

Using Claim 1 and Claim 2, we show that the expression in Equation (28) is negative. For this, first notice that because $R_p(p, \theta_1) \le 0 \le R_p(p, \theta_2)$, we have

$$D(p, \theta_1) \le -pD_p(p, \theta_1)$$
$$D(p, \theta_2) \ge -pD_p(p, \theta_2)$$

and therefore

$$D(p, \theta_1) - D(p, \theta_2) \le -p(D_p(p, \theta_1) - D_p(p, \theta_2)).$$

So because $1 + h_p(p) \ge 0$ by Claim 2, we have

$$(D(p, \theta_1) - D(p, \theta_2))(1 + h_p(p)) \le -p(D_p(p, \theta_1) - D_p(p, \theta_2))(1 + h_p(p)).$$

We can therefore give an upper bound on the expression in Equation (28) as follows

$$\begin{aligned} &(D(p, \theta_1) - D(p, \theta_2))(1 + h_p(p)) + (D_p(p, \theta_1) - D_p(p, \theta_2))(ph_p(p) + 2h(p)) \\ \le &(D_p(p, \theta_1) - D_p(p, \theta_2))(-p(1 + h_p(p)) + ph_p(p) + 2h(p)) \\ = &(D_p(p, \theta_1) - D_p(p, \theta_2))(2h(p) - p) \le 0, \end{aligned}$$

where the last inequality follows from Claim 1 and because

$$D_p(p, \theta_1) - D_p(p, \theta_2).$$

So to complete the proof, we only need to establish Claim 1 and Claim 2.

B.2.1 Proof of Claim 1

Let us examine the numerator and the denominator of $h(p)$ in turn. The numerator is

$$\begin{aligned} &R_p(p, \theta_2)CS_p(p, \theta_1) - R_p(p, \theta_1)CS_p(p, \theta_2) \\ = &-(D(p, \theta_2) + pD_p(p, \theta_2))D(p, \theta_1) + (D(p, \theta_1) + pD_p(p, \theta_1))D(p, \theta_2) \\ = &p(D(p, \theta_2)D_p(p, \theta_1) - D(p, \theta_1)D_p(p, \theta_2)). \end{aligned}$$

The denominator is

$$
\begin{aligned}
&R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)\\
=&(D(p,\theta_2) + pD_p(p,\theta_2))(2D_p(p,\theta_1) + pD_{pp}(p,\theta_1))\\
-&(D(p,\theta_1) + pD_p(p,\theta_1))(2D_p(p,\theta_2) + pD_{pp}(p,\theta_2))\\
=&2(D_p(p,\theta_1)D(p,\theta_2) - D_p(p,\theta_2)D(p,\theta_1))\\
+&p(D_{pp}(p,\theta_1)D(p,\theta_2) - D_{pp}(p,\theta_2)D(p,\theta_1))\\
+&p^2(D_{pp}(p,\theta_1)D_p(p,\theta_2) - D_{pp}(p,\theta_2)D_p(p,\theta_1)).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&2h(p) - p =\\
&+\frac{1}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)}\times\\
&\bigg(2p(D(p,\theta_2)D_p(p,\theta_1) - D(p,\theta_1)D_p(p,\theta_2))\\
&- 2p(D_p(p,\theta_1)D(p,\theta_2) - D_p(p,\theta_2)D(p,\theta_1))\\
&- p^2(D_{pp}(p,\theta_1)D(p,\theta_2) - D_{pp}(p,\theta_2)D(p,\theta_1))\\
&- p^3(D_{pp}(p,\theta_1)D_p(p,\theta_2) - D_{pp}(p,\theta_2)D_p(p,\theta_1))\bigg)\\
&= \frac{p^2}{R_p(p,\theta_2)R_{pp}(p,\theta_1) - R_p(p,\theta_1)R_{pp}(p,\theta_2)}\times\\
&\bigg(- (D_{pp}(p,\theta_1)D(p,\theta_2) - D_{pp}(p,\theta_2)D(p,\theta_1))\\
&- p(D_{pp}(p,\theta_1)D_p(p,\theta_2) - D_{pp}(p,\theta_2)D_p(p,\theta_1))\bigg).
\end{aligned}
$$

