Selling Consumer Data for Profit: Optimal Market-Segmentation Design and its Consequences

Kai Hao Yang*

December 3, 2020

Abstract

A data broker sells *market segmentations* to a producer with private production cost who sells a product to a unit mass of consumers with heterogeneous values. In this setting, I completely characterize the revenue-maximizing mechanisms for the data broker. In particular, every optimal mechanism induces *quasi-perfect price discrimination*—the data broker sells the producer a market segmentation described by a cost-dependent cutoff, such that all the consumers with values above the cutoff end up buying and paying their values while the rest of consumers do not buy. The characterization of optimal mechanisms leads to additional economically relevant implications. I show that the induced market outcomes remain unchanged even if the data broker becomes more active in the product market by gaining the ability to contract on prices; or by becoming an exclusive retailer, who purchases the product and owns the exclusive right to sell the consumers directly. Moreover, vertical integration between the data broker and the producer increases total surplus while leaving the consumer surplus unchanged, due to the fact that consumer surplus is zero under any optimal mechanism for the data broker.

KEYWORDS: Price discrimination, market segmentation, mechanism design, virtual cost, quasi-perfect segmentation, quasi-perfect price discrimination, surplus extraction, outcome-equivalence

JEL CLASSIFICATION:D42, D82, D61, D83, L12

*Cowles Foundation for Economic Research, Yale University, kaihao.yang@yale.edu. I am indebted to my advisor Phil Reny for his constant support and encouragement, and to my thesis committee: Ben Brooks, Emir Kamenica and Doron Ravid for their invaluable guidance and advice. I appreciate the helpful comments and suggestions from Mohammad Akborpour, Dirk Bergemann, Isa Chaves, Yeon-Koo Che, Alex Frankel, Andrew Gianou, Andreas Kleiner, Jacob Leshno, Elliot Lipnowski, Alejandro Maelli, Roger Myerson, Barry Nalebuff, Michael Ostrovsky, Marek Pycia, Daniel Rappoport, Eric Rasmusen, Ilya Segal, Andy Skrzypacz, Wenji Xu and Weijie Zhong. I also thank the participants of several conferences and seminars at which this paper was presented. All errors are my own.

1 Introduction

1.1 Motivation

In the information era, the abundance of personal data has moved the scope of price discrimination far beyond its traditional boundaries such as geography, age, or gender. Extensive usage of consumer data allows one to identify many characteristics of consumers that are relevant to predicting their values, and therefore to create numerous sorts of *market segmentations*—a way to split the market demand into several sub-demands that (horizontally) sum back to the market demand—to facilitate price discrimination. Consequently, "data brokers", with their ownership of massive amount of consumer data and advanced information technology, are able to create such market segmentations and eventually sell these segmentations as products to producers. For instance, online platforms such as Facebook sell[1] a significant amount of consumer information collected via its own platform, including personal characteristics, traveling plans, lifestyles, and text messages. Alternatively, data companies such as Acxiom and Datalogix gather and sell personal information such as government records, financial activities, online activities and medical records to retailers (Federal Trade Commission, 2014).

This paper studies the design of optimal selling mechanisms of a data broker. In this paper, I consider a model where there is one producer with privately known constant marginal cost, who produces and sells a single product to a unit mass of consumers. The consumers have unit demand and the distribution of their values is described by commonly known market demand. Into this environment, I introduce a *data broker*, who does not know the producer's marginal cost of production but can sell *any* market segmentation to the producer via *any* selling mechanism. As the data broker is uncertain about the production cost, and only affects the product market indirectly by selling consumer data to the producer, it is not obvious how the data broker should sell market segmentations to the producer, what market segmentations will be created, and how the sale of consumer data affects economic welfare and allocative outcomes.

As the main result, I completely characterize the revenue-maximizing mechanisms for the data broker. The optimal mechanisms feature *quasi-perfect price discrimination*, an outcome where all the purchasing consumers pay exactly their values, although not every consumer with values above the marginal cost buys the product. Specifically, Theorem 1

[1]In practice, "selling" consumer data can take a wide variety of forms, which include not only traditional physical transactions but also integrated data-sharing agreements/activities. For instance, in a recent full-scale investigation by *The New York Times*, Facebook has formed ongoing partnerships with other firms, including Netflix, Spotify, Apple and Microsoft, and granted these companies accesses to different aspects of consumer data *"in ways that advanced its own interests."* See full news coverage at https://www.nytimes.com/2018/12/18/technology/facebook-privacy.html

shows that every optimal mechanism must create *quasi-perfect segmentations* described by a cost-dependent cutoff. That is, all the consumers with values above the cutoff are separated from each other whereas the consumers with values below the cutoff are pooled with the separated high-value consumers. When pricing optimally under this segmentation, the producer only sells to high-value consumers and induces quasi-perfect price discrimination. Moreover, the cutoff function under any optimal mechanism is exactly the minimum of the (ironed) *virtual marginal cost* function and the optimal uniform price as a function of marginal cost. With proper regularity conditions, Theorem 2 further shows that there is an optimal mechanism where the low-value consumers are pooled uniformly with the separated high-values. In other words, the distribution of consumer values conditional on being below the cutoff remains the same as the market demand in every market segment.

Several economic implications follow accordingly. As the defining feature of quasi-perfect price discrimination, under any optimal mechanism, all the consumers pay their values conditional on buying. This implies that the consumer surplus under any optimal mechanism is zero (Theorem 3). In other words, in terms of consumer surplus, it is *as if* all the information about the consumers' values were revealed to the producer. Furthermore, Theorem 1 also allows a comparison between data brokership and uniform pricing, where no consumer data can be shared. I show that data brokership always increases total surplus (Theorem 4), and can even be Pareto-improving compared with uniform pricing if the data broker has to purchase the data from the consumers (before they learn their values, see Theorem 5).

Another set of relevant questions pertain to how different market regimes would affect market outcomes. More specifically, how would the market outcomes differ if the data broker, instead of merely supplying consumer data to the producers, plays a more active role in the product market? The characterization given by Theorem 1 allows for comparisons across (i) data brokership; (ii) *vertical integration*, where all the private information about production cost is revealed and the data broker merges with the producer; (iii) *exclusive retail*, where the data broker negotiates with the producer and purchases the product, as well as the exclusive right to sell the product, from the producer; and (iv) *price-controlling data brokership*, where the data broker can contract with the producer on prices in addition to providing consumer data. Using the main characterization, I show that vertical integration between the data broker and the producer increases total surplus while leaving the consumer surplus unchanged (Theorem 6). In terms of market outcomes (i.e., data broker's revenue, producer's profit, consumer surplus and the allocation of the product), I show that data brokership is equivalent to both exclusive retail and price-controlling data brokership (Theorem 7).

The rest of this paper is organized as follows. In this section, I continue to discuss related literatures. Followingly, Section 2 provides an illustrative example and Section 3 introduces the model. In Section 4, I characterize the optimal mechanisms of the data broker. In

Section 5, I discuss the consequences of data brokership. Section 6 studies an extension where the feasible market segmentations are limited. Section 7 provides further discussions and Section 8 concludes.

1.2 Related Literature

This paper is related to various streams of literature. In the literature of price discrimination, numerous studies center around the welfare effects of price discrimination. Some of them provide conditions under which third-degree price discrimination increases or decreases total surplus and output (see, for instance, Varian (1985), Aguirre, Cowan, and Vickers (2010) and Cowan (2016)), while Bergemann, Brooks, and Morris (2015) show that any surplus division between the consumers and a monopolist can be achieved by some market segmentation.[2] In those papers, market segmentation is treated as an exogenous object, whereas in this paper, market segmentation is determined endogenously by a data broker's revenue-maximization behavior. Additionally, Ali, Lewis, and Vasserman (2020) study the welfare effect of third-degree price discrimination when the consumers can disclose information about their values voluntarily; while Wei and Green (2020) study another channel of price discrimination (via providing differential information), which does not involve market segmentation.

The current paper is also related to the recent literature of the sale of information by a monopolistic information intermediary. Admati and Pfleiderer (1985) and Admati and Pfleiderer (1990) consider a monopoly who sells information about an asset in a speculative market. Bergemann and Bonatti (2015) explore a pricing problem of a data provider who provides data to facilitate targeted marketing. Bergemann, Bonatti, and Smolin (2018) solve a mechanism design problem in which the designer sells experiments to a decision maker who has private information about his belief. Yang (2019) considers a model where an intermediary can provide information about the product to the consumers and charge the seller for such services. In this regard, the current paper studies the revenue-maximizing mechanism of a data broker who sells consumer information to a producer to facilitate price discrimination.[3]

Methodologically, this paper is related to the literature of mechanism design and information design (see, for instance, Mussa and Rosen (1978), Myerson (1981), Kamenica and Gentzkow (2011) and Bergemann and Morris (2016)), and can be regarded as a mechanism

[2]See also: Haghpanah and Siegel (2020), who further consider segmentations in environments that feature second-degree price discrimination.

[3]Relatedly, Acemoglu, Makhdoumi, Malekian, and Ozdaglar (2019), Bergemann, Bonatti, and Gan (2020) and Ichihashi (2020) examine environments where a data broker buys data from the consumers and then sells the consumer data to downstream firms. Segura-Rodriguez (2020) studies the an environment where information is restricted to a parametrized family and the data-buying firm uses the purchased information to solve a (private) forecast problem.

design problem where the information structure is also part of the design object (see, for instance, Bergemann and Pesendorfer (2007), Yamashita (2017) and Dworczak (2020)).

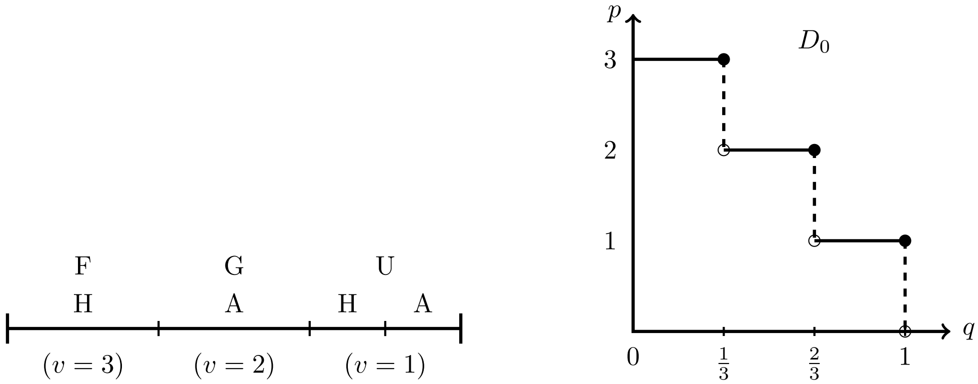
Among the aforementioned papers, Bergemann, Brooks, and Morris (2015), Bergemann, Bonatti, and Smolin (2018) and Yang (2019) are the closest to this paper. Specifically, Bergemann, Brooks, and Morris (2015) explore the welfare implications of different market segmentations, while this paper introduces a data broker who designs the market segmentation in order to maximize revenue. Bergemann, Bonatti, and Smolin (2018) study an environment where the agent has private information about his prior belief and characterize the optimal mechanism in a binary-action, binary-state case; or in a binary-type case. In comparison, this paper studies a revenue maximization problem where the agent's private information is directly payoff-relevant, has a rich action space, and allows for a large class of priors, including those with infinite support. Finally, Yang (2019) solves for optimal mechanisms of an intermediary that can provide information about the product to the consumers, while in this paper I consider the case where an intermediary sells information about the consumers' values to the producer.

2 An Illustrative Example

To fix ideas, consider the following example. A publisher sells an advanced textbook for graduate study. Her (constant) marginal cost of production c is her private information and takes two possible values, $1/4$ or $3/4$, with equal probability. There is a unit mass of consumers with three possible occupations: faculty, undergraduate, and graduate. Each of them makes up $1/3$ of the entire population. It is common knowledge that the textbook has has value $v = 1$ for an undergraduate student, value $v = 2$ for a graduate student and value $v = 3$ for a faculty member. In addition, suppose that among all the undergraduate students, $1/2$ live in houses and $1/2$ live in apartments, whereas all the graduate students live in apartments and all the faculty members live in houses. This economy can be represented by Figure 1, where Figure 1a plots the partitions of the consumers induced by their occupations and residence types and Figure 1b plots the (inverse) market demand D_0.

Suppose that there is a data broker who owns all the data about the consumers (e.g., income, medical records, occupations and residential information) and thus is able to provide any partition on the line in Figure 1a to the publisher so that the publisher can charge different prices to different groups of consumers. How should the data broker sell these data to the publisher? An intuitive guess is that the data broker should sell the most informative data. That is, he should provide the publisher with occupation data so that each consumer's value can be fully revealed. Upon receiving such data, the publisher is able to perfectly price discriminate the consumers. In other words, the value-revealing data creates a *market segmentation* that decomposes the market into three market segments, and each market

Figure 1: Representation of the market



(a) Partitioning consumers

(b) Market demand D_0

segment enables the publisher to perfectly identify the value of the consumers in that market segment. As a result, if the price of the value-revealing data is τ and if the publisher with cost $c \in \{1/4, 3/4\}$ buys the data, her net profit would be

$$\frac{1}{3}(1-c) + \frac{1}{3}(2-c) + \frac{1}{3}(3-c) - \tau.$$

Alternatively, if the publisher with cost c does not buy any data, she would then charge an optimal uniform price (either $1, 2$ or 3, since these are the only possible consumer values) and earn profit

$$\max\left\{(1-c), \frac{2}{3}(2-c), \frac{1}{3}(3-c)\right\}.$$

Therefore, for any τ, the publisher with cost c would buy the value-revealing data if and only if

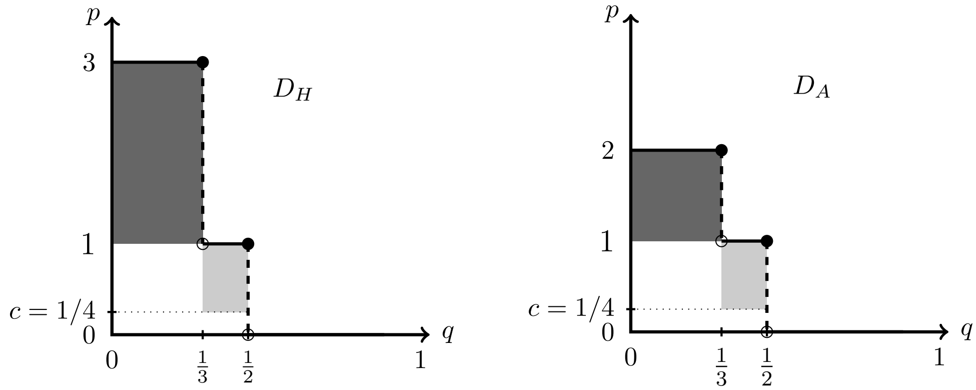
$$\frac{1}{3}(1-c) + \frac{1}{3}(2-c) + \frac{1}{3}(3-c) - \tau \geq \max\left\{(1-c), \frac{2}{3}(2-c), \frac{1}{3}(3-c)\right\},$$

which simplifies to $\tau \leq (2-c)/3$. Thus, since $c \in \{1/4, 3/4\}$, when $\tau \leq 5/12$, the publisher would always buy the value-revealing data regardless of her marginal cost. When $5/12 < \tau \leq 7/12$, the publisher would buy the data only if $c = 1/4$. Therefore, charging a price $\tau = 5/12$ gives the data broker revenue $5/12$ whereas charging a price $\tau = 7/12$ gives the data broker revenue $7/12 \times 1/2 = 7/24 < 5/12$. Hence the optimal price for the value-revealing data is $5/12$ and it gives the data broker revenue $5/12$.

However, the data broker can in fact improve his revenue by creating a menu consisting of not just the value-revealing data. To see this, consider the following menu of data

$$\mathcal{M}^* = \left\{\left(\text{residential data}, \tau = \frac{1}{3}\right), \left(\text{value-revealing data}, \tau = \frac{7}{12}\right)\right\}.$$

Figure 2: Market segmentation induced by residential data



Notice that the residential data creates a market segmentation with two segments described by two demand functions, D_H and D_A. Segment D_H contains all of the consumers with $v = 3$ and $1/2$ of the consumers with $v = 1$ (i.e., those who live in houses), while segment D_A contains all of the consumers with $v = 2$ and $1/2$ of the consumers with $v = 1$ (i.e., those who live in apartments). Figure 2 plots this market segmentation. From Figure 2, it can be seen that $D_H + D_A = D_0$. Moreover, for the publisher with $c = 1/4$, the difference in profit between charging price 3 (2) and charging price 1 in segment D_H (D_A) is exactly the difference between the area of the darker region and the area of the lighter region depicted in Figure 2. Therefore, since the area of the lighter region is smaller than the area of the darker region, charging a price of 3 (2) is better than charging a price of 1 in segment D_H (D_A). Thus, as there are only two possible values in each segment, charging a price of 3 (2) is optimal for the publisher under segment D_H (D_A). This is also the case when her cost is $c = 3/4$, since the area of the lighter region would decrease and the area of the darker region would remain unchanged when the marginal cost changes from $1/4$ to $3/4$. As a result, regardless of her marginal cost, the publisher will sell to all the consumers with values $v = 3$ and $v = 2$ by charging exactly their values upon receiving the residential data.

With this observation, it then follows that when $c = 1/4$, the publisher would prefer buying the value-revealing data (at the price of $\tau = 7/12$) whereas when $c = 3/4$, the publisher would prefer buying the residential data (at the price of $\tau = 1/3$). Therefore, when menu $\mathcal{M}^*$ is provided, the data broker's revenue is

$$(0.5)\frac{1}{3} + (0.5)\frac{7}{12} = \frac{11}{24} > \frac{5}{12},$$

which is higher than what can be obtained by selling value-revealing data alone. The intuition behind such an improvement is simple. When selling the value-revealing data alone, the publisher with lower marginal cost retains more rents because the data broker would have to incentivize the high-cost publisher to purchase. However, by creating a menu containing both the value-revealing data and the residential data, the data broker can further screen the publisher. To see this, notice that even though the residential data becomes less informative than the value-revealing data, the only extra benefit of the value-revealing data is for the publisher to be able to price discriminate the consumers with $v = 1$. Thus, when the publisher's marginal cost is high (i.e., $c = 3/4$), the additional information given by the value-revealing data is less useful to the publisher because the gains from selling to consumers with $v = 1$ are small. By contrast, when the publisher has a low marginal cost (i.e., $c = 1/4$), the value-revealing data is more valuable to the publisher since the gains from selling to consumers with $v = 1$ are larger. Therefore, by providing a menu that contains two different datasets with different prices, the data broker can screen the publisher and extract more revenue from the publisher with lower marginal cost than by selling the value-revealing data alone.

In fact, as it will be shown in Section 4, $\mathcal{M}^*$ is an optimal mechanism of the data broker. The optimal mechanism $\mathcal{M}^*$ has several notable features. First, when $c = 3/4$, the high-value consumers ($v = 2$ and $v = 3$) are separated from each other whereas the low-value consumers ($v = 1$) are pooled together with the high-value consumers. This induces a market outcome where consumers with values $v = 2$ and $v = 3$ buy the textbook by paying their values, whereas the consumers with $v = 1$ do not buy, even if their value is greater than the publisher's marginal cost $3/4$. In other words, in order to maximize revenue, the data broker would sometimes discourage (ex-post) efficient trades. Second, all the purchasing consumers are paying exactly about their values, which implies that consumer surplus is zero. Finally, even though every purchasing consumer pays their value, the high-cost publisher never learns exactly each individual consumer's value. These features are not specific to this simple example. In fact, all of them hold in a general class of environments, which will be explored in Section 4.

3 Model

3.1 Notation

The following notation is used throughout the paper. For any Polish space X, $\Delta(X)$ denotes the set of probability measures on X where X is endowed with the Borel σ-algebra and $\Delta(X)$ is endowed with the with weak-* topology. When $X = [\underline{x}, \overline{x}] \subseteq \mathbb{R}$ is an interval, let $\mathcal{D}(X)$ denote the collection of nonincreasing and upper-semicontinuous functions $D : \mathbb{R}_+ \to [0, 1]$ such that $D(\underline{x}) = 1$, $D(\overline{x}^+) = 0$.[4] Since $\mathcal{D}(X)$ and $\Delta(X)$ are bijective,[5] for any $D \in \mathcal{D}(X)$,

[4] As a convention, for any function f defined on $\mathbb{R}_+$, $f(x^+)$ denotes the right limit of f at x.

[5] This is because for any $D \in \mathcal{D}(X)$, the right limit of $1 - D$ is nondecreasing and right-continuous.

let $m^D \in \Delta(X)$ be the probability measure associated with D and define the integral

$$\int_A h(x)D(\mathrm{d}x) := \int_A h(x)m^D(\mathrm{d}x),$$

for any measurable $h : X \to \mathbb{R}$. Then, endow $\mathcal{D}(X)$ with the weak-* topology and the Borel σ-algebra using this integral (details in Appendix A). Also, write $\mathrm{supp}(D) := \mathrm{supp}(m^D)$.

3.2 Primitives

There is a single product, a unit mass of consumers with unit demand, a producer for this product (she), and a *data broker* (he). Across the consumers, their values v for the product are distributed according to a *market demand* $D_0 \in \mathcal{D} := \mathcal{D}(V)$, where $D_0(p)$ denotes the share of consumers whose values are above p and $V = [\underline{v}, \overline{v}] \subset \mathbb{R}_+$ is a compact interval. Each consumer knows their own value. For the rest of the paper, D_0 is said to be *regular* if the function $p \mapsto (p - c)D_0(p)$ is single-peaked on $\mathrm{supp}(D_0)$ for all $c \geq 0$.[6]

The producer has a constant marginal cost of production $c \in C = [\underline{c}, \overline{c}] \subset \mathbb{R}_+$ for some $0 \leq \underline{c} < \overline{c} < \infty$. The marginal cost c is private information to the producer and follows a cumulative distribution G, where G has a density $g > 0$ and induces a *virtual (marginal) cost* function ϕ_G, defined as $\phi_G(c) := c + G(c)/g(c)$ for all $c \in C$. Henceforth, G is said to be *regular* if ϕ_G is increasing.

The data broker can create any *market segmentation* (using consumer data), which is a probability measure $s \in \Delta(\mathcal{D})$ that satisfies the following condition

$$\int_{\mathcal{D}} D(p)s(\mathrm{d}D) = D_0(p),\ \forall p \in V. \tag{1}$$

That is, a segmentation is a way to split the market demand D_0 into different *market segments* that average back to the market demand.[7] Let $\mathcal{S}$ denote the set of segmentations.

[6] If D_0 is (strictly) decreasing on V, then this is equivalent to saying that the marginal revenue function induced by D_0 is decreasing. If, furthermore, D_0 is absolutely continuous, then this is equivalent to saying that $1 - D_0$ is regular in the sense of Myerson (1981).

[7] As illustrated in the motivating example, different consumer data induce different partitions of consumers' characteristics and therefore different ways to split D_0 into a collection of demand functions that sum up to D_0. Thus, given a market segmentation s, each market segment $D \in \mathrm{supp}(s)$ can be interpreted as a group of consumers who share some common characteristics (e.g., house residents). Notice that by allowing the data broker to provide *any* market segmentation, it is implicitly assumed that the data broker always has sufficient data to identify each consumer's value and is able to segment the consumers according to their values arbitrarily. In Section 6, I consider an extension where the data broker has imperfect information about the consumers' values.

3.3 Timing of the Events

First, the data broker proposes a *mechanism*, which contains a set of available messages that the producer can send, as well as mappings that specify the market segmentation and the amount of transfers as functions of the messages. Next, the producer decides whether to participate in the mechanism. If she opts out, she only operates under D_0 without any further segmentations and pays nothing. If the producer participates in the mechanism, she sends a message from the message space, pays the associated transfer, and then operates under the associated market segmentation.

Given any segmentation $s \in \mathcal{S}$, the producer engages in price discrimination by choosing a price $p \geq 0$ in each segment $D \in \text{supp}(s)$.[8] To maximize profit, for any segment $D \in \text{supp}(s)$, the producer with marginal cost c solves

$$\max_{p \in \mathbb{R}_+} (p - c)D(p).$$

For any $c \in C$ and any $D \in \mathcal{D}$, let $\boldsymbol{P}_D(c)$ denote the set of optimal prices for the producer with marginal cost c under market segment D. As a convention, regard $\boldsymbol{P}$ as a correspondence on $\mathcal{D} \times C$ and if $\boldsymbol{p}$ is a selection for $\boldsymbol{P}$, write $\boldsymbol{p} \in \boldsymbol{P}$.[9] Furthermore, for any $c \in C$ and any $D \in \mathcal{D}$, let

$$\pi_D(c) := \max_{p \in \mathbb{R}_+} (p - c)D(p)$$

denote the maximized profit of the producer. Also, let

$$\overline{\boldsymbol{p}}_D(c) := \max \boldsymbol{P}_D(c)$$

be the largest optimal price for the producer with marginal cost c under market segment D.[10] For conciseness, let $\overline{\boldsymbol{p}}_0 := \overline{\boldsymbol{p}}_{D_0}$.

Throughout Section 4 and Section 5, I impose the following technical assumption on the market demand D_0 and the distribution G.

[8] It is without loss of generality to restrict attention to posted price mechanisms even though the producer has private information about c when designing selling mechanisms. This is because the environment features independent private values and quasi-linear payoffs, and both the producer's and the consumers' payoffs are monotone in their types. By Proposition 8 of Mylovanov and Tröger (2014), it is as if c is commonly known when the producer designs selling mechanisms. Therefore, since the consumers have unit demand, according to Myerson (1981) and Riley and Zeckhauser (1983), it is without loss to restrict attention to posted price mechanisms.

[9] For notational conveniences, I restrict the feasible prices for each producer to a large enough compact interval $\overline{V} \subset \mathbb{R}_+$ such that $V \subsetneq \overline{V}$. With this restriction, $\boldsymbol{P}_D(c)$ would be a subset of a compact interval for all $D \in \mathcal{D}$ and for all $c \in C$. Since V itself is bounded, this restriction is simply a notational convention and does not affect the model at all.

[10] $\overline{\boldsymbol{p}}$ is well-defined under the notational convention stated in footnote 9, as $\boldsymbol{P}_D$ is a closed (implied by upper-semicontinuity of D) subset of a compact set $\bar{V}$.

Assumption 1. *The function* $c \mapsto \max\{g(c)(\phi_G(c) - \overline{\boldsymbol{p}}_0(c)), 0\}$ *is nondecreasing.*

Assumption 1 permits a wide class of (D_0, G) and includes many common examples.[11] Also, it does not require regularities of either D_0 or G (nor is it implied by regularities of D_0 and G). In Section 7, I will further discuss this assumption, including how the results rely on it, its relaxations, as well as several economically interpretable sufficient conditions.

3.4 Mechanism

When proposing mechanisms, by the revelation principle (Myerson, 1979), it is without loss to restrict the data broker's choice of mechanisms to incentive compatible and individually rational *direct mechanisms* that ask the producer to report her marginal cost and then provide her with the segmentation and determine the transfer accordingly.[12]

Formally, a mechanism is a pair (σ, τ), where $\sigma : C \to \mathcal{S}$, $\tau : C \to \mathbb{R}$ are measurable functions. Given a mechanism (σ, τ), for each report $c \in C$, $\sigma(c) \in \mathcal{S}$ stands for the market segmentation provided to the producer, and $\tau(c) \in \mathbb{R}$ stands for the amount the producer pays to the data broker. Moreover, any measurable $\sigma : C \to \mathcal{S}$ is called a *segmentation scheme* (or sometimes, a *scheme*).

A mechanism (σ, τ) is incentive compatible if for all $c, c' \in C$,

$$\int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c) - \tau(c) \geq \int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c') - \tau(c'). \tag{IC}$$

Also, since the producer can always sell to the consumers by charging a uniform price, a mechanism (σ, τ) is individually rational if for all $c \in C$,

$$\int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c) - \tau(c) \geq \pi_{D_0}(c). \tag{IR}$$

Henceforth, a mechanism (σ, τ) is said to be *incentive feasible* if it is incentive compatible and individually rational. A segmentation scheme σ is said to be implementable if there exists a measurable $\tau : C \to \mathbb{R}$ such that (σ, τ) is incentive feasible. The goal of the data broker is to maximize expected revenue $\mathbb{E}_G[\tau(c)]$ by choosing an incentive feasible mechanism.

The data broker's revenue maximization problem exhibits several noticeable features. First, the object being allocated is infinite dimensional. After all, the data broker sells market segmentations to the producer as opposed to a one-dimensional quality or quantity variable in classical mechanism design problems (e.g., Mussa and Rosen (1978), Myerson (1981) and Maskin and Riley (1984)). In particular, it is not clear whether there exists a partial order on

[11] For instance, if D_0 is linear demand and G is uniform; or if both D_0 and G are exponential on some intervals; or if D_0 and G are such that $D_0(v) = (1-v)^\beta$, $G(c) = c^\alpha$, for all $v \in [0,1]$, $c \in [0,1]$, for any $\alpha, \beta > 0$; or if D_0 and G take one of the aforementioned forms.

[12] Henceforth, unless otherwise noted, a *mechanism* stands for a *direct mechanism*.

the space of market segmentations that would lead to the single-crossing property commonly assumed in low-dimensional screening problems.[13] Second, the producer's outside option is type-dependent. This is because the producer has direct access to the consumers, and is only buying the additional information about the consumers' values.

As another remark, the model introduced above is equivalent to a model where there is one producer with private cost c and one consumer with private value v, where c and v are independently drawn from G and m^{D_0}, respectively. With this interpretation, a segmentation $s \in \mathcal{S}$ is then equivalent to a Blackwell experiment that provides the producer with information regarding the consumer's private value. Throughout the paper, the analyses and results are stated in terms of the version with a continuum of consumers, yet every statement and interpretation has an equivalent counterpart in the version with one consumer who has a private value.

4 Optimal Segmentation Design

In what follows, I characterize the data broker's optimal mechanisms. To this end, I first introduce a crucial class of mechanisms. Then I characterize the optimal mechanisms by this class.

4.1 Quasi-Perfect Segmentations and Quasi-Perfect Price Discrimination

As illustrated in the motivating example, to elicit private information from the producer, the data broker may sometimes wish to discourage sales even when there are gains from trade. In addition, the data broker would wish to extract as much surplus as possible by providing market segmentations under which all the purchasing consumers pay their values. These two features jointly lead to a specific form of market segmentation, which will be referred as *quasi-perfect segmentations*.

Definition 1. For any $c \in C$ and any $\kappa \geq c$, a segmentation $s \in \mathcal{S}$ is a *κ-quasi-perfect segmentation* for c if for s-almost all $D \in \mathcal{D}$, either $D(c) = 0$, or the set $\{v \in \text{supp}(D) : v \geq \kappa\}$ is a singleton and is a subset of $\boldsymbol{P}_D(c)$.

A κ-quasi-perfect segmentation for c is a segmentation that separates all the consumers with $v \geq \kappa$ while pooling the rest of the consumers with each of them, in a way so that every

[13] It is noteworthy that although Sinander (2020) studies a similar problem of allocating Blackwell experiments to a one-dimensional type space and shows that any Blackwell-monotone allocation rule is implementable (see Proposition 2 of Sinander (2020)), a key assumption is violated here. That is, for any $c \in C$, $\pi'_D(c)$ is not continuous in D in general. In fact, in this setting, Blackwell-monotone allocation rules may not be implementable.

market segment with positive trading volume[14] contains one and only one consumer-value $v \geq \kappa$ and that this v is an optimal price for the producer with marginal cost c. Notice that a κ-quasi-perfect segmentation for c induces *κ-quasi-perfect price discrimination* when the producer's marginal cost is c and when she charges the largest optimal price in (almost) all segments. Namely, a consumer with value v buys the product if and only if $v \geq \kappa$ and all purchasing consumers pay exactly their values. For instance, in the example given by Section 2, the residential data creates a 2-quasi-perfect segmentation for $c \in \{1/4, 3/4\}$. With Definition 1, I now define the following:

Definition 2. Given any function $\psi : C \to \mathbb{R}$ with $c \leq \psi(c)$ for all $c \in C$:

1. A segmentation scheme σ is a *ψ-quasi-perfect scheme* if for G-almost all $c \in C$, $\sigma(c)$ is a $\psi(c)$-quasi-perfect segmentation for c.

2. A mechanism (σ, τ) is a *ψ-quasi-perfect mechanism* if σ is a ψ-quasi-perfect scheme and if the producer with marginal cost $\overline{c}$, when reporting truthfully, has net profit $\pi_{D_0}(\overline{c})$.

4.2 Characterization of the Optimal Mechanisms

With the definitions above, the main characterization of this paper can be stated. For any $c \in C$, define $\overline{\varphi}_G(c) := \min\{\varphi_G(c), \overline{\boldsymbol{p}}_0(c)\}$, where φ_G is the ironed virtual cost function.[15]

Theorem 1 (Optimal Mechanism)**.** *The set of optimal mechanisms is nonempty and is exactly the set of incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanisms. Furthermore, every optimal mechanism induces $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G-almost all $c \in C$.*

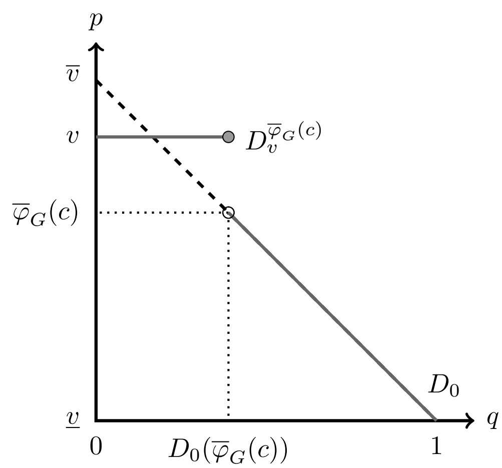
From the definition of quasi-perfect segmentations, there are some degrees of freedom regarding the ways to pool the low-value consumers with the high-values. Indeed, according to Theorem 1, any mechanism is optimal as long as the low-value consumers are pooled with the high-values in a way such that it is incentive feasible and is $\overline{\varphi}_G$-quasi-perfect. Therefore, there might multiple optimal mechanisms.

Nevertheless, the outcome induced by any optimal mechanism is unique. Under any optimal mechanism, for (almost) all marginal cost $c \in C$, a consumer with value v buys the product if and only if $v \geq \overline{\varphi}_G(c)$ and all the purchasing consumers pay their values. In other words, the multiplicity only accounts for the off-path incentives. Furthermore, there is always an explicit way to construct an optimal mechanism (see details in the Online Appendix). In

[14] Notice that when the producer's marginal cost is c, no trade occurs in market segment D if and only if $D(c) = 0$.

[15] Ironing in the sense of Myerson (1981).

Figure 3: Market segment $D_v^{\overline{\varphi}_G(c)}$



fact, when the market demand D_0 is regular, this construction is particularly straightforward: The low-value consumers are spread uniformly across all the market segments. More specifically, for any $c \in C$ and for any $v \geq \overline{\varphi}(c)$, define market segment $D_v^{\overline{\varphi}_G(c)} \in \mathcal{D}$ as

$$D_v^{\overline{\varphi}_G(c)}(p) := \begin{cases} D_0(p), & \text{if } p \in [\underline{v}, \overline{\varphi}_G(c)] \\ D_0(\overline{\varphi}_G(c)), & \text{if } p \in (\overline{\varphi}_G(c), v] \\ 0, & \text{if } p \in (v, \overline{v}] \end{cases}, \tag{2}$$

for all $p \in V$. Then, for any $c \in C$ and for any $p \in [\overline{\varphi}_G(c), \overline{v}]$, let

$$\sigma^* \left(\left\{ D_v^{\overline{\varphi}_G(c)} : v \geq p \right\} \middle| c \right) := \frac{D_0(p)}{D_0(\overline{\varphi}_G(c))}. \tag{3}$$

In other words, for any $c \in C$, $\sigma^*(c)$ induces market segments $\{D_v^{\overline{\varphi}_G(c)}\}_{v \in [\overline{\varphi}_G(c), \overline{v}]}$, which belong to a one-dimensional family indexed by $v \in [\overline{\varphi}_G(c), \overline{v}]$ and are distributed according to the market demand D_0 conditional on $[\overline{\varphi}_G(c), \overline{v}]$ under $\sigma^*(c)$ (Notice that this implies $\sigma^*(c) \in \mathcal{S}$). [16] Figure 3 illustrates σ^* by plotting the (inverse) demand[17] of a generic market segment $D_v^{\overline{\varphi}_G(c)}$ induced by $\sigma^*(c)$ (the dashed line represents the market demand D_0). This inverse demand has a jump at $D_0(\overline{\varphi}_G(c))$. To the left of $D_0(\overline{\varphi}_G(c))$, all the consumer values are concentrated at v, whereas the distribution of the consumer values to the right of $D_0(\overline{\varphi}_G(c))$ remains the same as that under D_0.

With this definition, it turns out that when D_0 is regular, as it will be shown below, there exists a unique transfer scheme $\tau^* : C \to \mathbb{R}$ such that (σ^*, τ^*) is an incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism. Thus, by Theorem 1, (σ^*, τ^*) is optimal. Henceforth, I refer the mechanism (σ^*, τ^*) as the *canonical* $\overline{\varphi}_G$-quasi-perfect mechanism.

[16] Notice that $\sigma^* : C \to \mathcal{S}$ is well-defined and measurable since for all $c \in C$, $v \mapsto D_v^{\overline{\varphi}_G(c)}$ is a measurable function from V to $\mathcal{D}$ and since $D_0 \circ \overline{\varphi}_G : C \to [0, 1]$ is also measurable.

[17] See Appendix A for the formal definition of inverse demands.

Theorem 2. *Suppose that D_0 is regular. Then the canonical $\overline{\varphi}_G$-quasi-perfect mechanism (σ^*, τ^*) is optimal.*

In what follows, I will outline the main ideas of the proof of Theorem 1 (which also lead to the proof of Theorem 2). Details of the proof can be found in Appendix B. I first derive a revenue-equivalence formula and characterize the incentive compatible mechanisms. Next, I identify an upper bound $\bar{R}$ for the data broker's revenue. Then I construct a feasible mechanism that attains $\bar{R}$, which would in turn imply every incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism is optimal. Finally, I argue that any mechanism that gives revenue $\bar{R}$ must be $\overline{\varphi}_G$-quasi-perfect.