Because the denominator of h is negative, it is sufficient to show that

$$(D_{pp}(p,\theta_1)D(p,\theta_2) - D_{pp}(p,\theta_2)D(p,\theta_1)) + p(D_{pp}(p,\theta_1)D_p(p,\theta_2) - D_{pp}(p,\theta_2)D_p(p,\theta_1))$$

is negative. Substituting the definition of the demands, this expression is

$$c(c-1)(\theta_1 - p)^{c-2}(\theta_2 - p)^{c-2}(\theta_2 - \theta_1)(\theta_1 + \theta_2 - p(2 + c)).$$

It is therefore sufficient to show that $\theta_1 + \theta_2 - p(2+c) \geq 0$. For this, notice that because $p \leq \frac{\theta_2}{1+c}$ and $\theta_1 \geq \frac{\theta_2}{1+c}$, we have

$$\begin{aligned}\theta_1 + \theta_2 - p(2+c) &\geq \theta_1 + \theta_2 - \frac{\theta_2(2+c)}{1+c} \\ &\geq \frac{\theta_2(2+c)}{1+c} - \frac{\theta_2(2+c)}{1+c} = 0,\end{aligned}$$

completing the proof.

B.2.2 Proof of Claim 2

Using the definitions of the demand curves, we have

$$h(p) = \frac{p(\theta_1 - p)(\theta_2 - p)}{p^2 c(1+c) - p(1+c)(\theta_1+\theta_2) + 2\theta_1\theta_2}.$$

Let us write $h(p) = f(p)/g(p)$ where

$$\begin{aligned}f(p) &= p(\theta_1 - p)(\theta_2 - p) \\ g(p) &= p^2 c(1+c) - p(1+c)(\theta_1+\theta_2) + 2\theta_1\theta_2.\end{aligned}$$

Therefore,

$$1 + h'(p) = 1 + \frac{f'(p)g(p) - g(p)f'(p)}{g^2(p)} = \frac{g^2(p) + f'(p)g(p) - g(p)f'(p)}{g^2(p)}.$$

So in order to prove that $1 + h'(p) \geq 0$, we need to prove that

$$g^2(p) + f'(p)g(p) - g(p)f'(p) \geq 0.$$

We do so in three steps.

Step 1. We show that $g^2(p) + f'(p)g(p) - g(p)f'(p)$ is decreasing in p. So it is sufficient to establish the inequality at the highest possible price, $p = \frac{\theta_2}{1+c}$.

Step 2. We show that $g^2(p) + f'(p)g(p) - g(p)f'(p)$ is increasing in θ_1. So it is sufficient to establish the inequality at the lowest possible θ_1, $\theta_1 = \frac{\theta_2}{1+c}$.

Step 3. We show that at $p = \frac{\theta_2}{1+c}$, $\theta_1 = \frac{\theta_2}{1+c}$, the expression is zero.

Let us now explain each step in detail.

Step 1. The derivative of $g^2(p) + f'(p)g(p) - g(p)f'(p)$ with respect to p is

$$\begin{aligned}
&2(p^2c(1+c) - p(1+c)(\theta_1+\theta_2) + 2\theta_1\theta_2)(2pc(1+c) - (1+c)(\theta_1+\theta_2)) \\
&+ (6p - 2(\theta_1+\theta_2))(p^2c(1+c) - p(1+c)(\theta_1+\theta_2) + 2\theta_1\theta_2) \\
&- 2c(1+c)p(\theta_1-p)(\theta_2-p).
\end{aligned}$$