To highlight the main insights and avoid unnecessary complications, in this subsection, I impose some further assumptions in addition to Assumption 1. More precisely, throughout the remaining part of Section 4.2, I assume that D_0 and G are regular and that

$$\phi_G(c) \leq \overline{\boldsymbol{p}}_0(c), \; \forall c \in C. \tag{4}$$

Notice that (4) is a sufficient condition for Assumption 1. With these additional conditions, $\overline{\varphi}_G(c) = \phi_G(c)$ for all $c \in C$ and hence $\overline{\varphi}_G$ can be replaced by the virtual cost function ϕ_G. Among these assumptions, regularity of G is purely for conciseness and can be relaxed by ironing ϕ_G. Regularity of D_0 simplifies the construction of the mechanism that attains $\bar{R}$. Without regularity of D_0, the construction is more involved and can be found in the Online Appendix. Lastly, (4) allows a straightforward construction of the revenue upper bound $\bar{R}$. Without (4), the upper bound $\bar{R}$ may not be attainable and must be replaced by a tighter upper bound, which will be discussed in Section 5. Nonetheless, it is noteworthy that all the lemmas stated in this section do not rely on any of these assumptions, nor on Assumption 1.

The Revenue Equivalence Formula and an Upper Bound for Revenue

Even though the data broker's problem is more complex compared with a standard monopolistic screening problem due to the high-dimensionality nature of market segmentations, a revenue-equivalence formula can still be derived by properly invoking the envelope theorem. To see this, notice that for any incentive compatible mechanism (σ, τ), the indirect utility of a producer with marginal cost c is

$$\begin{aligned} U(c) &:= \int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c) - \tau(c) \\ &= \max_{c' \in C} \left[\int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c') - \tau(c') \right]. \end{aligned}$$

By the envelope theorem, the derivative of U is simply the partial derivative of the objective function evaluated at the optimum. That is,

$$U'(c) = \int_{\mathcal{D}} \pi'_D(c)\sigma(\mathrm{d}D|c).$$

Moreover, since $\pi_D(c)$ is the optimal profit of the producer with marginal cost c under segment D, again by the envelope theorem, for all $c \in C$,

$$\pi'_D(c) = -D(\overline{\boldsymbol{p}}_D(c)). \tag{5}$$

Together,

$$U(c) = U(\overline{c}) + \int_c^{\overline{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) \right) \mathrm{d}z,\ \forall c \in C.$$

Therefore, under any incentive compatible mechanism (σ, τ), if a producer with marginal cost c misreports a marginal cost c' and sets prices optimally, the deviation gain can be written as

$$\begin{aligned}
&U(c) - \left(\int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c') - \tau(c') \right) \\
=& \int_{\mathcal{D}} [\pi_D(c) - \pi_D(c')]\sigma(\mathrm{d}D|c') - (U(c) - U(c')) \\
=& \int_c^{c'} \left[\int_{\mathcal{D}} -\pi'_D(z)\sigma(\mathrm{d}D|c') - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) \right] \mathrm{d}z \\
=& \int_c^{c'} \left[\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|c') - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) \right] \mathrm{d}z.
\end{aligned}$$

Together, these lead to Lemma 1 below.

Lemma 1. *A mechanism (σ, τ) is incentive compatible if and only if:*

1. *For all $c \in C$,*

$$\tau(c) = \int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c) - \int_c^{\overline{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z)\, \mathrm{d}z \right) - U(\overline{c}).$$

2. *For all $c, c' \in C$,*

$$\int_c^{c'} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))(\sigma(\mathrm{d}D|z) - \sigma(\mathrm{d}D|c')) \right) \mathrm{d}z \geq 0.$$

Furthermore, $\overline{\boldsymbol{p}}$ can be replaced by any $\boldsymbol{p} \in \boldsymbol{P}$ for the "only if" part.

The proof of Lemma 1 can be found in Appendix B. It formalizes the heuristic arguments above by using the envelope theorem of Milgrom and Segal (2002). In essence, condition 1 in Lemma 1 is a generalized revenue-equivalence formula stating that the transfer τ must be determined by σ up to a constant, whereas condition 2 in Lemma 1 is reminiscent of Lemma 1 of Pavan, Segal, and Toikka (2014), and is sometimes referred as the *integral monotonicity* condition that guarantees global incentive compatibility in various mechanism design problems with multi-dimensional allocation spaces (see, for instance, Rochet (1987), Carbajal and Ely (2013), Pavan, Segal, and Toikka (2014), Bergemann and Valimaki (2019)).

From Lemma 1, for any incentive compatible mechanism (σ, τ), the data broker's expected revenue can be written as

$$\mathbb{E}_G[\tau(c)] = \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c))\, D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - U(\overline{c}), \tag{6}$$

which can be interpreted as the expected *virtual profit* net of a constant. That is, maximizing the data broker's expected revenue by choosing an incentive feasible mechanism (σ, τ) is equivalent to maximizing the expected virtual profit—the profit of the producer if her marginal cost c is replaced by the virtual marginal cost $\phi_G(c)$ while she still prices optimally according to marginal cost c—by choosing an implementable scheme σ.

With (6), there is an immediate upper bound for the data broker's revenue. First notice that since the producer's outside option is $\pi_{D_0}(c)$ when her cost is c, for an incentive compatible mechanism (σ, τ) to be individually rational, it must be that $U(\overline{c}) \geq \bar{\pi} := \pi_{D_0}(\overline{c})$. Moreover, for any $c \in C$,

$$\begin{aligned}\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) &\leq \int_{\mathcal{D}} \max_{p \in \mathbb{R}_+} [(p - \phi_G(c))D(p)]\sigma(\mathrm{d}D|c) \\ &\leq \int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v),\end{aligned}$$

where the second inequality holds because the last term is the total gains from trade in the economy when the producer's marginal cost is $\phi_G(c)$. Together with (6), it then follows that

$$\begin{aligned}\bar{R} :=& \int_C \left(\int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \bar{\pi} \\ \geq& \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - U(\overline{c}) \\ =& \mathbb{E}_G[\tau(c)].\end{aligned}$$

In other words, the upper bound $\bar{R}$ is constructed by ignoring the individual rationality constraints and the global incentive compatibility constraints (i.e., condition 2 in Lemma 1), and by compelling the producer to charge prices that are optimal when her marginal cost is replaced by the virtual marginal cost.

Attaining $\bar{R}$

By the definition of quasi-perfect segmentations, for any nondecreasing function $\psi : C \to \mathbb{R}_+$ and for any ψ-quasi-perfect scheme σ, given any truthful report $c \in C$, $\sigma(c)$ must induce $\psi(c)$-quasi-perfect price discrimination when the producer charges the largest optimal price in (almost) every segment. This means that all the consumers with $v \geq \psi(c)$ would buy the product by paying exactly their values whereas all the consumers with values $v < \psi(c)$ would not buy. As a result, all the surplus of consumers with $v \geq \psi(c)$ would be extracted and the trade volume must be the share of consumers with $v \geq \psi(c)$.[18] Namely, for all $c \in C$,

$$\int_{\mathcal{D}} \overline{\boldsymbol{p}}_D(c) D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) = \int_{\{v \geq \psi(c)\}} v D_0(\mathrm{d}v) \tag{7}$$

and

$$\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) = D_0(\psi(c)). \tag{8}$$

Therefore, if there is an incentive feasible ϕ_G-quasi-perfect mechanism (σ, τ), then by Lemma 1, the data broker can attain revenue

$$\begin{aligned}\mathbb{E}[\tau(c)] =& \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{\pi} \\ =& \int_C \left(\int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \bar{\pi} \\ =& \bar{R}.\end{aligned} \tag{9}$$

However, not every quasi-perfect scheme is implementable. To ensure incentive compatibility, the integral inequality given by condition 2 in Lemma 1 must be satisfied. While this condition involves a continuum of constraints and is difficult to check, the following lemma provides a simpler sufficient condition.

Lemma 2. *For any nondecreasing function $\psi : C \to \mathbb{R}_+$ with $\psi(c) \geq c$ for all $c \in C$, and for any ψ-quasi-perfect scheme σ, there exists a transfer scheme $\tau : C \to \mathbb{R}$ such that (σ, τ) is incentive compatible if for any $c \in C$,*

$$\psi(z) \leq \overline{\boldsymbol{p}}_D(z), \tag{10}$$

for (Lebesgue)-almost all $z \in [\underline{c}, c]$ and for all $D \in \mathrm{supp}(\sigma(c))$.

Essentially, Lemma 2 is a sufficient condition that reduces the integral inequalities in Lemma 1 to pointwise inequalities. Details about the proof can be found in Appendix B. The crucial step is to notice that for a ψ-quasi-perfect scheme, there are always no downward-deviation incentives (i.e., a producer with cost c would never have an incentive to misreport $c' < c$), as a higher-cost producer would find the gains from reducing the cutoff less beneficial than the increment in transfer. Furthermore, the pointwise condition (10) is sufficient to rule out upward-deviation incentives. Together, Lemma 2 then follows.

After simplifying the incentive constraints, the following lemma then provides a crucial sufficient condition for there to exist an incentive compatible ψ-quasi-perfect mechanism.

[18] Formal arguments are in the proof of Lemma 5, which can be found in the Online Appendix.

Lemma 3. *For any nondecreasing function $\psi : C \to \mathbb{R}_+$ such that that $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$, there exists a ψ-quasi-perfect scheme σ that satisfies (10).*

A direct consequence of Lemma 2 and Lemma 3 is that there exists an incentive compatible ϕ_G-quasi-perfect mechanism (σ, τ), provided that G is regular and (4) holds. Furthermore, for any $c \in C$, (4) also implies that

$$\int_c^{\overline{c}} D_0(\phi_G(z))\,\mathrm{d}z \geq \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z.$$

Together, by Lemma 1 and (5), after possibly adding a constant to τ so that the indirect utility of the producer with cost $\overline{c}$ equals to $\overline{\pi}$, (σ, τ) is an incentive feasible ϕ_G-quasi-perfect mechanism, which in turn implies that (σ, τ) is optimal. Combined with (9), it then follows that any incentive feasible ϕ_G-quasi-perfect mechanism is optimal.

The proof of Lemma 3 is by construction. For arbitrary $D_0 \in \mathcal{D}$, the desired segmentation scheme is constructed by first approximating D_0 with a sequence of step functions $\{D_n\} \subseteq \mathcal{D}$ that converges to D_0, and then by finding a desired ψ-quasi perfect scheme σ_n of each D_n. Together with a continuity property of quasi-perfect mechanisms and optimal prices, the limit of $\{\sigma_n\}$ is then a desired ψ-quasi-perfect scheme. Detailed arguments for this general case can be found in the Online Appendix. Here, I provide a simpler proof for the case where D_0 is regular.

Proof of Lemma 3 (regular D_0). For any $c \in C$ and for any $v \in [\psi(c), \overline{v}]$, let $D_v^{\psi(c)} \in \mathcal{D}$ be defined as (2) with $\overline{\varphi}_G(c)$ replaced by $\psi(c)$. Also, let $\sigma^* : C \to \Delta(\mathcal{D})$ be defined as (3) with $\overline{\varphi}_G$ replaced by ψ. By construction, $\sigma^*(c) \in \mathcal{S}$ for all $c \in C$. Furthermore, σ^* is a ψ-quasi-perfect scheme satisfying (10). To see this, for any $c \in C$, let $p^{\psi(c)} := \min\{v \in \mathrm{supp}(D_0) : v \geq \psi(c)\}$. By the hypothesis that $\psi(c) \leq \overline{\boldsymbol{p}}_0(c)$, it must be $p^{\psi(c)} \leq \overline{\boldsymbol{p}}_0(c)$. This in turns implies that, by regularity of D_0 (i.e., singled-peakedness of $p \mapsto (p-c)D_0(p)$), $(p-c)D_0(p) \leq (p^{\psi(c)}-c)D_0(p^{\psi(c)})$ for all $p \leq \psi(c)$. Therefore, for any $v \in [\psi(c), \overline{v}] \cap \mathrm{supp}(D_0)$, since $D_0(p^{\psi(c)}) = D_0(\psi(c)) = D_v^{\psi(c)}(v)$ and since $D_v^{\psi(c)}(p) = D_0(p)$ for all $p \leq p^{\psi(c)}$, it must be that

$$(p-c)D_v^{\psi(c)}(p) = (p-c)D_0(p) \leq (p^{\psi(c)}-c)D_0(p^{\psi(c)}) \leq (v-c)D_0(p^{\psi(c)}) = (v-c)D_v^{\psi(c)}(v),$$

for all $p \leq p^{\psi(c)}$, where the second inequality follows from the fact that $v \geq p^{\psi(c)}$ for all $v \in [\psi(c), \overline{v}] \cap \mathrm{supp}(D_0)$. Therefore, since $(p^{\psi(c)}, v) \cap \mathrm{supp}(D_v^{\psi(c)}) = \emptyset$, it follows that $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(c) = v$ and hence $\sigma^*(c)$ is indeed a $\psi(c)$-quasi-perfect segmentation for c.

Furthermore, for any $z \leq c$ and for any $v \geq \psi(c)$, since $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}$ is nonincreasing, it must be that either $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z) = v$ or $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z) < \psi(c)$. In the former case, since ψ is nondecreasing, it then follows that $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z) = v \geq \psi(c) \geq \psi(z)$, as desired. In the latter case, since

$D_v^{\psi(c)}(p) = D_0(p)$ for all $p \leq \psi(c)$ and since $p \mapsto (p-z)D_0(p)$ is singled-peaked, $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z)$ must be the largest optimal price for the producer under D_0 as well. That is, $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z) = \overline{\boldsymbol{p}}_0(z)$. Combined with the hypothesis that $\psi(z) \leq \overline{\boldsymbol{p}}_0(z)$, this then implies that $\psi(z) \leq \overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z)$, as desired. As a result, σ^* is indeed a ψ-quasi-perfect scheme satisfying (10). ■

Combining Lemma 1, Lemma 2 and Lemma 3, it then follows that there exists an incentive feasible ϕ_G-quasi-perfect mechanism and hence the data broker can attain revenue $\bar{R}$, proving the first part of Theorem 1 (under the regularity assumptions and (4)). In fact, even without the assumptions that G is regular and that (4) holds, as long as D_0 is regular, the proof above still implies the canonical $\overline{\varphi}_G$-quasi-perfect mechanism (σ^*, τ^*) defined by (3) and Lemma 1 is incentive feasible, which, together with Theorem 1, proofs Theorem 2.

Uniqueness

To see why any optimal mechanism of the data broker is ϕ_G-quasi-perfect, suppose that (σ, τ) is optimal. Then,

$$\begin{aligned}\bar{R} &= \int_C \left(\int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \bar{\pi} \\ &= \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{\pi}, \end{aligned} \tag{11}$$

which in turn implies that for (almost) all $c \in C$,

$$\int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) = \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c), \tag{12}$$

since the left-hand side is the efficient surplus in an economy where the producer's cost is $\phi_G(c)$ and hence must be an upper-bound of the right-hand side. (11) then implies that the right-hand side of (12) must attain this upper bound for (almost) all $c \in C$.

It then follows that σ must be a ϕ_G-quasi-perfect mechanism. Indeed, if σ is not a ϕ_G-quasi-perfect scheme, then there must be a positive G-measure of $c \in C$ and a positive $\sigma(c)$-measure of $D \in \mathrm{supp}(\sigma(c))$ such that either $D(v) > 0$ for some $v > \overline{\boldsymbol{p}}_D(c)$, or $D(\phi_G(c)) \neq D(\overline{\boldsymbol{p}}_D(c))$. That is, either there are some consumers with $v \geq \phi_G(c)$ who do not buy the product or buy the product at a price below v, or there are some consumers with $v < \phi_G(c)$ who end up buying the product. This contradicts (12). As a result, (σ, τ) must be a ϕ_G-quasi-perfect mechanism. Moreover, (σ, τ) must also induce quasi-perfect price discrimination since $\overline{\boldsymbol{p}}$ can be replaced with any $\boldsymbol{p} \in \boldsymbol{P}$ according to Lemma 1.

4.3 Further Remarks

Theorem 1 underlines a noteworthy feature of the optimal mechanisms. According to Theorem 1, for any optimal mechanism (σ, τ), the segmentation scheme σ does not generate value-revealing segmentations in general. Specifically, for any report c such that $\overline{\varphi}_G(c) > \underline{v}$, there are market segments $D \in \text{supp}(\sigma(c))$ containing consumers with distinct values. The reason is that in order to attain the desired upper bound, the data broker has to incentivize the producer to not sell to any consumers with values $v \in [c, \overline{\varphi}_G(c))$. Consumers with values below the desirable threshold $\overline{\varphi}_G(c)$ must be assigned to the same segment as some consumers with values above $\overline{\varphi}_G(c)$. By properly pooling the low-value consumers with the high-value ones while separating all the high-value consumers at the same time, the data broker is able to incentivize the producer to only sell to the consumers with the highest value in each market segment and induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for all c. This in turn enables the data broker to elicit private information by discouraging trade and extract surplus from the purchasing consumers at the same time.

As an example, consider the case where D_0 is linear and G is a uniform distribution with $V = C = [0, 1]$. It then follows $\overline{\varphi}_G(c) = 2c$ for all $c \in [0, 1/3]$ and $\overline{\varphi}_G(c) = (1 + c)/2$ for all $c \in (1/3, 1]$. In this case, the canonical $\overline{\varphi}_G$-quasi-perfect mechanism is described by a uniform distribution across market segments $\{D_v^{\overline{\varphi}_G(c)}\}_{v\in[\overline{\varphi}_G(c),1]}$, where each market segment $D_v^{\overline{\varphi}_G(c)}$ is defined by (2). As another example, notice that the optimal menu $\mathcal{M}^*$ in Section 2, which consists of the value-revealing data (with a price of 7/12) and the residential data (with a price of 1/3), implements the canonical quasi-perfect mechanism with a desirable cutoff function. Indeed, the residential data induces a 2-quasi-perfect segmentation for $c = 3/4$ as it only separates the high-value consumers (graduate and faculty) and pools the low-value consumers (undergraduate) with them uniformly. Meanwhile, the value-revealing data induces a 1-quasi-perfect segmentation for $c = 1/4$. According to the characterization above, since market demand D_0 is regular and since the virtual costs are 1/4 and 5/4 (for costs 1/4 and 3/4, respectively),[19] the menu $\mathcal{M}^*$ is indeed optimal.

Finally, recall that the upper bound $\bar{R}$ is derived by (i) ignoring the global incentive constraints (i.e., condition 2 of Lemma 1); (ii) compelling the producer to charge prices that are optimal with respect to the virtual cost $\phi_G(c)$, as opposed to her true cost c; and (iii) ignoring the individual rationality constraints. As shown above, under (4) and regularity assumptions for both D_0 and G, all three constraints end up being not binding under the optimal mechanism (σ^*, τ^*). While it is a general feature that (i) and (ii) do not bind even without these simplifying assumptions, the mechanism constructed above might violate the individual rationality constraints (iii) when (4) fails. Therefore, another (tighter) upper bound needs to be considered when extending the arguments above to the case when (4) does not necessarily hold, which will be discussed at the end of Section 5.

[19]Although the characterization is stated for cost distributions that admit densities, as in standard mechanism design problems, there is a straightforward analogous notion of virtual cost function when the cost distribution has atoms.

5 Consequences of Consumer-Data Brokership

5.1 Surplus Extraction

One of the most pertinent questions about consumer-data brokership is how it affects consumer surplus. Are the data broker's possession of consumer data and the ability to sell them to a producer detrimental for the consumers? If so, to what extent? Meanwhile, can the consumers benefit from the fact that the data broker does not have retail access to the consumers and only affects the product market indirectly by selling data to the producer? The following result, as an implication of Theorem 1, answers a certain aspect of this question.

Theorem 3 (Surplus Extraction)**.** *Consumer surplus is zero under any optimal mechanism.*

Theorem 3 follows directly from the characterization given by Theorem 1. According to Theorem 1, any optimal mechanism must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for (almost) all $c \in C$, which means that every purchasing consumer must be paying their values. Notably, Theorem 3 provides an unambiguous assertion about the consumer surplus under data brokership. According to Theorem 3, even though the data broker does not sell the product to the consumers directly and only affects the market by creating market segmentations for the producer, it is as if the consumers are perfectly price discriminated and all the surplus is extracted away (even though the optimal mechanisms do not perfectly reveal consumers' values in general). This means that the consumers do not benefit from the gap between the ownership of production technology and ownership of consumer data.

5.2 Comparisons with Uniform Pricing

Although Theorem 3 indicates data brokership is undesirable for the consumers, it does not imply that data brokership is detrimental to the entire economy. After all, by facilitating price discrimination, data brokership may increase total surplus compered with uniform pricing where no information about the consumers' values is revealed. Theorem 1, together with Proposition 1, allows for such a comparison.

Proposition 1. *The data broker's optimal revenue is no less than the consumer surplus under uniform pricing.*

An immediate consequence of Proposition 1 is that total surplus under data brokership is greater compared with uniform pricing, as summarized below.

Theorem 4 (Total Surplus Improvement)**.** *Data brokership always increases total surplus compared with uniform pricing.*

In other words, in terms of total surplus, data brokership is always better than the environment where no information about the consumers' values can be disclosed, even though data brokership is harmful to the consumers.

Another implication of Proposition 1 pertains to the source of consumer data. So far, it has been assumed that the data broker owns all the consumer data and is able to perfectly predict each consumer's value. In contrast, a different ownership structure of consumer data can be considered. In this alternative setting, the data broker does not have any data in the first place and has to purchase them from the consumers.[20] Proposition 1 immediately implies that, if the data broker has to purchase data by compensating the consumers with monetary transfers *before* they learn their values,[21] then the optimal mechanism would be to purchase all the data by paying the consumers their ex-ante surplus under uniform pricing and then use any optimal mechanism characterized by Theorem 1 to sell these data to the producer. Furthermore, since the data broker's revenue is greater than the consumer surplus under uniform pricing according to Proposition 1, and since the producer always has an outside option of uniform pricing, this outcome is in fact Pareto improving compared with uniform pricing in the ex-ante sense, as stated below.[22]

Theorem 5 (Data Ownership)**.** *If the data broker has to purchase data from the consumers and if such purchase occurs before consumers learn their values, then data brokership is Pareto improving compared with uniform pricing in the ex-ante sense.*

[20]For simplicity, a "purchase" of data here means that the data broker gains access to *all* the consumer data, in the sense that he can provide any segmentation of D_0 to the producer once he makes the purchase. In an earlier version of this paper (Yang, 2020c), I further extend the model and allow the data broker to make a take-it-or-leave-it offer to purchase *any* kind of consumer data and then sell them to the producer. (i.e., offer any segmentation of D_0 that is a mean-preserving contraction of the segmentation induced by the purchased data.)

[21]It is crucial here the data broker purchases *before* the consumers learn their value, since otherwise he would also have to screen the consumers to elicit their private information. This assumption is plausibly suitable for online activities. After all, in online settings, consumers often do not consider their values about a particular product when they agree that their personal data such as browsing histories, IP address and cookies, can be collected by the data brokers. Nevertheless, other purchase timing would also be a relevant question, which can be explored in future research.

[22]Jones and Tonetti (2020) also conclude that granting consumers ownership of their own data is welfare-improving. However, their results are derived in a monopolistic competition setting and the main driving force is the non-rival property of data, whereas Theorem 5 is derived under a monopoly setting and the main rationale is that consumer data facilitate price discrimination, which in turn increases sales and thus enhances efficiency.

5.3 Comparisons across Market Regimes

In addition to its welfare implications, the characterization of Theorem 1 provides further insights about the comparisons across different regimes of the market. Indeed, other than selling consumer data to the producer, there are several other market regimes under which the data broker can profit from the consumer data he owns. Therefore, it would be policy-relevant to compare the outcomes induced by these different market regimes. In what follows, I introduce several market regimes in addition to data brokership, including **vertical integration**, **exclusive retail**, and **price-controlling data brokership**. I then compare the implications among these different regimes using the characterization provided by Theorem 1.

Vertical Integration— The producer's marginal cost of production becomes common knowledge (for exogenous reasons such as regulation or technological improvements) and the data broker vertically integrates with the producer. That is, the vertically integrated entity is able to produce the product and sell to the consumers via perfect price discrimination.

Exclusive Retail— The producer's marginal cost of production remains private. The data broker negotiates with the producer to purchase the product and the exclusive right to sell the product. Specifically, the data broker can offer a menu, where each item in this menu specifies the quantity $q \in [0, 1]$ that the producer has to produce and supply to the data broker, as well as the amount of payment $t \in \mathbb{R}$ the data broker has to pay to the producer. If the producer chooses an item (q, t) from this menu, the producer receives profit $t - cq$ while the data broker pays t and can sell at most q units exclusively to the consumers through any market segmentation. If the producer rejects this menu, she retains her optimal uniform profit and the data broker receives zero.

Price-Controlling Data Brokership— The producer's marginal cost of production is private information. The data broker, in addition to being able to create market segmentations and sell them to the producer, can further specify what price should be charged in each market segment as a part of the contract. If the producer rejects, she retains her optimal uniform pricing profit and the data broker receives zero. Specifically, the data broker offers a mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ such that for all $c, c' \in C$,

$$\int_{\mathcal{D}\times\mathbb{R}_+} (p-c)D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\sigma(\mathrm{d}D|c) - \tau(c) \geq \int_{\mathcal{D}\times\mathbb{R}_+} (p-c)D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c')\sigma(\mathrm{d}D|c') - \tau(c')$$

and for all $c \in C$,

$$\int_{\mathcal{D}\times\mathbb{R}_+} (p-c)D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\sigma(\mathrm{d}D|c) - \tau(c) \geq \pi_{D_0}(c),$$

where for each $c \in C$, $\sigma(c) \in \mathcal{S}$ is the market segmentation provided to the producer, $\tau(c) \in \mathbb{R}$ is the payment from the producer to the data broker, and $\boldsymbol{\gamma}(c) : \mathcal{D} \to \Delta(\mathbb{R}_+)$ is a transition kernel so that $\boldsymbol{\gamma}(\cdot|D,c)$ specifies the distribution from which prices charged in segment D must be drawn.

With these definitions, for each market regime, there is an associated profit maximization problem. Henceforth, two market regimes are said to be *outcome-equivalent* if every solution of the profit maximization problems associated with either market regime induces the same market outcome (i.e., consumer surplus, producer's profit, data broker's revenue and the allocation of the product).

An immediate consequence of Theorem 1 is the comparison between data brokership and vertical integration. To see this, recall that any optimal mechanism (σ,τ) of the data broker must induce $\overline{\varphi}_G$-quasi-perfect price discrimination but not perfect price discrimination in general, as $\overline{\varphi}_G(c) > c$ for all $c > \underline{c}$. Thus, whenever there are some consumers with values between c and $\overline{\varphi}_G(c)$ for a positive measure of c, no optimal mechanism would lead to an efficient allocation, because there would be some consumers who end up not buying the product even though their values are greater than the marginal cost. Together with Theorem 3, this means that vertical integration between the data broker and producer strictly increases total surplus while leaving the consumer surplus unchanged when $\mathrm{supp}(D_0) = V$ and when there is no common knowledge of gains from trade. After all, consumer surplus is always zero under both regimes, whereas the integrated entity after vertical integration does not create any friction and would perfectly price discriminate the consumers whose values are above the marginal cost.

Theorem 6 (Vertical Integration)**.** *Compared with data brokership, vertical integration strictly increases total surplus and leaves the consumer surplus unchanged if D_0 is strictly decreasing and $\underline{v} < \overline{c}$.*

For other market regimes, it is noteworthy that since prices are contractable under price-controlling data brokership, for any mechanism $(\sigma,\tau,\boldsymbol{\gamma})$, the producer's private marginal cost affects her profit only through the quantity produced and sold to the consumers induced by $(\sigma,\boldsymbol{\gamma})$. This effectively reduces allocation space under price-controlling data brokership to a one-dimensional quantity space, which is the same as the allocation space under exclusive retail. In fact, as stated in Lemma 4 below, price-controlling data brokership is always equivalent to exclusive retail.

Lemma 4. *Exclusive retail and price-controlling data brokership are outcome-equivalent.*

With Lemma 4, to compare exclusive retail and price-controlling data brokership with data brokership, it suffices to compare only price-controlling data brokership with data brokership. This comparison is particularly convenient since the price-controlling data broker's revenue maximization problem is a relaxation of the data broker's. After all, with the extra ability to contract on prices, the constraints in the price-controlling data broker's problem are clearly weaker. Nevertheless, as an implication of Theorem 1 and Proposition 2 below, it turns out that the data broker's optimal revenue is in fact the same as the price-controlling data broker's optimal revenue.

Proposition 2. *Any optimal mechanism of the price-controlling data broker induces $\bar{\varphi}_G(c)$-quasi-perfect price discrimination for G-almost all $c \in C$. In particular, the optimal revenue is*

$$R^* = \int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \bar{\pi}.$$

According to Theorem 1 and Lemma 1, the optimal revenue of the data broker must also be R^*. This means that the additional ability to control prices does not benefit the data broker at all. In fact, as stated by Theorem 7 below, this ability is entirely irrelevant in terms of market outcomes.

Theorem 7 (Outcome-Equivalence)**.** *Exclusive retail, price-controlling data brokership and data brokership are outcome-equivalent.*

In other words, Theorem 7 means that even though the data broker only affects the product market indirectly by selling consumer data, the market outcomes he induces are the same as those when he has more control over the product market (by either becoming a price-controlling data broker or an exclusive retailer). More specifically, from the data broker's perspective, having control over how the product is sold in addition to consumer data adds no extra value to his revenue. As for the producer, preserving the retail access to consumers and the right to sell the product is in fact not more profitable. In addition, the allocation of the product induced by a data broker is the same as that induced by an exclusive retailer. Therefore, the channel through which the product is sold to the consumers does not affect the amount of products being produced, nor does it affect to whom the product is sold. Overall, Theorem 7 provides a way to gauge how powerful the ability to design and sell market segmentations is: According to Theorem 7, this ability is so powerful that being able to further contract on outcomes in the product market provides no additional value to the data broker.

As another remark, the fact that the price-controlling data broker's optimal revenue R^* is an upper bound for the data broker's optimal revenue completes the intuition behind the proof of Theorem 1 without the additional assumption (4) imposed in Section 4.2. To see this, since the price-controlling data broker's optimal mechanisms always induce $\overline{\varphi}_G$-quasi-perfect price discrimination for (almost) all $c \in C$ according to Proposition 2, proving Theorem 1 is essentially reduced to finding an incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism. Meanwhile, by the definition of $\overline{\varphi}_G$, $c \leq \overline{\varphi}_G(c) \leq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$, and hence $\overline{\varphi}_G$ satisfies the condition required by Lemma 3. As a result, combining Lemma 2 and Lemma 3, there is indeed an incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism, which, by definition, generates revenue R^*, and hence is optimal. As noted at the end of the previous section, while ϕ_G-quasi perfect mechanisms may not be individual rational when (4) fails, $\overline{\varphi}_G$-quasi-perfect mechanisms implied by Lemma 3 are indeed individually rational. In fact, the reason the price controlling data broker's revenue R^* (as opposed to $\bar{R}$ in Section 4) becomes the correct upper bound when (4) does not hold is precisely because some individual rationality constraints may be binding under the price-controlling data broker's optimal mechanism (see more discussions in Section 7).

6 Extension: Restricted Market Segmentations

Thus far, it has been assumed that the data broker is able to create *any* market segmentation, including the value-revealing segmentation that perfectly discloses consumers' values. Although it is not implausible—given the advancement of information technology—that a data broker is (or at least will soon be) able to almost perfectly predict consumers' values, it is still crucial to explore the economic implications when the data broker does not have perfect information about consumers' values. This section extends the baseline model in Section 3 and restricts the data broker's ability in creating market segmentations.

To model this restriction, let Θ be a finite set of consumer characteristics that can be disclosed by the data broker. Suppose that among the consumers, their characteristics $\theta \in \Theta$ are distributed according to $\beta_0 \in \Delta(\Theta)$. These characteristics are informative of the consumers' values but there may still be variations in values among the consumers who share the same characteristics. Specifically, given any $\theta \in \Theta$, suppose that among the consumers who share characteristic θ, their values are distributed according a demand $D_\theta \in \mathcal{D}$ (i.e., $D_\theta(p)$ denotes the share of consumers with values above p among those with characteristic θ). Moreover, suppose that $\{\mathrm{supp}(D_\theta)\}_{\theta\in\Theta}$ forms a partition of V and that $\mathrm{supp}(D_\theta)$ is an interval for all $\theta \in \Theta$. In other words, the available consumer characteristics is only partially informative of the consumers' values in a way that any particular characteristic can only identify which interval a particular consumer's value belongs to. As a result, even when θ is perfectly revealed, the producer would still be unable to perfectly identify each consumer's value. For any $p \in V$, let

$$D_0(p) := \sum_{\theta\in\Theta} D_\theta(p)\beta_0(\theta).$$

$D_0 \in \mathcal{D}$ then describes the market demand in this environment.

In this environment, a market segmentation is defined by $s \in \Delta(\Delta(\Theta))$ such that

$$\int_{\Delta(\Theta)} \beta(\theta)s(\mathrm{d}\beta) = \beta_0(\theta),$$

for all $\theta \in \Theta$. A market segmentation s induces market segments $\{D_\beta\}_{\beta \in \text{supp}(s)}$ and

$$\sum_{\beta \in \text{supp}(s)} D_\beta(p)s(\beta) = D_0(p),$$

for all $p \in V$, where $D_\beta(p) := \sum_{\theta \in \Theta} D_\theta(p)\beta(\theta)$ for any $\beta \in \Delta(\Theta)$ and any $p \in V$.

When the consumers' values can never be fully disclosed, it is clear that their surplus will increase. After all, it is no longer possible for the producer to charge the consumers their values as the additional variation in values given by D_θ always allows some consumers to buy the product at a price below their values. Nevertheless, as shown in Theorem 8, under any optimal mechanism, consumer surplus must be lower than the case when all the information about θ is revealed to the producer. That is, the main implication of Theorem 3— for the consumers, the presence of a data broker is no better than a scenario where their data is fully revealed to the producer—is still valid even when the consumers retain some private information.

Theorem 8. *For any $(\{D_\theta\}_{\theta \in \Theta}, \beta_0)$ and for any cost distribution G, an optimal mechanism always exists. Furthermore, the consumer surplus under any optimal mechanism of the data broker is lower than the case when θ is fully disclosed.*

The intuition behind Theorem 8 is simple. Since there are only finitely many characteristics and since $\{\text{supp}(D_\theta)\}_{\theta \in \Theta}$ forms a partition of V, identifying the consumers' characteristic θ effectively enables the producer to categorize the consumers into finitely many "blocks" so that every possible value belongs to one and only one block. As a result, when changing prices within each block of values, the trading volume is only affected by purchasing decisions of the consumers whose values are within that block. Such separability allows the data broker to always construct a mechanism that (strictly) increases its revenue if the consumer surplus is higher than that when the characteristic θ is fully-revealed.[23]

In addition to the surplus extraction result, the characterization of the optimal mechanisms can be generalized as well. With proper regularity conditions, there is an optimal mechanism analogous to the canonical $\overline{\varphi}_G$-quasi-perfect mechanism introduced in Section 4. To state this result, given any $(\{D_\theta\}_{\theta \in \Theta}, \beta_0)$, for each $\theta \in \Theta$, write $\text{supp}(D_\theta)$ as $[l(\theta), u(\theta)]$. For any $p \in V$, let $\theta_p \in \Theta$ be the unique θ such that $p \in (l(\theta), u(\theta)]$. For any $c \in C$, let $\widehat{\boldsymbol{p}}_0(c)$ be the largest optimal price for the producer with marginal cost $c \in C$ under the demand whose support contains $\overline{\boldsymbol{p}}_0(c)$.[24] Also, let $\widehat{\varphi}_G(c) := \min\{\varphi_G(c), \widehat{\boldsymbol{p}}_0(c)\}$ for all $c \in C$. Furthermore, given any function $\psi : C \to \mathbb{R}_+$, say that a mechanism (σ, τ) is a canonical

[23] A more detailed argument can be found in the proof, which is provided in the Online Appendix

[24] That is, $\widehat{\boldsymbol{p}}_0(c) := \overline{\boldsymbol{p}}_{D_{\theta_{\overline{\boldsymbol{p}}_0(c)}}}(c)$. Notice that $\widehat{\boldsymbol{p}}_0(c) \leq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$. Moreover, in the case where the data broker can disclose all the information about the value v, $\widehat{\boldsymbol{p}}_0(c) = \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$.

ψ-quasi-perfect segmentation if the producer with marginal cost $\bar{c}$, when reporting truthfully, recevies $\bar{\pi}$, and if for any $c \in C$, and for any $\beta \in \text{supp}(\sigma(c))$, either

$$\beta(\theta') = \beta^{\theta}_{\psi(c)}(\theta') := \begin{cases} \beta_0(\theta'), & \text{if } u(\theta') < \psi(c) \text{ and } u(\theta) \geq \psi(c) \\ \sum_{\{\hat{\theta}: u(\hat{\theta}) \geq \psi(c)\}} \beta_0(\hat{\theta}), & \text{if } u(\theta') \geq \psi(c) \text{ and } \theta' = \theta \\ 0, & \text{otherwise} \end{cases}, \tag{13}$$

for any $\theta, \theta' \in \Theta$; or

$$\text{supp}(\beta) = \{\theta' : l(\theta') \leq \psi(c)\} \cup \{\theta\} \tag{14}$$

for some $\theta \in \Theta$ with $l(\theta) \geq \psi(c)$ and

$$\beta(\theta') = \beta_0(\theta'). \tag{15}$$

for all $\theta' \in \Theta$ such that $u(\theta') < \psi(c)$.

With these definitions, Theorem 9 below prescribes an optimal mechanism for the data broker.