The last term is positive because $p \le \theta_1, \theta_2$. So it is sufficient to show that the first two terms are negative, that is

$$\begin{aligned}
&(p^2c(1+c) - p(1+c)(\theta_1+\theta_2) + 2\theta_1\theta_2)\times \\
&(2pc(1+c) - (1+c)(\theta_1+\theta_2) + 6p - 2(\theta_1+\theta_2))
\end{aligned}$$

is negative. The first term is positive because it is $g(p) \ge 0$. So we want to show that the second term is negative. For this, given $p \le \frac{\theta_2}{1+c}$ and $\theta_1 \ge \frac{\theta_2}{1+c}$, let us write

$$\begin{aligned}
&(2pc(1+c) - (1+c)(\theta_1+\theta_2) + 6p - 2(\theta_1+\theta_2)) \\
=&p(4c(1+c)+6) - (\theta_1+\theta_2)(4+2c) \\
\le&\frac{\theta_2}{1+c}(4c(1+c)+6) - (\frac{\theta_2}{1+c} + \theta_2)(4+2c) \\
=&\frac{2\theta_2(c-1)^2}{1+c} \le 0.
\end{aligned}$$

Step 2. At $p = \frac{\theta_2}{1+c}$, the derivative of $g^2(p) + f'(p)g(p) - g(p)f'(p)$ with respect to θ_1 is

$$2\theta_1(\theta_2^2\frac{3c}{1+c}) + \theta_2^3(\frac{-5c-c^2}{(1+c)^2}).$$

Because this expression is decreasing in θ_1, to show that it is positive, it is sufficient to show that it is positive at the lowest possible θ_1, $\theta_1 = \frac{\theta_2}{1+c}$, at which the expression is

$$\frac{\theta_2^3(c-c^2)}{(1+c)^2} \ge 0.$$

Step 3. At $p = \frac{\theta_2}{1+c}, \theta_1 = \frac{\theta_2}{1+c}$, both g and f are zero. To see this, consider f,

$$f(p) = p(\theta_1 - p)(\theta_2 - p),$$

which is zero because $\theta_1 = p$. Now consider g,

$$\begin{aligned}
g(p) &= p^2c(1+c) - p(1+c)(\theta_1+\theta_2) + 2\theta_1\theta_2 \\
&= \frac{\theta_2^2c}{1+c} - \frac{\theta_2^2(2+c)}{1+c} + \frac{2\theta_2^2}{1+c} = 0.
\end{aligned}$$

Therefore,

$$g^2(p) + f'(p)g(p) - g(p)f'(p) = 0.$$

B.3 Example 3

For a CES demand curve $D(p, \theta) = (c + p)^{-\theta}$, the optimal monopoly price is given by

$$p^*(\theta) = \frac{c}{\theta - 1}.$$

Thus, we require $c > 0$ and $\theta > 1$ to ensure that a finite monopoly price exists when the monopolist faces each demand curve. To make the exposition of the proof more intuitive, we define $\theta_1 = \theta_H$ and $\theta_2 = \theta_L$, which means $\theta_H > \theta_L$, and use (θ_H, θ_L) instead of (θ_1, θ_2). Thus, the statement of the example is

$$|\theta_H - \theta_L| < \frac{1}{2} \Rightarrow \alpha - \text{IMB holds for } \alpha \geq \frac{1}{2}$$

Let $x = \theta_H - \theta_L$, $0 < x \leq \frac{1}{2}$. The derivative of the expression in Equation (3) simplifies to

$$-\frac{(c+p)^{-\theta_L - x - 1}\left((c - \theta_L p + p)(c+p)^x - c + \theta_L p + px - p\right)}{2\left(2c^2 - cp(2\theta_L + x - 3) + (\theta_L - 1)p^2(\theta_L + x - 1)\right)^2} \times$$
$$\Big(2c^4 + 8c^3 p - c^2 p^2(\theta_L(3\theta_L + 2) + 3\theta_L x + x - 11) + cp^3(2\theta_L + x - 3)(\theta_L(\theta_L + x) - 2)$$
$$- (\theta_L - 1)p^4(\theta_L + x - 1)(\theta_L(\theta_L + x) - 1)\Big)$$

The denominator is positive. Thus, in order to show that the derivative is negative for all p, we need to show that the numerator is positive.

1. $(c + p)^{-(\theta_L + x + 1)} > 0$: always true.

2. $(c - \theta_L p + p)(c + p)^x - c + \theta_L p + px - p > 0$.

   Simplify the above expression to get

$$\begin{aligned} &(c - (\theta_L - 1)p)(c+p)^x - (c - (\theta_L - 1)p) + px \\ &= ((c+p)^x - 1)\,(c - (\theta_L - 1)p) + px \end{aligned}$$

   As $p^*(\theta_H) < p < p^*(\theta_L)$ and $p^*(\theta_L) = \frac{c}{\theta_L - 1}$, $c - (\theta_L - 1)p > 0$.

   If $((c + p)^x - 1) > 0$, we are done. So suppose $((c + p)^x - 1) < 0$. In this case, because

$c-(\theta_L-1)p$ is maximized at $p=p^*(\theta_H)=p^*(\theta_L+x)=\frac{c}{\theta_L+x-1}$, we have

$$\begin{aligned}
&\left((c+p)^x-1\right)\left(c-(\theta_L-1)p\right)+px \\
\geq &\left((c+p)^x-1\right)\left(c-(\theta_L-1)p^*(\theta_H)\right)+px \\
= &\frac{cx\left((c+p)^x-1\right)}{\theta_L+x-1}+px. \qquad (29)
\end{aligned}$$

The derivative of this expression with respect to p is given by

$$\frac{x^2c(c+p)^{x-1}}{\theta_L+x-1}+x,$$

which is positive. Thus, Equation 29 is minimized when p takes its minimum value at $p=p^*(\theta_H)=p^*(\theta_L+x)$. In this case, Equation 29 simplifies to

$$\frac{cx\left(c\left(\frac{1}{\theta_L+x-1}+1\right)\right)^x}{\theta_L+x-1}>0.$$

Thus, the minimum value of $\left((c+p)^x-1\right)\left(c-(\theta_L-1)p\right)+px$ is greater than zero, which completes the proof.

3.

$$\begin{aligned}
&\Big(2c^4+8c^3p-c^2p^2(\theta_L(3\theta_L+2)+3\theta_Lx+x-11)+cp^3(2\theta_L+x-3)(\theta_L(\theta_L+x)-2) \\
&-(\theta_L-1)p^4(\theta_L+x-1)(\theta_L(\theta_L+x)-1)\Big)>0 \qquad (30)
\end{aligned}$$

We show that the derivative of Equation 30 with respect to x is negative, which means that it takes its lowest value at the highest value of x, $x=\frac{1}{2}$. The derivative is given by

$$p^2\left(\theta_L^2p(3(c+p)-2px)+\theta_L(c+p)(2px-3c)-(c+p)^2-2\theta_L^3p^2\right).$$

Dividing this expression by p^2 does not change its sign. So we want to show that the following expression is negative

$$\theta_L^2p(3(c+p)-2px)+\theta_L(c+p)(2px-3c)-(c+p)^2-2\theta_L^3p^2. \qquad (31)$$

Let us first evaluate Equation 31 at the minimum and maximum p i.e. $p^*(\theta_L+x)$ and $p^*(\theta_L)$. At $p^*(\theta_L+x)$, the expression is $-\frac{c^2(\theta_L+x)(2(\theta_L-1)\theta_L+(\theta_L+1)x)}{(\theta_L+x-1)^2}$ while at $p^*(\theta_L)$, it is given by $-\frac{2c^2\theta_L^2}{\theta_L-1}$. Thus, the two end points are always negative.

As such, in order for Equation 31 to be negative for all $p^*(\theta_L+x)\leq p\leq p^*(\theta_L)$ it is sufficient to show that it is convex in p. The second derivative of Equation 31 with respect to p is

$$-2(\theta_L-1)(\theta_L(2\theta_L+2x-1)-1)=-2(\theta_L-1)\left(2\theta_L(\theta_L+x-1)(\theta_L-1)\right)>0.$$

Therefore, Equation 30 is minimized for the highest value of x which is $x=\frac{1}{2}$.

Evaluating Equation 30 at $x=\frac{1}{2}$ gives

$$\frac{1}{4}(c-(\theta_L-1)p)\left(8c^3+8c^2(\theta_L+3)p+2c\left(-2\theta_L^2+\theta_L+9\right)p^2+\left(4\theta_L^3-5\theta_L+2\right)p^3\right)$$

To show that the above expression is positive, it is sufficient to show that each of its two constituents are positive

(a) $(c-(\theta_L-1)p)>0$: as we are only interested in $p^*(\theta_L+\frac{1}{2})<p<p^*(\theta_L)$ and $p^*(\theta_L)=\frac{c}{\theta_L-1}$, this is true.