Theorem 9. *For any $(\{D_\theta\}_{\theta \in \Theta}, \beta_0)$ and any distribution of marginal cost G such that the function $c \mapsto \max\{(\phi_G(c) - \widehat{\boldsymbol{p}}_0(c)), 0\}$ is nondecreasing and that D_0 is regular, there is a canonical $\widehat{\varphi}_G$-quasi-perfect mechanism that is optimal.*

7 Discussions

7.1 Sufficient Conditions and Relaxations of Assumption 1

As noted in Section 4, Assumption 1 has a sufficient condition (4). To better understand (4), recall that $\phi_G(c)$ is the actual marginal cost c plus the information rent $G(c)/g(c)$. Meanwhile, $\overline{\boldsymbol{p}}_0(c)$ can be written as $\overline{\boldsymbol{p}}_0(c) = c + \xi_0(c)$, where $\xi_0(c) := \overline{\boldsymbol{p}}_0(c) - c$ is the *monopoly mark-up* that the producer charges under uniform pricing. From this perspective, (4) is equivalent to $G(c)/g(c) \leq \xi_0(c)$, for all $c \in C$. That is, the information rent that the producer retains due to asymmetric information about her marginal cost is less than her monopoly mark-up. Furthermore, since (4) means that the optimal uniform price must be greater than the virtual cost, (4) also can be interpreted as that the gains from trade are large enough.[25]

Although the results derived above rely on Assumption 1, the main purpose of Assumption 1 is to ensure that as a revenue upper bound, the price-controlling data broker's problem has a closed form solution. After all, by Lemma 4, the price-controlling data broker's problem is essentially a nonlinear screening problem with one-dimensional allocation space and type-dependent outside options. A common feature of such problem is that the characterization

[25] A formal argument can be found in an earlier version of this paper (Yang, 2020c), where gains from trade are measured by a demand shifter that moves the market demand to the right on the real line.

of the optimal mechanisms involves Lagrange multipliers in general (see, for instance, Lewis and Sappington (1989) and Jullien (2000)). Assumption 1, however, yields a closed form solution for the price-controlling data broker's problem (Proposition 2), which in turn allows an explicit construction of an incentive feasible mechanism for the data broker that attains the revenue upper bound. Consequently, many of the results, including the main characterization, the surplus extraction result and the associated implications can be extended to environments without Assumption 1.[26]

7.2 Creating Market Segmentations by Partitioning underlying Characteristics

Throughout the paper, a market segmentation is formalized as a probability measure $s \in \mathcal{S}$ that splits the market demand D_0 into several segments $D \in \mathcal{D}$, which aligns with the literature of price discrimination. However, a more practical way to describe a market segmentation—especially in environments where segmentations are generated by consumer data—is to define it as a partition on a set of consumers' characteristics that are correlated with their values of a product.

Clearly, with too few underlying characteristics, the ways to split the market demand would be limited. For instance, in the motivating example, if the only available characteristic is the residence type, then the market demand can only be split in the way described by Figure 2. For the data broker to be able to design *any* market segmentation, it is implicitly required that the underlying characteristics should be "rich enough" (i.e., the data broker has a large enough dataset). In a companion note (Yang, 2020a), I formalize this observation, which guarantees that the data broker can generate any market segmentation $s \in \mathcal{S}$ by partitioning an underlying characteristic space, provided that it is "rich enough". From this perspective, while how the data broker should sell consumer data when there is only a limited set of available characteristics remains an open question, the results in this paper can be regarded as what the data broker can possibly achieve when he has an access to sufficiently large datasets.

7.3 Source of Asymmetric Information

The results in previous sections are derived under an information structure where the producer has private information about her marginal cost. Although this informational assump-

[26]In an earlier version of this paper (Yang, 2020c), I provide a generalized version of Theorem 1 when D_0 is continuous. Specifically, I show that there exists a nondecreasing function φ^* (may not necessarily be of a closed form) such that every optimal mechanism must be a φ^*-quasi-perfect mechanism. Furthermore, I prove a strengthened version of Theorem 3, which does not rely on any assumptions about D_0 and G and ensures both the existence of an optimal mechanism, as well as the fact that any optimal mechanism must yield zero consumer surplus.

tion captures certain the features in retail markets, it apparently does not capture all of them. Specifically, one salient informational asymmetry between a data broker and a producer in the real world is that producers often know more about how consumers' characteristics are related to their values for a particular product—perhaps due to their industry-specific knowledge that is too costly for the data broker to acquire. While optimal selling mechanisms for the data broker under this more general environment remain an open question, the methodology developed in this paper can still provide some insights. In particular, under a parameterized information structure where the producer has private information about the market condition (as opposed to her marginal cost), all the results derived in this paper continue to hold.

Consider the following alternative information structure. There is a unit mass of consumers with unit demand for a single product. Each consumer has value $v - \xi$, where $v \in [\underline{v}, \bar{v}] = V \subseteq \mathbb{R}_+$ is heterogeneous across consumers and distributed according to $D_0 \in \mathcal{D}$, while $\xi \in [0, \underline{v}]$ is the same across consumers. All the consumers and the producer (with a commonly known marginal cost that is normalized to zero) know ξ, while the data broker only knows that ξ is drawn from a distribution G. The interpretation is that the producer knows more about the market condition (i.e., a "demand shifter" described by ξ) than the data broker does. In this setting, market segmentations are defined as before: A market segmentation is a probability measure $s \in \mathcal{S} \subseteq \Delta(\mathcal{D})$. It then follows that under market condition ξ, the demand in a market segment $D \in \mathcal{D}$ at price p is given by $D(p + \xi)$ (i.e., $D(p+\xi)$ is the share of consumers in segment D who are willing to buy the product at price p when the market condition is ξ). Thus, given a demand shifter ξ, under any market segment $D \in \mathcal{D}$, the producer's pricing problem is given by

$$\max_{p \geq 0} pD(p + \xi),$$

which, by letting $p' = p + \xi$, is equivalent to

$$\max_{p' \geq 0} (p' - \xi)D(p') = \pi_D(\xi).$$

As a result, the model above where the producer privately knows a demand shifter is equivalent to the original model where the producer has a private marginal cost ξ, and hence all the results derived above continue to hold in this alternative setting.

7.4 Policy Implications

The results above have several broader policy implications. First, in terms of welfare, although Theorem 3 implies that data brokership is undesirable for the consumers, Theorem 4 shows that the total surplus is always higher in the presence of a data broker compared with an environment where no information about the consumers' values can be disclosed. As a result, the answer to whether a data broker is beneficial must depend on the objective of the policymaker and the kinds of redistributional policy tools available. If the policymaker's objective is to simply maximize total surplus, or if redistributional tools such as lump-sum transfers are available, then it is indeed beneficial to allow a data broker to sell consumer data. By contrast, if the policymaker is additionally concerned with consumer surplus, and if no effective redistributional policies are accessible, then the presence of a data broker can be fairly unfavorable. However, Theorem 5 prescribes a potential way to improve welfare: If the data broker had to purchase the data from the consumers, and if the purchase took place before the consumers learn their values, then data brokership would be Pareto-improving compared with uniform pricing. As a result, if the policymaker can establish the consumers' property right of their own data,[27] as well as a channel for the data broker to compensate the consumers, then not only the consumers can secure their surplus as if their data is not used for price discrimination (via compensation), but also the entire economy can benefit from data brokership, because less deadweight loss will be generated.

Furthermore, the discussions in Section 5.3 facilitates the evaluation of whether a certain market regime is desirable than another. According to Theorem 6, it can be beneficial when the policymaker reveals the producer's private marginal cost and encourages vertical integration, as all the informational frictions would be eliminated without affecting the consumer surplus. Meanwhile, the equivalence result given by Theorem 7 implies that as long as the producer bears the production cost, however active the data broker is in the product market does not affect market outcomes at all. On the one hand, this means that the data broker has no incentive to become more active in the product market in addition to selling consumer data. In fact, together with other potential costs that are abstracted away from the model (e.g., inventory costs, shipping costs and other transaction costs), participating directly in product market can be less profitable than merely selling consumer data to the producer. On the other hand, this implies that even if the data broker does become more active, it raises no further concerns to the policymaker. Thus, any policy intervention that prohibits the data broker entering the product market by either gaining control over prices (e.g., by establishing an online platform and allows the producer to trade on this platform while controlling the prices) or obtaining the exclusive right to (re)-sell the product would be unnecessary. However, another interpretation of this result is that even if the data broker is not active in the product market at all, the policymaker should be equally concerned as if the data broker were very active.

[27] For instance, just as what is stipulated by the recent regulation of the European Union, General Data Protection Regulation (GDPR, Art. 7), consumers' property right for their own data can be better protected by prohibiting all the processing of personal data unless the data subject has consented the use.

8 Conclusion

In this paper, I consider a model where a data broker sells consumer data and creates market segmentations and characterize the optimal mechanisms of the data broker. I conclude that consumer surplus is always zero, that data brokership generates more total surplus than uniform pricing, and that the ability to control prices in the product market is irrelevant. I also study an extension where the data broker can only create a limited set of market segmentations and find qualitatively similar results.

Several topics remain to be explored by future studies. First, although private information about a demand shifter is equivalent to private information about marginal cost, a model with more general specifications of the producer's private information on how consumer data can be used to predict their values is worth exploring. Second, while the extension considers the case where the data broker can only create a limited set of market segmentations, it is restricted to the partitional environment introduced in Section 6. A natural direction is then to study a setting where the feasible market segmentation is restricted by an arbitrary Blackwell upper bound. Lastly, while both the data broker and the producer are assumed to be monopolists in this paper, it would be economically relevant to explore the consequences of consumer-data brokership under different market structures.

Appendix

A Details of $\mathcal{D}$

Below I first discuss more formally about the properties of the set $\mathcal{D}$. Recall that $\mathcal{D} = \mathcal{D}([\underline{v}, \overline{v}])$ is the collection of nonincreasing and left-continuous functions D on $[\underline{v}, \overline{v}]$ such that $D(\underline{v}) = 1$ and $D(\overline{v}^+) = 0$. Since for every $D \in \mathcal{D}$, there exists a unique probability measure $m^D \in \Delta(V)$ such that $D(p) = m^D(\{v \geq p\})$ for all $p \in V$, I define the topology on $\mathcal{D}$ by the following notion of convergence: For any $\{D_n\} \subseteq \mathcal{D}$ and any $D \in \mathcal{D}$, $\{D_n\} \to D$ if and only if for any bounded continuous function $f : V \to \mathbb{R}$,

$$\lim_{n \to \infty} \int_V f(v) m^{D_n}(\mathrm{d}v) = \int_V f(v) m^D(\mathrm{d}v).$$

This corresponds to the weak-* topology on $\Delta(V)$ and hence this topology on $\mathcal{D}$ is also called the weak-* topology. As a result, $\mathcal{D}$ is a Polish space. Furthermore, notice that under this topology, $\{D_n\} \to D$ if and only if $\{D_n(p)\} \to D(p)$ for all $p \in V$ at which D is continuous. Finally, for any $D \in \mathcal{D}$, let $\mathcal{S}_D$ denote the collection of $s \in \Delta(\mathcal{D})$ such that (1) holds with D_0 replaced by D (so that $\mathcal{S}_{D_0} = \mathcal{S}$). Also, let D^{-1} denote the inverse demand of D. That is,

$$D^{-1}(q) := \sup\{p \in V : D(p) \geq q\}, \; \forall q \in [0, 1]. \tag{16}$$

B Proofs for Optimal Mechanisms

This section contains proofs of the main results regarding the optimal mechanisms (i.e., Theorem 1 and Theorem 2). To this end, I first solve for the price-controlling data broker's optimal mechanism (Proposition 2) and use this as an upper bound for the data broker's revenue. I then construct an incentive feasible mechanism for the data broker that attains this bound and establish uniqueness (Theorem 1)

B.1 Crucial Properties of Quasi-Perfect Schemes

The following lemma summarizes some crucial properties of quasi-perfect segmentation schemes. The proofs of these properties are mostly technical and are not directly related to the arguments of the proofs of main results, and therefore are relegated to the Online Appendix.

Lemma 5. *Consider any nondecreasing function $\psi : C \to \mathbb{R}_+$ with $c \leq \psi(c)$ for all $c \in C$. Suppose that for any $c \in C$, $\sigma(c) \in \mathcal{S}$ is a $\psi(c)$-quasi-perfect segmentation for c. Then,*

1. $\int_{\mathcal{D}} D(p) \sigma(\mathrm{d}D|c) = D_0(p)$ *for all* $p \in V$ *and for all* $c \in C$.
2. $\sigma : C \to \Delta(\mathcal{D})$ *is measurable.*
3. $\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) = D_0(\psi(c))$ *for all* $c \in C$.
4. $\int_{\mathcal{D}} \overline{\boldsymbol{p}}_D(c) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) = \int_{\{v \geq \psi(c)\}} v D_0(\mathrm{d}v)$ *for all* $c \in C$.

B.2 Proof of Proposition 2

To solve for the price-controlling data broker's optimal mechanism, it is useful to introduce the revenue-equivalence formula for the price-controlling data broker.

Lemma 6. *For the price-controlling data broker, a mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ is incentive compatible if and only if:*

1. *There exists some $\bar{\tau} \in \mathbb{R}$ such that for any $c \in C$,*

$$\tau(c) = \int_{\mathcal{D}} \int_{\mathbb{R}_+} (p-c) D(p) \boldsymbol{\gamma}(\mathrm{d}p|D,c)\sigma(\mathrm{d}D|c) - \int_c^{\bar{c}} \int_{\mathcal{D}} \int_{\mathbb{R}_+} D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,z)\sigma(\mathrm{d}D|z)\,\mathrm{d}z - \bar{\tau}.$$

2. *The function $c \mapsto \int_{\mathcal{D}} \int_{\mathbb{R}_+} D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\sigma(\mathrm{d}D|c)$ is nonincreasing.*

The proof of Lemma 6 follows directly from the standard envelope arguments and therefore is omitted. In addition to Lemma 6, since both prices and market segmentations can be contracted by the price-controlling data broker, and since the producer's private information is one dimensional, the price controlling data broker's problem can effectively be summarized by a one dimensional screening problem where the data broker contracts on quantity (sold via perfect price discrimination), as stated in Lemma 7 below.

Lemma 7. *There exists an incentive feasible mechanism that maximizes the price-controlling data broker's revenue. Furthermore, the price-controlling data broker's revenue maximization problem is equivalent to the following:*

$$\sup_{\boldsymbol{q} \in \mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \phi_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) - \bar{\pi} \tag{17}$$
$$s.t. \;\; \bar{\pi} + \int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z \geq \bar{\pi} + \int_c^{\bar{c}} D_0(\boldsymbol{p}_0(z))\,\mathrm{d}z,$$

where $\mathcal{Q}$ is the collection of nonincreasing functions that map from C to $[0,1]$.

The proof of Lemma 7 can be found in the Online Appendix. Essentially, the argument is to summarize σ and $\boldsymbol{\gamma}$ by

$$\boldsymbol{q}(c) = \int_{\mathcal{D} \times \mathbb{R}_+} D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\sigma(\mathrm{d}D|c),$$

for all $c \in C$. As the producer's private information is one dimensional, it turns out that it is sufficient for the price-controlling data broker to design quantity $\boldsymbol{q}$ and then prescribe perfect price discrimination subject to a capacity constraint $\boldsymbol{q}(c)$, for all $c \in C$. By the revenue equivalence formula (Lemma 6), the objective function of (17) equals to the broker's expected revenue given $\boldsymbol{q}$; the monotonicity condition $\boldsymbol{q} \in \mathcal{Q}$ corresponds to global incentive compatibility constraints; and the inequality constraints in (17) are equivalent to the individual rationality constraints.

With Lemma 7, the price-controlling data broker's revenue maximization problem can be solved explicitly.

Proof of Proposition 2. Let R^* be the value of (17) and consider the dual problem of (17). By weak duality, it suffices to find a Borel measure μ^* on C and a feasible $\boldsymbol{q}^* \in \mathcal{Q}$ such that $\boldsymbol{q}^*$ is a solution of

$$\sup_{\boldsymbol{q}\in\mathcal{Q}} \left[\int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \phi_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) - \bar{\pi} + \int_C \left(\int_c^{\bar{c}} (\boldsymbol{q}(z) - D_0(\overline{\boldsymbol{p}}_0(z)))\,\mathrm{d}z \right) \mu^*(\mathrm{d}c) \right] \tag{18}$$

and that

$$\int_C \left(\int_c^{\bar{c}} (\boldsymbol{q}^*(z) - D_0(\overline{\boldsymbol{p}}_0(z)))\,\mathrm{d}z \right) \mu^*(\mathrm{d}c) = 0. \tag{19}$$

To this end, define $M^* : C \to [0,1]$ as the following:

$$M^*(c) := \lim_{z \downarrow c} g(z)(\phi_G(z) - \overline{\boldsymbol{p}}_0(z))^+, \; \forall c \in C. \tag{20}$$

By definition, M^* is right-continuous. Also, by Assumption 1, M^* is nondecreasing and hence M^* a CDF. Let μ^* be the Borel measure induced by M^*. Notice that $\operatorname{supp}(\mu^*) = [c^*, \bar{c}]$, where $c^* := \inf\{c \in C : \phi_G(c) > \overline{\boldsymbol{p}}_0(c)\}$.

For any $\boldsymbol{q} \in \mathcal{Q}$, by interchanging the order of integrals and then rearranging, (18) can be written as

$$\sup_{\boldsymbol{q}\in\mathcal{Q}} \left[\int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\phi}_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) - \bar{\pi} - \int_C M^*(c) D_0(\overline{\boldsymbol{p}}_0(c))\,\mathrm{d}c \right], \tag{21}$$

where $\overline{\phi}_G := \min\{\phi_G, \overline{\boldsymbol{p}}_0\}$.

To solve (21), let φ_G be the ironed virtual cost. That is, φ_G is defined by the following procedure: Let $h : [0,1] \to \mathbb{R}_+$ be defined as $h(q) := \phi_G(G^{-1}(q))$, and define $H : [0,1] \to \mathbb{R}_+$, $K : [0,1] \to \mathbb{R}_+$ as $H(q) := \int_0^q h(s)\,\mathrm{d}s$ and $K := \operatorname{conv}(H)$. Then, for every $q \in [0,1]$ let $k(q) := K'(q)$ and define φ_G as $\varphi_G(c) := k(G(c))$. Also, let $\overline{\varphi}_G := \min\{\varphi_G, \overline{\boldsymbol{p}}_0\}$. Now notice that for any $\boldsymbol{q} \in \mathcal{Q}$, and for any $c \in C$,

$$\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\phi}_G(c))\,\mathrm{d}q = \int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\varphi}_G(c))\,\mathrm{d}q + (\overline{\varphi}_G(c) - \overline{\phi}_G(c))\boldsymbol{q}(c). \tag{22}$$

Moreover, using integration by parts, since $K(0) = H(0)$ and $K(G(c^*)) = H(G(c^*))$ (by Assumption 1),

$$\int_C (\overline{\varphi}_G(c) - \overline{\phi}_G(c))\boldsymbol{q}(c) G(\mathrm{d}c) = \int_{\underline{c}}^{c^*} (\varphi_G(c) - \phi_G(c))\boldsymbol{q}(c) G(\mathrm{d}c) = -\int_{\underline{c}}^{c^*} (K(G(c)) - H(G(c)))\boldsymbol{q}(\mathrm{d}c) \leq 0, \tag{23}$$

where the first equality follows from the observation that $\overline{\phi}_G(c) = \phi_G(c)$ and $\overline{\varphi}_G(c) = \varphi_G(c)$ for all $c \leq c^*$, and $\overline{\varphi}_G(c) = \overline{\phi}_G(c) = \overline{\boldsymbol{p}}_0(c)$ for all $c > c^*$, which is due to Assumption 1, and the inequality follows from the fact that $K = \operatorname{conv}(H)$ and that $\boldsymbol{q}$ is nonincreasing for any $\boldsymbol{q} \in \mathcal{Q}$.

Meanwhile, notice that

$$\int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\varphi}_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) \leq \int_C \left(\int_0^{D_0(\overline{\varphi}_G(c))} (D_0^{-1}(q) - \overline{\varphi}_G(c))\,\mathrm{d}q \right), \; \forall \boldsymbol{q} \in \mathcal{Q}. \tag{24}$$

In addition, since $\overline{\varphi}_G(c) = \overline{\phi}_G(c) = \overline{\boldsymbol{p}}_0(c)$ for all $c \in (c^*, \bar{c}]$ and since $K(G(c)) < H(G(c))$ on an interval $[c_1, c_2] \subseteq [\underline{c}, c^*]$ if and only if $\overline{\varphi}_G$ is constant on that interval, which implies that $D_0 \circ \overline{\varphi}_G$ is constant on that interval, it must be that

$$\int_C (\overline{\varphi}_G(c) - \overline{\phi}_G(c)) D_0(\overline{\varphi}_G(c)) G(\mathrm{d}c) = -\int_{\underline{c}}^{c^*} (K(G(c)) - H(G(c))) D_0 \circ \overline{\varphi}_G(\mathrm{d}c) = 0. \tag{25}$$

Together with (22), and (23), (24) for any $\boldsymbol{q} \in \mathcal{Q}$,

$$\int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\phi}_G(c))\, \mathrm{d}q \right) G(\mathrm{d}c) \leq \int_C \left(\int_0^{D_0(\overline{\varphi}_G(c))} (D_0^{-1}(q) - \overline{\phi}_G(c))\, \mathrm{d}q \right) G(\mathrm{d}c).$$

Also, since $\overline{\varphi}_G$ is nondecreasing by definition, $D_0 \circ \overline{\varphi}_G$ is indeed a solution of (21) and hence a solution of (18).

Moreover, since $\overline{\varphi}_G \leq \overline{\boldsymbol{p}}_0$, for all $c \in C$, $\int_c^{\overline{c}} D_0(\overline{\varphi}_G(z))\, \mathrm{d}z \geq \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z))\, \mathrm{d}z$. Therefore, $D_0 \circ \overline{\varphi}_G \in \mathcal{Q}$ is feasible in the primal problem (17). Meanwhile, since $M^*(c) = 0$ for all $c \in [\underline{c}, c^*)$ and since $\overline{\varphi}_G(c) = \overline{\boldsymbol{p}}_0(c)$ for all $c \in (c^*, \overline{c}]$, the complementary slackness condition (19) is also satisfied. Together, $D_0 \circ \overline{\varphi}_G$ is indeed a solution of (17). Finally, by definition of D_0^{-1}, it then follows that

$$R^* = \int_C \left(\int_0^{D_0(\overline{\varphi}_G(c))} (D_0^{-1}(q) - \phi_G(c))\, \mathrm{d}q \right) G(\mathrm{d}c) - \bar{\pi} = \int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \bar{\pi}.$$

The see that any solution of the price-controlling data broker's problem must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G almost all $c \in C$, consider any optimal mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ of the price-controlling data broker. By optimality, it must be that $\mathbb{E}_G[\tau(c)] = R^*$ and that the indirect utility of the producer with marginal cost $\overline{c}$ is $\bar{\pi}$. Thus, by Lemma 7, it must be that

$$\int_C \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \phi_G(c)) D(p) \boldsymbol{\gamma}(\mathrm{d}p|D, c) \right) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) = \int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c), \tag{26}$$

which is equivalent to

$$\begin{aligned} &\int_C \left(\int_{\mathcal{D} \times \mathbb{R}_+} (p - \overline{\varphi}_G(c)) D(p) \boldsymbol{\gamma}(\mathrm{d}p|D, c) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) + \int_C (\overline{\varphi}_G(c) - \phi_G(c)) \boldsymbol{q}_{\boldsymbol{\gamma}}^{\sigma}(c) G(\mathrm{d}c) \\ = &\int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \overline{\varphi}_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) + \int_C (\overline{\varphi}_G(c) - \phi_G(c)) D_0(\overline{\varphi}_G(c)) G(\mathrm{d}c), \end{aligned} \tag{27}$$

where $\boldsymbol{q}_{\boldsymbol{\gamma}}^{\sigma}(c) := \int_{\mathcal{D} \times \mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p|D, c) \sigma(\mathrm{d}D|c)$ for all $c \in C$. Moreover, since for any $c \in C$,

$$\int_{\mathcal{D} \times \mathbb{R}_+} (p - \overline{\varphi}_G(c)) D(p) \boldsymbol{\gamma}(\mathrm{d}p|D, c) \sigma(\mathrm{d}D|c) \leq \int_{\mathcal{D}} \max_{p \in \mathbb{R}_+} [(p - \overline{\varphi}_G(c)) D(p)] \sigma(\mathrm{d}D|c) \leq \int_V (v - \overline{\varphi}_G(c))^+ D_0(\mathrm{d}v), \tag{28}$$

it must be that

$$\int_C (\overline{\varphi}_G(c) - \phi_G(c)) \boldsymbol{q}_{\boldsymbol{\gamma}}^{\sigma}(c) G(\mathrm{d}c) \geq \int_C (\overline{\varphi}_G(c) - \phi_G(c)) D_0(\overline{\varphi}_G(c)) G(\mathrm{d}c).$$

Meanwhile, since $(\sigma, \tau, \boldsymbol{\gamma})$ is incentive compatible, Lemma 6 implies that $\boldsymbol{q}_{\boldsymbol{\gamma}}^{\sigma}$ is nonincreasing in c. Together with (23) and (25), we have

$$\int_C (\overline{\phi}_G(c) - \phi_G(c)) \boldsymbol{q}_{\boldsymbol{\gamma}}^{\sigma}(c) G(\mathrm{d}c) \geq \int_C (\overline{\phi}_G(c) - \phi_G(c)) D_0(\overline{\varphi}_G(c)) G(\mathrm{d}c). \tag{29}$$

Furthermore, since $\overline{\phi}_G(c) = \overline{\boldsymbol{p}}_0(c) \leq \phi_G(c)$ for all $c \in (c^*, \overline{c}]$ and $\overline{\phi}_G(c) = \phi_G(c)$, for all $c \in [\underline{c}, c^*]$, by the definition of M^* given by (20), together with integration by parts, (29) is equivalent to

$$\int_C \left(\int_c^{\overline{c}} \left(\boldsymbol{q}_{\boldsymbol{\gamma}}^{\sigma}(z) - D_0(\overline{\boldsymbol{p}}_0(z)) \right) \mathrm{d}z \right) M^*(\mathrm{d}c) \leq 0 \tag{30}$$

Lastly, since $(\sigma, \tau, \boldsymbol{\gamma})$ is individually rational, for any $c \in C$,

$$\int_c^{\overline{c}} \left(\boldsymbol{q}_{\boldsymbol{\gamma}}^{\sigma}(z) - D_0(\overline{\boldsymbol{p}}_0(z))\right) \mathrm{d}z \geq 0.$$

Thus, as M^* is the CDF of a Borel measure, (30) must hold with equality, which in turn implies that (29) must hold with equality. Together with (27), (28) must hold with equality for G-almost all $c \in C$. Therefore, $(\sigma, \tau, \boldsymbol{\gamma})$ must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G-almost all $c \in C$, as desired. ∎

B.3 Proof of Lemma 1

Proof of Lemma 1. For necessity, consider any incentive compatible mechanism (σ, τ). First notice that, by Proposition 1 of Yang (2020b), $\pi_D : C \to \mathbb{R}_+$ is convex and continuous on C for any $D \in \mathcal{D}$ with $\pi_D'(c) = -D(\boldsymbol{p}_D(c))$ for all $\boldsymbol{p} \in \boldsymbol{P}$ and for almost all $c \in C$. Moreover, since for any $D \in \mathcal{D}$ and for any $\boldsymbol{p} \in \boldsymbol{P}$, $|\pi_D'(c)| = |D(\boldsymbol{p}_D(c))| \leq 1$, for almost all $c \in C$, the order of integral and differential can be interchanged. That is, for any $c, c' \in C$,

$$\frac{\mathrm{d}}{\mathrm{d}c} \int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c') = \int_{\mathcal{D}} \pi_D'(c)\sigma(\mathrm{d}D|c') = -\int_{\mathcal{D}} D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c'). \tag{31}$$

As such, for any $c' \in C$, the function $c \mapsto \int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c')$ is convex and, by (31), has an almost-everywhere derivative $-\int_{\mathcal{D}} D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c')$, for any $\boldsymbol{p} \in \boldsymbol{P}$. Now let $u(c, c') := \int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c') - \tau(c')$ for all $c, c' \in C$ be the producer's profit if her report is c' and marginal cost is c. By the Lebesgue dominated convergence theorem, $u(\cdot, c')$ is convex and continuous on C for all $c' \in C$ as π_D is convex and continuous for all $D \in \mathcal{D}$. Furthermore, since the mechanism (σ, τ) is incentive compatible, by the envelope theorem (Milgrom and Segal, 2002), let $U(c) := u(c, c)$, we then have

$$U(c) = U(\overline{c}) - \int_c^{\overline{c}} \frac{\partial}{\partial c} u(z, z)\, \mathrm{d}z = U(\overline{c}) + \int_c^{\overline{c}} \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(z))\sigma(\mathrm{d}D|z)\right) \mathrm{d}z. \tag{32}$$

Assertion 1 then follows after rearranging.

Furthermore, for any mechanism (σ, τ) satisfying assertion 1 (and hence (32)) with any $\boldsymbol{p} \in \boldsymbol{P}$, we have

$$\begin{aligned}
U(c) - u(c, c') =& (U(c) - U(c')) + \int_{\mathcal{D}} (\pi_D(c) - \pi_D(c'))\sigma(\mathrm{d}D|c') \\
=& \int_c^{c'} \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(z))\sigma(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\boldsymbol{p}_D(z))\sigma(\mathrm{d}D|c')\right) \mathrm{d}z \\
=& \int_c^{c'} \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(z))(\sigma(\mathrm{d}D|z) - \sigma(\mathrm{d}D|c'))\right) \mathrm{d}z,
\end{aligned}$$

where the second equality follows from the fundamental theorem of calculus and (31). Therefore, for any mechanism (σ, τ) satisfying assertion 1 with any $\boldsymbol{p} \in \boldsymbol{P}$, $U(c) \geq u(c, c')$ for all $c, c' \in C$ if and only if assertion 2 holds. This completes the proof. ∎

B.4 Proof of Lemma 2

Proof of Lemma 2. Given any nondecreasing function $\psi : C \to \mathbb{R}_+$, and any ψ-quasi-perfect scheme $\sigma : C \to \mathcal{S}$, suppose that for any $c \in C$, $\psi(z) \leq \overline{\boldsymbol{p}}_D(z)$, for Lebesgue almost all $z \in [\underline{c}, c]$ and for all $D \in \mathrm{supp}(\sigma(c))$.

Then, for any $c, c' \in C$ with $c < c'$,

$$\begin{aligned}\int_c^{c'} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))(\sigma(\mathrm{d}D|z) - \sigma(\mathrm{d}D|c'))\right) \mathrm{d}z =& \int_c^{c'} \left(D_0(\psi(z)) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|c')\right) \mathrm{d}z \\ \geq& \int_c^{c'} \left(D_0(\psi(z)) - \int_{\mathcal{D}} D(\psi(z))\sigma(\mathrm{d}D|c')\right) \mathrm{d}z \\ =& \int_c^{c'} (D_0(\psi(z)) - D_0(\psi(z)))\, \mathrm{d}z \\ =& 0,\end{aligned}$$

where the first equality follows from assertion 3 of Lemma 5, the inequality follows from the hypothesis, and the second equality follows from $\sigma(z) \in \mathcal{S}$ for all $z \in [c, c']$. Meanwhile, for any $c, c' \in C$ with $c > c'$,

$$\begin{aligned}\int_{c'}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))(\sigma(\mathrm{d}D|c) - \sigma(\mathrm{d}D|z))\right) \mathrm{d}z =& \int_{c'}^{c} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|c) - D_0(\psi(z))\right) \mathrm{d}z \\ =& \int_c^{c'} (\min\{D_0(\psi(c)), D_0(z)\} - D_0(\psi(z)))\, \mathrm{d}z \\ \geq& 0,\end{aligned}$$

where the first equality again follows from assertion 3 of Lemma 5, and the second equality follows from the fact that $c < c'$ and from the definition of quasi-perfect segmentations.[28] Therefore, by Lemma 1, there exists a transfer τ such that (σ, τ) is incentive compatible, as desired. ∎

B.5 Proof of Theorem 1

Proof of Theorem 1. I first show that the data broker's optimal revenue must be the same as the price-controlling data broker's optimal revenue R^*. Since R^* is an upper bound of the data broker's revenue under any incentive feasible mechanism, it suffices to find an incentive feasible mechanism for the data broker that gives revenue R^*. To this end, notice that since $c \leq \overline{\varphi}_G(c) \leq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$ and $\overline{\varphi}_G : C \to \mathbb{R}_+$ is nondecreasing, by Lemma 3, there exists a $\overline{\varphi}_G$-quasi-perfect scheme $\bar{\sigma} : C \to \mathcal{S}$ that satisfies (10). Together with Lemma 2, there exists a transfer $\bar{\tau}$ such that $(\bar{\sigma}, \bar{\tau})$ is incentive compatible. Meanwhile, by possibly adding a constant to the transfer $\bar{\tau}$ so that the indirect utility of the producer with cost $\bar{c}$, $U(\bar{c})$, equals to $\bar{\pi}$ under the mechanism $(\bar{\sigma}, \bar{\tau})$, it must be that, for any $c \in C$,

$$\begin{aligned}\int_{\mathcal{D}} \pi_D(c)\bar{\sigma}(\mathrm{d}D|c) - \bar{\tau}(c) =& U(\bar{c}) + \int_c^{\bar{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\bar{\sigma}(\mathrm{d}D|z)\right) \mathrm{d}z \\ =& \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\varphi}_G(z))\, \mathrm{d}z \\ \geq& \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\, \mathrm{d}z \\ =& \pi_{D_0}(c),\end{aligned}$$

[28] More specifically, for any $c \in C$, since $\sigma(c)$ is a $\psi(c)$-quasi-perfect segmentation for c, for any $z > c$ and for any $D \in \mathrm{supp}(\sigma(c))$, if $D(c) > 0$ and $\max(\mathrm{supp}(D)) \geq z$, then $\overline{\boldsymbol{p}}_D(z) = \overline{\boldsymbol{p}}_D(c)$ and hence $D(\overline{\boldsymbol{p}}_D(z)) = D_0(\psi(c)) = D_0(z)$; if $D(c) > 0$ and $\max(\mathrm{supp}(D)) < z$, then $D(\overline{\boldsymbol{p}}_D(z)) = 0$; while if $D(c) = 0$ then $D(z) = 0$.

where the first equality follows from Lemma 1, the second equality follows from assertion 3 of Lemma 5, the inequality follows from $\overline{\varphi}_G \leq \overline{\boldsymbol{p}}_0$ and the last equality follows from (5). As a result, $(\bar{\sigma}, \bar{\tau})$ is individually rational.

Furthermore, since $\bar{\sigma} : C \to \mathcal{S}$ is a $\overline{\varphi}_G$-quasi-perfect scheme, by assertion 3 and assertion 4 of Lemma 5, for any $c \in C$,

$$\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \bar{\sigma}(\mathrm{d}D|c) = \int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v). \tag{33}$$

and therefore, together with Lemma 1,

$$\begin{aligned}
\mathbb{E}[\bar{\tau}(c)] =& \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \bar{\sigma}(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{\pi} \\
=& \int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \bar{\pi} \\
=& R^*,
\end{aligned}$$

as desired.

Since the data broker's optimal revenue is R^* and since (33) holds for any $\overline{\varphi}_G$-quasi-perfect scheme σ, by Lemma 1, any incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism must give revenue R^* and hence is optimal.

Conversely, to see why any optimal mechanism must be a $\overline{\varphi}_G$-quasi-perfect mechanism, consider any optimal mechanism (σ, τ). As it is optimal and incentive compatible, by Lemma 1,

$$R^* = \mathbb{E}[\tau(c)] = \int_C \left(\int_{\mathcal{D}} (\boldsymbol{p}_D(c) - \phi_G(c)) D(\boldsymbol{p}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{\pi}, \tag{34}$$

for any $\boldsymbol{p} \in \boldsymbol{P}$. Also, since (σ, τ) is incentive compatible, for any $\boldsymbol{p} \in \boldsymbol{P}$, the function

$$c \mapsto \int_{\mathcal{D}} D(\boldsymbol{p}_D(c)) \sigma(\mathrm{d}D|c)$$

is nonincreasing on C.[29] Thus, by (23),

$$\int_C \overline{\phi}_G(c) \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) \geq \int_C \overline{\varphi}_G(c) \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c). \tag{35}$$

Moreover, since (σ, τ) is individually rational, by Lemma 1, it must be that

$$\int_c^{\overline{c}} \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(z)) \sigma(\mathrm{d}D|z) \right) \mathrm{d}z \geq \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z)) \, \mathrm{d}z, \; \forall c \in C. \tag{36}$$

Now suppose that (σ, τ) is not a $\overline{\varphi}_G$-quasi-perfect mechanism or it does not induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for a positive G-measure of c, then there exists $\boldsymbol{p} \in \boldsymbol{P}$, a positive G-measure of c and a positive $\sigma(c)$-measure of $D \in \mathcal{D}$ such that either $\boldsymbol{p}_D(c) < \overline{\boldsymbol{p}}_D(c)$, or $D(c) > 0$ and either $\#\{v \in \mathrm{supp}(D) :$

[29]To see this, notice that U is convex since it is a pointwise supremum of convex functions, which is because $\pi_D(c)$ is convex for all D. Lemma 1 implies that the derivative of U is nondecreasing and thus $c \mapsto \int_{\mathcal{D}} D(\boldsymbol{p}_D(c)) \sigma(\mathrm{d}D|c)$ must be nonincreasing.