(b) $(8c^3+8c^2(\theta_L+3)p+2c\left(-2\theta_L^2+\theta_L+9\right)p^2+\left(4\theta_L^3-5\theta_L+2\right)p^3)>0$

We will show that this expression is positive at minimum value of p, $p=p^*(\theta_L+\frac{1}{2})$, and that it is increasing in p.

First, the value of this expression at $p=p^*(\theta_H)=p^*(\theta_L+\frac{1}{2})=\frac{2c}{2\theta_L-1}$ is given by

$$\frac{16c^3(2\theta_L+1)^2}{(2\theta_L-1)^2}>0$$

which is positive.

Next, take the derivative with respect to p and simplify to get

$$\begin{aligned}&8c^2(\theta_L+3)+4c\left(-2\theta_L^2+\theta_L+9\right)p+3\left(4\theta_L^3-5\theta_L+2\right)p^2\\=&8c^2(\theta_L+3)+36cp-4cp\theta_L(2\theta_L-1)+3p^2(2\theta_L-1)(2\theta_L^2+\theta_L-2)\\=&8c^2(\theta_L+3)+36cp+(2\theta_L-1)(3p^2\theta_L(2\theta_L+1)-4c\theta_Lp-6p^2).\end{aligned}$$

The first two terms are positive. So in order to show that the derivative is positive, we show

$$3(2\theta_L+1)\theta_Lp^2-4c\theta_Lp-6p^2>0.$$

Dividing by p

$$\begin{aligned}&3(2\theta_L+1)\theta_Lp-4c\theta_L-6p\\=&6(\theta_L^2-1)p+3\theta_Lp-4c\theta_L=6\left(\theta_L^2-1\right)p+\theta_L(3p-4c)\end{aligned}$$

In order to show that the latter expression is positive, it is sufficient to show that it is positive at minimum value of p, $p = p^*(\theta_H) = p^*(\theta_L + \frac{1}{2}) = \frac{2c}{2\theta_L - 1}$:

$$
\begin{aligned}
6\left(\theta_L^2 - 1\right) p + \theta_L(3p - 4c) >& 6\left(\theta_L^2 - 1\right) p^*(\theta_L + \frac{1}{2}) + \theta_L(3p^*(\theta_L + \frac{1}{2}) - 4c) \\
=& 2c\Big(\theta_L + 3(1 - \frac{1}{2\theta_L - 1})\Big) > 0
\end{aligned}
$$

The last line is positive as $\theta_L > 1$, which implies $\frac{1}{2\theta_L - 1} < 1$.

Necessity of $\theta_H \leq \theta_L + \frac{1}{2}$. For any pair of (θ_H, θ_L), derivative of the expression in Equation (3) at $p^*(\theta_H)$ is given by

$$
-\frac{(2\theta_H - 2\theta_L + 1)(\theta_H - \theta_L)\left(\frac{c\theta_H}{\theta_H - 1}\right)^{1-\theta_H}}{2c\theta_H^2}
$$

As $\theta_H > \theta_L > 1$, this expression is always negative. Furthermore, for (θ_H, θ_L), derivative of the expression in Equation (3) at $p^*(\theta_L)$ is given by

$$
\frac{(2\theta_H - 2\theta_L - 1)(\theta_H - \theta_L)\left(\frac{c\theta_L}{\theta_L - 1}\right)^{1-\theta_H}}{2c\theta_L^2}
$$

As $\theta_L > 1$, this expression is strictly increasing in θ_H and is zero if $\theta_H = \theta_L + \frac{1}{2}$, which completes the proof.