$v \geq \overline{\varphi}_G(c)\} \neq 1$ or $\max(\mathrm{supp}(D)) \notin \boldsymbol{P}_D(c)$, which imply that there is a positive G-measure of c and a positive $\sigma(c)$-measure of D such that

$$\begin{aligned}\int_{\{v\geq\overline{\varphi}_G(c)\}}(v-\overline{\varphi}_G(c))D(\mathrm{d}v) \geq &\int_{\{v\geq\boldsymbol{p}_D(c)\}}(v-\overline{\varphi}_G(c))D(\mathrm{d}v)\\ =&(\boldsymbol{p}_D(c)-\overline{\varphi}_G(c))D(\boldsymbol{p}_D(c))+\int_{\{v\geq\boldsymbol{p}_D(c)\}}(v-\boldsymbol{p}_D(c))D(\mathrm{d}v)\\ \geq&(\boldsymbol{p}_D(c)-\overline{\varphi}_G(c))D(\boldsymbol{p}_D(c)),\end{aligned}$$

with at least one inequality being strict. Therefore,

$$\int_C\left(\int_{\mathcal{D}}(\boldsymbol{p}_D(c)-\overline{\varphi}_G(c))D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)<\int_C\left(\int_V(v-\overline{\varphi}_G(c))^+D_0(\mathrm{d}v)\right)G(\mathrm{d}c). \tag{37}$$

Meanwhile, since by (34)

$$\begin{aligned}&\int_C\left(\int_{\mathcal{D}}(\boldsymbol{p}_D(c)-\overline{\varphi}_G(c))D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)+\int_C(\overline{\varphi}_G(c)-\phi_G(c))\left(\int_{\mathcal{D}}D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)\\ =&\int_{\mathcal{D}}\left(\int_{\mathcal{D}}(\boldsymbol{p}_D(c)-\phi_G(c))D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)\\ =&\int_C\left(\int_{\{v\geq\overline{\varphi}_G(c)\}}(v-\phi_G(c))D_0(\mathrm{d}v)\right)G(\mathrm{d}c)\\ =&\int_C\left(\int_V(v-\overline{\varphi}_G(c))^+D_0(\mathrm{d}v)\right)G(\mathrm{d}c)+\int_C(\overline{\varphi}_G(c)-\phi_G(c))D_0(\overline{\varphi}_G(c))G(\mathrm{d}c),\end{aligned}$$

(25), (35) and (37) imply that

$$\begin{aligned}\int_C(\overline{\phi}_G(c)-\phi_G(c))\left(\int_{\mathcal{D}}D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)\geq&\int_C(\overline{\varphi}_G(c)-\phi_G(c))\left(\int_{\mathcal{D}}D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)\\ >&\int_C(\overline{\varphi}_G(c)-\phi_G(c))D_0(\overline{\varphi}_G(c))G(\mathrm{d}c)\\ =&\int_C(\overline{\phi}_G(c)-\phi_G(c))D_0(\overline{\varphi}_G(c))G(\mathrm{d}c),\end{aligned}$$

where the first inequality follows from (35) and the equality follows from (25). Furthermore, since $\overline{\phi}_G(c)=\phi_G(c)$ for all $c\in[\underline{c},c^*]$ and $\overline{\phi}_G(c)=\overline{\varphi}_G(c)=\overline{\boldsymbol{p}}_0(c)$ for all $c\in(c^*,\overline{c}]$, it then follows that

$$\int_{c^*}^{\overline{c}}(\phi_G(c)-\overline{\boldsymbol{p}}_0(c))\left(\int_{\mathcal{D}}D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)<\int_{c^*}^{\overline{c}}(\phi_G(c)-\overline{\boldsymbol{p}}_0(c))D_0(\overline{\boldsymbol{p}}_0(c))G(\mathrm{d}c),$$

Using integration by parts, this is equivalent to

$$\int_{c^*}^{\overline{c}}\left(\int_c^{\overline{c}}\left(\int_{\mathcal{D}}D(\boldsymbol{p}_D(z))\sigma(\mathrm{d}D|z)\right)\mathrm{d}z\right)M^*(\mathrm{d}c)<\int_{c^*}^{\overline{c}}\left(\int_c^{\overline{c}}D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z\right)M^*(\mathrm{d}c),$$

where M^* is defined in (20). However, by (36) and by the fact that M^* is a CDF of a Borel measure, which is due to Assumption 1,

$$\int_{c^*}^{\overline{c}}\left(\int_c^{\overline{c}}\left(\int_{\mathcal{D}}D(\boldsymbol{p}_D(z))\sigma(\mathrm{d}D|z)\right)\mathrm{d}z\right)M^*(\mathrm{d}c)\geq\int_{c^*}^{\overline{c}}\left(\int_c^{\overline{c}}D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z\right)M^*(\mathrm{d}c),$$

a contradiction. Therefore, σ must be a $\overline{\varphi}_G$-quasi-perfect scheme and must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G-almost all $c\in C$. Together with Lemma 1, and the fact that $U(\overline{c})=\bar{\pi}$ under any optimal mechanism, (σ,τ) must be a $\overline{\varphi}_G$-quasi-perfect mechanism. This completes the proof. ■

B.6 Proof of Theorem 2

Proof of Theorem 2. By the proof of Lemma 3 in the main text. When D_0 is regular, since $c \leq \overline{\varphi}_G(c) \leq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$, the canonical $\overline{\varphi}_G$-quasi-perfect scheme σ^* defined in (3) is implementable. Therefore, there exists τ^* such that (σ^*, τ^*) is an incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism. By Theorem 1, (σ^*, τ^*) is optimal. ∎

C Proofs of Other Main Results

C.1 Proof of Theorem 3

Proof of Theorem 3. Let (σ, τ) be any optimal mechanism. By Theorem 1, (σ, τ) must be a $\overline{\varphi}_G$-quasi-perfect mechanism and induces $\overline{\varphi}_G$-quasi-perfect price discrimination. Therefore, for any $\boldsymbol{p} \in \boldsymbol{P}$, for G-almost all $c \in C$ and for $\sigma(c)$-almost all $D \in \mathcal{D}$, $D(p) = 0$ for all $p > \boldsymbol{p}_D(c)$ and thus consumer surplus is

$$\int_C \left(\int_{\mathcal{D}} \left(\int_{\{v \geq \boldsymbol{p}_D(c)\}} (v - \boldsymbol{p}_D(c)) D(\mathrm{d}v) \right) \sigma^*(\mathrm{d}D|c) \right) G(\mathrm{d}c) = \int_C \left(\int_{\mathcal{D}} \left(\int_{\boldsymbol{p}_D(c)}^{\overline{v}} D(z)\, \mathrm{d}z \right) \sigma^*(\mathrm{d}D|c) \right) G(\mathrm{d}c) = 0,$$

as desired. ∎

C.2 Proof of Proposition 1

Proof of Proposition 1. Since $\boldsymbol{P}_{D_0}(c)$ is a singleton for (Lebesgue)-almost all $c \in C$ and since G is absolutely continuous, consumer surplus under uniform pricing does not depend which selection $\boldsymbol{p} \in \boldsymbol{P}$ is used. Therefore, by Theorem 1, the difference between the data broker's optimal revenue and the consumer surplus under uniform pricing is

$$\begin{aligned}
&\int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \bar{\pi} - \int_C \left(\int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \overline{\boldsymbol{p}}_0(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}C) \\
=& \int_C \left((\overline{\boldsymbol{p}}_0(c) - \phi_G(c) D_0(\overline{\boldsymbol{p}}_0(c)) + \int_{[\overline{\varphi}_G(c), \overline{\boldsymbol{p}}_0(c))} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) - \bar{\pi} \\
=& \int_C \left(\int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z))\, \mathrm{d}z - \frac{G(c)}{g(c)} D_0(\overline{\boldsymbol{p}}_0(c)) \right) G(\mathrm{d}c) \\
&+ \int_C \left(\int_{[\overline{\varphi}_G(c), \overline{\boldsymbol{p}}_0(c))} (v - \varphi_G(c)) D_0(\mathrm{d}v) + \int_C (\varphi_G(c) - \phi_G(c))(D_0(\overline{\varphi}_G(c)) - D_0(\overline{\boldsymbol{p}}_0(c))) \right) G(\mathrm{d}c) \\
\geq& \int_C G(c)(D_0(\overline{\boldsymbol{p}}_0(c)) - D_0(\overline{\boldsymbol{p}}_0(c)))\, \mathrm{d}c + \int_C (\varphi_G(c) - \phi_G(c)) D_0(\overline{\varphi}_G(c)) G(\mathrm{d}c) - \int_C (\varphi_G(c) - \phi_G(c)) D_0(\overline{\boldsymbol{p}}_0(c)) G(\mathrm{d}c) \\
\geq& 0.
\end{aligned}$$

where the second equality follows from Lemma 1, the first inequality follows from the fact that $\overline{\varphi}_G(c) < \overline{\boldsymbol{p}}_0(c)$ if and only if $\varphi_G(c) < \overline{\boldsymbol{p}}_0(c)$, and the last inequality follows from (23) and (25). This completes the proof. ∎

C.3 Proof of Theorem 7

Proof of Theorem 7. By Lemma 4, whose proof can be found in the Online Appendix, it suffices to prove the outcome-equivalence between data brokership and price-controlling data brokership. By Proposition 2 and Theorem 1, both the data broker and the price-controlling data broker have optimal revenue R^*. Furthermore, for any optimal mechanism (σ, τ) of the data broker and any optimal mechanism $(\hat{\sigma}, \hat{\tau}, \hat{\boldsymbol{\gamma}})$ of the price-controlling data broker, both of them must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G-almost all $c \in C$. In particular, for G-almost all $c \in C$, all the consumers with $v \geq \overline{\varphi}_G(c)$ buys the product by paying their values and all the consumers with $v < \overline{\varphi}_G(c)$ do not buy the product. Thus, the consumer surplus and the allocation of the product induced by (σ, τ) and $(\hat{\sigma}, \hat{\tau}, \hat{\boldsymbol{\gamma}})$ are the same.

In addition, for any optimal mechanism (σ, τ) of the data broker, Theorem 1 implies that σ must be a $\overline{\varphi}_G$-quasi-perfect scheme and hence by assertions 3 and 4 of Lemma 5, and by Lemma 1, for Lebesgue almost all $c \in C$,

$$\int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c) - \tau(c) = \bar{\pi} + \int_c^{\bar{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) \right) \mathrm{d}z = \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\varphi}_G(z))\,\mathrm{d}z. \tag{38}$$

Meanwhile, for the price-controlling data broker's optimal mechanism $(\hat{\sigma}, \hat{\tau}, \hat{\boldsymbol{\gamma}})$, since, by Proposition 1, it induces $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for almost all $c \in C$, it must be that $\boldsymbol{q}^{\hat{\sigma}}_{\hat{\boldsymbol{\gamma}}}(c) = D_0(\overline{\varphi}_G(c))$. Together with Lemma 6, for any $c \in C$,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p-c)D(p)\hat{\boldsymbol{\gamma}}(\mathrm{d}p|D,c) \right) \hat{\sigma}(\mathrm{d}D|c) - \hat{\tau}(c) = \bar{\pi} + \int_c^{\bar{c}} \boldsymbol{q}^{\hat{\sigma}}_{\hat{\boldsymbol{\gamma}}}(z)\,\mathrm{d}z = \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\varphi}_G(z))\,\mathrm{d}z. \tag{39}$$

Thus, the producer's profit under both (σ, τ) and $(\hat{\sigma}, \hat{\tau}, \hat{\boldsymbol{\gamma}})$ are the same. This completes the proof. ∎

References

Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2019): "Too Much Data: Prices and Inefficiencies in Data Market," Working Paper.

Admati, A. R. and P. Pfleiderer (1985): "A Monopolistic Market for Information," *Journal of Economic Theory*, 39, 400–438.

——— (1990): "Direct and Indirect Sale of Information," *Econometrica*, 58, 901–928.

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

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

Bergemann, D. and A. Bonatti (2015): "Selling Cookies," *American Economic Journal: Microeconomics*, 7, 259–294.

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

Bergemann, D., A. Bonatti, and A. Smolin (2018): "The Design and Price of Information," *American Economic Review*, 108, 1–45.

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

Bergemann, D. and S. Morris (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487–522.

Bergemann, D. and M. Pesendorfer (2007): "Information Structures in Optimal Auctions," *Journal of Economic Theory*, 137, 580–609.

Bergemann, D. and J. Valimaki (2019): "Dynamic Mechanism Design: An Introduction," *Journal of Economic Literature*, 57, 235–274.

Carbajal, J. C. and J. Ely (2013): "Mechanism Design without Revenue Equivalence," *Journal of Economic Theory*, 148, 104–133.

Cowan, S. (2016): "Welfare-Increasing Third-Degree Price Discrimination," *RAND Journal of Economics*, 47, 326–340.

Dworczak, P. (2020): "Mechanism Design with Aftermarkets: Cutoff Mechanisms," *Econometrica*, Forthcoming.

Haghpanah, N. and R. Siegel (2020): "Pareto Improving Segmentation of Multi-product Markets," Working Paper.

Ichihashi, S. (2020): "Non-Competing Data Intermediaries," Working Paper.

Jones, C. and C. Tonetti (2020): "Nonrivalry and the Economics of Data," *American Economic Review*, 110, 2819–2858.

Jullien, B. (2000): "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93, 1–47.

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

Lewis, T. R. and D. E. M. Sappington (1989): "Countervailing Incentives in Agency Problems," *Journal of Economic Theory*, 49, 294–313.

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

Federal Trade Commission (2014): "Data Brokers: A Call for Transparency and Accountability," https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf (accessed June 20, 2019).

Milgrom, P. and I. Segal (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583–601.

Mussa, M. and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–307.

Myerson, R. (1979): "Incentive Compatibility and the Bargaining Problem," *Econometrica*, 47, 61–73.

——— (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Mylovanov, T. and T. Tröger (2014): "Mechanism Design by an Informed Principal: The Quasi-Linear Private-Values Case," *Review of Economic Studies*, 81, 1668–1707.

Pavan, A., I. Segal, and J. Toikka (2014): "Dynamic Mechanism Design: A Myersonian Approach," *Econometrica*, 82, 601–653.

Riley, J. and R. Zeckhauser (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm," *Quarterly Journal of Economics*, 98, 267–289.

Rochet, J.-C. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-linear Context," *Journal of Mathematical Economics*, 16, 191–200.

Segura-Rodriguez, C. (2020): "Selling Data," Working Paper.

Sinander, L. (2020): "The Converse of Envelope Theorem," Working Paper.

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

Wei, D. and B. Green (2020): "(Reverse) Price Discrimination with Information Design," Working Paper.

Yamashita, T. (2017): "Optimal Public Information Disclosure by Mechanism Designer," Working Paper.

Yang, K. H. (2019): "Equivalence in Business Models of Informational Intermediary," Working Paper.

——— (2020a): "A Note on Generating Arbitrary Joint Distributions using Partitions," Working Paper.

——— (2020b): "A Note on Topological Properties of Outcomes in a Monopoly Market," Working Paper.

——— (2020c): "Selling Consumer Data for Profit: Optimal Market-Segmentation Design and its Consequences," Discussion Paper 2258, Cowles Foundation for Research in Economics, Yale University.

Selling Consumer Data for Profit:
Optimal Market-Segmentation Design and its Consequences
Online Appendix

Kai Hao Yang*

December 3, 2020

Proof of Lemma 3

The proof of Lemma 3 relies on the following technical lemma

Lemma A.1. *Consider any function $\psi : C \to \mathbb{R}_+$ with $c \le \psi(c)$ for all $c \in C$. Given any $\{D_n\} \subset \mathcal{D}$ and $\{\sigma_n\}$ such that $\sigma_n : C \to \mathcal{S}_{D_n}$ is measurable for all $n \in \mathbb{N}$. Suppose that $\{\sigma_n\} \to \sigma$ pointwise and $\{D_n\} \to D_0$ for some $\sigma : C \to \Delta(\mathcal{D})$ and $D_0 \in \mathcal{D}$. Then σ is measurable and $\sigma(c) \in \mathcal{S}_{D_0}$ for all $c \in C$. Moreover, suppose further that σ_n is a ψ-quasi-perfect scheme for all $n \in \mathbb{N}$. Then σ is a ψ-quasi-perfect scheme.*

Proof. First notice that since for all $n \in \mathbb{N}$ and for all $c \in C$, $\sigma_n(c) \in \mathcal{S}_{D_n}$ and since $\{\sigma_n\} \to \sigma$ pointwise, σ is measurable. Moreover, since $\{D_n\} \to D_0$ and $\{\sigma_n\} \to \sigma$, for any bounded continuous function $f : V \to \mathbb{R}$ and for any $c \in C$,

$$\begin{aligned}
\int_V f(v)\left(\int_{\mathcal{D}} D(\mathrm{d}v)\sigma(\mathrm{d}D|c)\right) &= \int_{\mathcal{D}}\left(\int_V f(v)D(\mathrm{d}v)\right)\sigma(\mathrm{d}D|c)\\
&= \lim_{n\to\infty}\int_{\mathcal{D}}\left(\int_V f(v)D(\mathrm{d}v)\right)\sigma_n(\mathrm{d}D|c)\\
&= \lim_{n\to\infty}\int_V f(v)\left(\int_{\mathcal{D}} D(\mathrm{d}v)\sigma_n(\mathrm{d}D|c)\right)\\
&= \lim_{n\to\infty}\int_V f(v)D_n(\mathrm{d}v)\\
&= \int_V f(v)D_0(\mathrm{d}v),
\end{aligned}$$

where the first and the third equality follow from interchanging the order of integrals, the second equality follows from the fact that the integrand in the parentheses is a bounded continuous function of D and from weak-*convergence of $\{\sigma_n(c)\}$, the fourth equality is due to the fact that $\sigma_n(c) \in \mathcal{S}_{D_n}$, and the last equality follows from the weak-* convergence of $\{D_n\}$. Thus, by the Riesz representation theorem,

$$\int_{\mathcal{D}} D(p)\sigma(\mathrm{d}D|c) = D_0(p),\ \forall p \in V,\ c \in C.$$

*Cowles Foundation for Economic Research, Yale University, kaihao.yang@yale.edu.

This proves that $\sigma(c) \in \mathcal{S}_{D_0}$ for all $c \in C$.