The concavity comparison of Aguirre et al. (2010). Aguirre et al. (2010) use a condition that is violated in our example. In particular, the condition requires that the demand curve with a lower monopoly price, θ_1, is more convex in the sense that

$$
\frac{D_{pp}(p, \theta_1)}{D_p(p, \theta_1)} \geq \frac{D_{pp}(p, \theta_2)}{D_p(p, \theta_2)}.
$$

In our example,

$$
\begin{aligned}
D_p(p, \theta_i) &= -\theta_i(c + p)^{-\theta_i - 1}, \\
D_{pp}(p, \theta_i) &= \theta_i(\theta_i + 1)(c + p)^{-\theta_i - 2}, \\
\frac{D_{pp}(p, \theta_1)}{D_p(p, \theta_1)} &= -\frac{1 + \theta_i}{c + p}.
\end{aligned}
$$

Therefore, their ranking requires that

$$
\frac{1 + \theta_1}{c + p} \leq \frac{1 + \theta_2}{c + p},
$$

which is violated because $\theta_1 > \theta_2$.

B.4 Example 4

Consider IMG first. We find conditions for 1-IMG, implying α-IMG holds for any α. Recall that Corollary 4 says, as summarized in Figure 3, that 1-IMG holds if the marginal revenue curves are concave and the more elastic demand curve, θ_1, has

1. a higher level, $D(p, \theta_1) \geq D(p, \theta_2)$, and,
2. a more concave revenue curve, $R_{pp}(p, \theta_2) \leq R_{pp}(p, \theta_1)$, or equivalently a less steep marginal revenue curve.

In this example, both marginal revenue curves are linear because $R_{ppp}(p, \theta) = 0$. Additionally, the two marginal revenue curves have the same slope, -2. Therefore, the only condition left to check is $D(p, \theta_1) \geq D(p, \theta_2)$ to ensure that the cross-types price change effect is positive. This condition is

$$c_1 - c_2 \geq (a_2 - a_1)p \tag{32}$$

for all prices in $[\frac{a_1}{2}, \frac{a_2}{2}]$. Because $a_2 \geq a_1$, the right hand side of the above inequality is increasing in p and takes its highest value at $p = p^*(\theta_2) = \frac{a_2}{2}$, so the above condition becomes

$$c_1 - c_2 \geq (a_2 - a_1)\frac{a_2}{2}. \tag{33}$$

To summarize, 1-IMG holds if c_1 is sufficiently larger than c_2, as measured by Equation (33). This is intuitively because as $c_1 - c_2$ increases, the level of demand for the first type θ_1 increases compared to the level of demand for the second type θ_2, which increases the cross-group price change effect. Because the marginal revenue curves are linear and parallel to each other, the price curvature effect is zero. And the within-type price effect is positive because consumers surplus is a convex function of price.

Now consider IMB. We find conditions for $\frac{1}{2}$-IMB, implying α-IMB holds for any $\alpha \geq \frac{1}{2}$. Recall that Corollary 4 says, as summarized in Figure 6, that $\frac{1}{2}$-IMB holds if the marginal revenue curves are convex and the more elastic demand curve, θ_1, has

1. a lower derivative, $D_p(p, \theta_1) \leq D_p(p, \theta_2)$, and,
2. a more convex revenue curve, $R_{pp}(p, \theta_2) \geq R_{pp}(p, \theta_1)$, or equivalently a more steep marginal revenue curve,

and an additional condition that ensures that the within-type price change effect is negative holds. In this example the marginal revenue curves are linear and have the same slope, so the only conditions left to check are regarding the ranking of the derivatives and the technical condition $D_p(p, \theta) + pD_{pp}(p, \theta) \leq 0$ that implies that the within-type price change effect is negative.

The condition $D_p(p, \theta_2) \leq D_p(p, \theta_1)$, ensuring the cross-types price change effect is negative, simply becomes

$$c_1 \leq c_2.$$

The condition $D_p(p, \theta) + pD_{pp}(p, \theta) \leq 0$, ensuring the within-type price change effect is negative, becomes

$$c_i \leq p^2$$

for all prices in $[\frac{a_1}{2}, \frac{a_2}{2}]$ and each $i \in \{1, 2\}$, which given the condition $c_1 \leq c_2$ can be summarized as $c_2 \leq \frac{a_1^2}{4}$. To summarize, $\frac{1}{2}$-IMB holds if each c_i is small enough and c_1 is no more than c_2.