Now suppose that σ_n is a ψ-quasi-perfect scheme for all $n \in \mathbb{N}$ and suppose that, by way of contradiction, $\sigma : C \to \mathcal{S}_{D_0}$ is not a ψ-quasi-perfect scheme. Then there exists a positive G-measure of c and a positive $\sigma(c)$-measure of $D \in \mathcal{D}$ such that $D(c) > 0$ and either $\#\{v \in \text{supp}(D) : v \geq \psi(c)\} \neq 1$ or $\max(\text{supp}(D)) \notin \boldsymbol{P}_D(c)$ (i.e., $D(\overline{\boldsymbol{p}}_D(c)) > 0$). As such, there is a positive G-measure of c and a positive $\sigma(c)$-measure of D such that

$$\begin{aligned}\int_{\{v \geq \psi(c)\}} (v - \psi(c))D(\mathrm{d}v) \geq & \int_{\{v \geq \overline{\boldsymbol{p}}_D(c)\}} (v - \psi(c))D(\mathrm{d}v) \\ =&(\overline{\boldsymbol{p}}_D(c) - \psi(c))D(\overline{\boldsymbol{p}}_D(c)) + \int_{\{v \geq \overline{\boldsymbol{p}}_D(c)\}} (v - \overline{\boldsymbol{p}}_D(c))D(\mathrm{d}v) \\ \geq&(\overline{\boldsymbol{p}}_D(c) - \psi(c))D(\overline{\boldsymbol{p}}_D(c)),\end{aligned}$$

with at least one inequality being strict. Thus, there exists a positive G-measure of $c \in C$ such that

$$\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \psi(c))D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) < \int_V (v - \psi(c))^+ D_0(\mathrm{d}v).$$

However, by Theorem 12 of Hart and Reny (2019) and Corollary 2 of Yang (2020a), for Lebesgue almost all $c \in C$,

$$\begin{aligned}&\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \psi(c))D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) \\ =&\int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c) - (\psi(c) - c)\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) \\ \geq&\lim_{n\to\infty} \int_{\mathcal{D}} \pi_D(c)\sigma_n(\mathrm{d}D|c) - \liminf_{n\to\infty}(\psi(c) - c)\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma_n(\mathrm{d}D|c) \\ =&\limsup_{n\to\infty}\left[\int_{\mathcal{D}} \pi_D(c)\sigma_n(\mathrm{d}D|c) - (\psi(c) - c)\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma_n(\mathrm{d}D|c)\right] \\ =&\limsup_{n\to\infty}\int_V (v - \psi(c))^+ D_n(\mathrm{d}v) \\ =&\lim_{n\to\infty}\int_V (v - \psi(c))^+ D_n(\mathrm{d}v) \\ =&\int_V (v - \psi(c))^+ D_0(\mathrm{d}v),\end{aligned} \tag{A.1}$$

a contradiction. Here, the first inequality follows from the fact that $\{\sigma_n(c)\} \to \sigma(c)$, Theorem 12 of Hart and Reny (2019) and Corollary 2 of Yang (2020a); the second equality follows from the properties of the $\liminf$ and $\limsup$ operators;[1] the third equality follows from the fact that $\sigma_n(c) \in \mathcal{S}_{D_n}$ and is a $\psi(c)$-quasi-perfect segmentation for c; and the last two equalities follow from the fact that the function $(v - \psi(c))^+$ is bounded and continuous in v and that $\{D_n\} \to D_0$. Therefore, σ must be a ψ-quasi-perfect scheme. ■

[1] More precisely, this follows from the following properties: For any real sequences $\{a_n\}$, $\{b_n\}$,

$$-\liminf_{n\to\infty} b_n = \limsup_{n\to\infty}(-b_n).$$

Moreover, if $\{a_n\}$ is convergent, then

$$\limsup_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

Proof of Lemma 3 (general D_0). I first prove the lemma for D_0 being a step function with finitely many steps. Consider any step function $D \in \mathcal{D}$ with $|\text{supp}(D)| < \infty$ and any nondecreasing function $\psi : C \to \mathbb{R}_+$ such that $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_D(c)$ for all $c \in C$ and fix any $c \in C$, let

$$V^+ := \{v \in \text{supp}(D) : v \geq \psi(c)\}$$

and let

$$\hat{c} := \inf\{z \in C : \overline{\boldsymbol{p}}_D(z) \geq \psi(c)\}.$$

Since $\overline{\boldsymbol{p}}_D$ is nondecreasing, it then follows $\overline{\boldsymbol{p}}_D(z) \geq \psi(c)$ for all $z \in [\hat{c}, \overline{c}]$ and $\overline{\boldsymbol{p}}_D(z) \leq \psi(c)$ for all $z \in [\underline{c}, \hat{c})$. Moreover, since $\psi(c) \leq \overline{\boldsymbol{p}}_D(c)$, $\hat{c} \leq c$. Furthermore, by definition of $\hat{c}$, it must be either $\hat{c} = \underline{c}$ or $\hat{c} > \underline{c}$ and $\underline{\boldsymbol{p}}_D(\hat{c}) < \psi(c) \leq \overline{\boldsymbol{p}}_D(\hat{c})$, where $\underline{\boldsymbol{p}}_D(\hat{c}) := \min \boldsymbol{P}_D(\hat{c})$, since otherwise, if $\hat{c} > \underline{c}$ and $\underline{\boldsymbol{p}}_D(\hat{c}) \geq \psi(c)$, then for $\varepsilon > 0$ small enough, as $|\text{supp}(D)| < \infty$, $\overline{\boldsymbol{p}}_D(\hat{c} - \varepsilon) = \underline{\boldsymbol{p}}_D(\hat{c}) \geq \psi(c)$, contradicting to the definition of $\hat{c}$. Consider first the case where $\hat{c} > \underline{c}$. In this case, for each $v \in V^+$, define $\hat{m}^v$ recursively as the following

$$\hat{m}^v(v') := \begin{cases} 0, & \text{if } v' \geq \psi(c) \text{ and } v' \neq v \\ m^D(v'), & \text{if } v' = v \\ \beta^*(v|v')m^D(v'), & \text{if } \underline{\boldsymbol{p}}_D(\hat{c}) \leq v' < \psi(c) \\ \alpha^*(v)m^D(v'), & \text{if } v' < \underline{\boldsymbol{p}}_D(\hat{c}) \end{cases}, \forall v' \in \text{supp}(D), \forall v \in V^+,$$

where for all $v \in V^+$ and all $v' \in \text{supp}(D)$ s.t. $\underline{\boldsymbol{p}}_D(\hat{c}) \leq v' < \psi(c)$,

$$\beta^*(v|v') := \frac{(v - \hat{c})m^D(v) - (v' - \hat{c})\sum_{\hat{v} > v'} \hat{m}^v(\hat{v})}{\sum_{v \geq \psi(c)} \left[(v - \hat{c})m^D(v) - (v' - \hat{c})\sum_{\hat{v} > v'} \hat{m}^v(\hat{v})\right]},$$

and for all $v \in V^+$,

$$\alpha^*(v) := \frac{\sum_{\hat{v} \geq \underline{\boldsymbol{p}}_D(\hat{c})} \hat{m}^v(\hat{v})}{\sum_{\hat{v} \geq \underline{\boldsymbol{p}}_D(\hat{c})} m^D(\hat{v})}.$$

By construction,

$$\sum_{v \in V^+} \alpha^*(v) = \sum_{v \in V^+} \beta^*(v|v') = 1 \tag{A.2}$$

for all $v' \in \text{supp}(D)$ with $\underline{\boldsymbol{p}}_D(\hat{c}) \leq v' < \psi(c)$. As such,

$$\sum_{v \in V^+} \hat{m}^v(v') = m^D(v'), \forall v' \in \text{supp}(D). \tag{A.3}$$

Notice that since $\hat{c} \leq \underline{\boldsymbol{p}}_D(\hat{c}) < \psi(c) \leq \overline{\boldsymbol{p}}_D(\hat{c})$, it must be that

$$\sum_{v \geq \psi(c)} (v - \hat{c})m^D(v) \geq \sum_{v \geq \overline{\boldsymbol{p}}_D(\hat{c})} (v - \hat{c})m^D(v) \geq (\overline{\boldsymbol{p}}_D(\hat{c}) - \hat{c})D(\overline{\boldsymbol{p}}_D(\hat{c})) = (\underline{\boldsymbol{p}}_D(\hat{c}) - \hat{c})D(\underline{\boldsymbol{p}}_D(\hat{c})). \tag{A.4}$$

Now consider any $v' \in \text{supp}(D)$ such that $\underline{\boldsymbol{p}}_D(\hat{c}) \leq v' < \psi(c)$. Notice first that

$$\begin{aligned}
&\sum_{v \geq \psi(c)} \left[(v-\hat{c})m^D(v) - (v'-\hat{c}) \sum_{\hat{v}>v'} \hat{m}^v(\hat{v}) \right] \\
=& \sum_{v \geq \psi(c)} (v-\hat{c})m^D(v) - (v'-\hat{c}) \sum_{\hat{v}>v'} m^D(\hat{v}) \\
\geq& (\underline{\boldsymbol{p}}_D(\hat{c}) - \hat{c}) D(\underline{\boldsymbol{p}}_D(\hat{c})) - (v'-\hat{c}) \sum_{\hat{v}>v'} m^D(\hat{v}) \\
\geq& (v'-\hat{c}) \sum_{\hat{v} \geq v'} m^D(\hat{v}) - (v'-\hat{c}) \sum_{\hat{v}>v'} m^D(\hat{v}) \\
=& (v'-\hat{c}) m^D(v') \\
\geq& 0,
\end{aligned}$$

where the first equality follows from (A.3), the first inequality follows from (A.4), the second inequality follows from the fact that $\underline{\boldsymbol{p}}_D(\hat{c}) \in \boldsymbol{P}_D(\hat{c})$, and the last inequality follows from $\underline{\boldsymbol{p}}_D(\hat{c}) \geq \hat{c}$. As such, for any $v' \in \text{supp}(D)$ with $\underline{\boldsymbol{p}}_D(\hat{c}) \leq v' < \psi(c)$ and for any $v \in V^+$, if

$$(v-\hat{c})m^D(v) - (v'-\hat{c}) \sum_{\hat{v}>v'} \hat{m}^v(\hat{v}) \geq 0,$$

then $\beta^*(v|v') \geq 0$ and

$$\begin{aligned}
&\hat{m}^v(v') \leq \frac{(v-\hat{c})m^D(v) - (v'-\hat{c}) \sum_{\hat{v}>v'} \hat{m}^v(\hat{v})}{(v'-\hat{c})m^D(v')} m^D(v') \\
\iff& (v'-\hat{c})\hat{m}^v(v') + (v'-\hat{c}) \sum_{\hat{v}>v'} \hat{m}^v(\hat{v}) \leq (v-\hat{c})m^D(v) \\
\iff& (v'-\hat{c}) \sum_{\hat{v} \geq v'} \hat{m}^v(\hat{v}) \leq (v-\hat{c})\hat{m}^v(v),
\end{aligned}$$

which in turn implies that

$$(v-\hat{c})m^D(v) - (v''-\hat{c}) \sum_{\hat{v}>v''} \hat{m}^v(\hat{v}) > (v-\hat{c})m^D(v) - (v'-\hat{c}) \sum_{\hat{v} \geq v'} \hat{m}^v(\hat{v}) \geq 0,$$

where $v'' \in \text{supp}(D)$ is the largest element of $\{\hat{v} \in \text{supp}(D) : \underline{\boldsymbol{p}}_D(\hat{c}) \leq \hat{v} < v'\}$. Moreover, if $v' = \max\{\hat{v} \in \text{supp}(D) : \underline{\boldsymbol{p}}_D(\hat{c}) \leq \hat{v} < \psi(c)\}$, then clearly, for all $v \in V^+$,

$$(v-\hat{c})m^D(v) - \sum_{\hat{v}>v'} \hat{m}^v(v') = (v-v')m^D(v) \geq 0.$$

Therefore, by induction, for any $v' \in \text{supp}(D)$ such that $\underline{\boldsymbol{p}}_D(\hat{c}) \leq v' < \psi(c)$, it must be that $\beta^*(v|v') \geq 0$ for all $v \in V^+$ and that

$$(v'-\hat{c}) \sum_{\hat{v} \geq v'} \hat{m}^v(\hat{v}) \leq (v-\hat{c})\hat{m}^v(v). \tag{A.5}$$

Together with (A.2), this also ensures that

$$\alpha^* \in \Delta(V^+) \tag{A.6}$$

and

$$\beta^*(v') \in \Delta(V^+), \tag{A.7}$$

for all $v' \in \text{supp}(D)$ such that $\underline{\boldsymbol{p}}_D(\hat{c}) \le v' < \psi(c)$.

Meanwhile, for any $v' \in \text{supp}(D)$ with $v' \le \underline{\boldsymbol{p}}_D(\hat{c})$ and any $v \in V^+$, notice that by the definition of α^*,

$$\sum_{\hat{v} \ge v'} \hat{m}^v(\hat{v}) = \alpha^*(v) \sum_{v' \le \hat{v} < \underline{\boldsymbol{p}}_D(\hat{c})} m^D(\hat{v}) + \sum_{\hat{v} \ge \underline{\boldsymbol{p}}_D(\hat{c})} \hat{m}^v(\hat{v}) = \alpha^*(v) \sum_{\hat{v} \ge v'} m^D(\hat{v}). \tag{A.8}$$

Thus, for any $v' \in \text{supp}(D)$ with $v' < \underline{\boldsymbol{p}}_D(\hat{c})$ and any $v \in V^+$,

$$\begin{aligned}(v' - \hat{c}) \sum_{\hat{v} \ge v'} \hat{m}^v(\hat{v}) =& \alpha^*(v)(v' - \hat{c}) D(v') \\ \le& \alpha^*(v)(\underline{\boldsymbol{p}}_D(\hat{c}) - \hat{c}) D(\underline{\boldsymbol{p}}_D(\hat{c})) \\ =& (\underline{\boldsymbol{p}}_D(\hat{c}) - \hat{c}) \sum_{\hat{v} \ge \underline{\boldsymbol{p}}_D(\hat{c})} \hat{m}^v(\hat{v}) \\ \le& (v - \hat{c}) \hat{m}^v(v),\end{aligned} \tag{A.9}$$

where both equalities follow from (A.8), the first inequality follows from the fact that $\underline{\boldsymbol{p}}_D(\hat{c}) \in \boldsymbol{P}_D(\hat{c})$, and the last inequality follows from (A.5) by taking $v' = \underline{\boldsymbol{p}}_D(\hat{c})$.

Moreover, by (A.8), for any $z \in [\underline{c}, \hat{c})$, and any $v \in V^+$, since $\overline{\boldsymbol{p}}_D(z) \le \underline{\boldsymbol{p}}_D(\hat{c})$, it must be that for all $v' \le \overline{\boldsymbol{p}}_D(z)$,

$$\begin{aligned}(v' - z) \sum_{\hat{v} \ge v'} \hat{m}^v(\hat{v}) =& \alpha^*(v)(v' - z) D(v') \\ \le& \alpha^*(v)(\underline{\boldsymbol{p}}_D(z) - z) D(\underline{\boldsymbol{p}}_D(z)) \\ =& (\underline{\boldsymbol{p}}_D(z) - z) \sum_{\hat{v} \ge \underline{\boldsymbol{p}}_D(z)} \hat{m}^v(\hat{v}).\end{aligned} \tag{A.10}$$

Finally, if $\hat{c} = \underline{c}$, then define $\{\hat{m}^v\}_{v \in V^+}$ as

$$\hat{m}^v(v') := \begin{cases} m^D(v'), & \text{if } v' = v \\ 0, & \text{if } v' \ge \psi(c) \text{ and } v' \ne v \\ \alpha^*(v) m^D(v'), & \text{if } v' < \psi(c) \end{cases}, \forall v' \in V,\, v \in V^+,\, v \ge \overline{\boldsymbol{p}}_D(\underline{c})$$

and

$$\hat{m}^v(v') := \begin{cases} m^D(v'), & \text{if } v' = v \\ 0, & \text{if } v' \ne v \end{cases}, \forall v' \in V,\, v \in V^+,\, \psi(c) \le v < \overline{\boldsymbol{p}}_D(\underline{c})$$

where

$$\alpha^*(v) := \frac{m^D(v)}{\sum_{v' \ge \overline{\boldsymbol{p}}_D(\underline{c})} m^D(v')}.$$

Again,

$$\sum_{v \ge \overline{\boldsymbol{p}}_D(\underline{c})} \alpha^*(v) = 1 \tag{A.11}$$

and hence

$$\sum_{v \in V^+} \hat{m}^v(v') = m^D(v'),\, \forall v' \in V. \tag{A.12}$$

Then, for any $v \geq \overline{\boldsymbol{p}}_D(\underline{c})$ and any $v' \in \text{supp}(D)$ with $v' < \psi(c)$,

$$(v' - \underline{c}) \sum_{\hat{v} \geq v'} \hat{m}^v(v') = \alpha^*(v)(v' - \underline{c})D(v') \leq \alpha^*(v)(\overline{\boldsymbol{p}}_D(\underline{c}) - \underline{c})D(\overline{\boldsymbol{p}}_D(\underline{c})) \leq (v - \underline{c})\hat{m}^v(v). \tag{A.13}$$

Together, in both of the cases above, from the constructed $\{\hat{m}^v\}_{v \in V^+}$, for each $v \in V^+$, let

$$m^v(v') := \frac{\hat{m}^v(v')}{\sum_{\hat{v} \in V} \hat{m}^v(\hat{v})}, \; \forall v' \in \text{supp}(D)$$

and let $D_v(p) := m^v([p, \overline{v}])$ for all $p \geq 0$, by (A.6), (A.7) and (A.11), in each case, $D_v \in \mathcal{D}$ for all $v \in V^+$. Now define $\sigma(c) \in \Delta(\mathcal{D})$ by

$$\sigma(D_v|c) := \sum_{v' \in V} \hat{m}^v(v'), \; \forall v \in V^+.$$

By (A.3) and (A.12), in each case, $\sigma(c) \in \mathcal{S}_D$. Furthermore, since m^v is proportional to $\hat{m}^v$ for all $v \in V^+$, (A.5), (A.9) and (A.13) ensure that in each case, $\sigma(c)$ is a $\psi(c)$-quasi-perfect segmentation for $\hat{c}$. Meanwhile, since $\hat{c} \leq c \leq \psi(c)$, $\sigma(c)$ is also a $\psi(c)$-quasi-perfect segmentation for c. Finally, since m^v is proportional to $\hat{m}^v$, (A.10) implies that for any $z \in [\underline{c}, \hat{c})$,

$$\overline{\boldsymbol{p}}_{D'}(z) \geq \overline{\boldsymbol{p}}_D(z), \; \forall D' \in \text{supp}(\sigma(c)).$$

Meanwhile, by the conclusion that $\sigma(c)$ is a $\psi(c)$-quasi-perfect segmentation for $\hat{c} \leq c$, for any $z \in [\hat{c}, c]$, since $c \leq \psi(c)$ and since $\overline{\boldsymbol{p}}_D$ is nondecreasing for any $D' \in \mathcal{D}$,

$$\overline{\boldsymbol{p}}_{D'}(z) \geq \overline{\boldsymbol{p}}_{D'}(\hat{c}) \geq \psi(c), \; \forall D' \in \text{supp}(\sigma(c)).$$

Together with the fact that ψ is nondecreasing and that $\psi \leq \overline{\boldsymbol{p}}_D$, it then follows that for any $z \in [\underline{c}, c]$ and for any $D \in \text{supp}(\sigma(c))$, $\psi(z) \leq \overline{\boldsymbol{p}}_D(z)$. Since $c \in C$ is arbitrary, this ensures that there exists a ψ-quasi-perfect scheme $\sigma : C \to \mathcal{S}_D$ that satisfies (10).

Now consider any $D_0 \in \mathcal{D}$ and any nondecreasing $\psi : C \to \mathbb{R}_+$ with $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$. I first construct a sequence of step functions $\{D_n\} \subseteq \mathcal{D}$ such that $\{D_n\} \to D_0$ and that $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_{D_n}(c)$ for all $c \in C$ and for all $n \in \mathbb{N}$. To this end, for each $n \in \mathbb{N}$, first partition V by $\underline{v} = v_0 < v_1 < \ldots < v_n = \overline{v}$ and let $V_k := [v_{k-1}, v_k]$. Then define D_n by $D_n(p) := D_0(v_k)$, for all $p \in V_k$, for all $k \in \{1, \ldots, n\}$ (i.e., by moving all the masses on interval V_k to the top v_k). By construction, it must be that $\overline{\boldsymbol{p}}_{D_n}(c) \geq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$ and for all $n \in \mathbb{N}$ and hence $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_{D_n}(c)$ for all $c \in C$ and for all $n \in \mathbb{N}$. Also, by construction, $\{D_n\} \to D_0$, as desired.

As such, for each $n \in \mathbb{N}$, there exists a ψ-quasi-perfect scheme σ_n such that for all $c \in C$,

$$\psi(z) \leq \overline{\boldsymbol{p}}_D(z)$$

for all $D \in \text{supp}(\sigma_n(c))$ and for all $z \in [\underline{c}, c]$. Furthermore, according to Helly's selection theorem, by possibly taking a subsequence,[2] $\{\sigma_n\} \to \sigma$ for some $\sigma : C \to \Delta(\mathcal{D})$. By Lemma A.1, $\sigma(c) \in \mathcal{S}$ for all $c \in C$ and σ is a ψ-quasi-perfect scheme.

It then remains to show that σ satisfies (10). To this end, fix any $c \in C$ and consider any $D \in \text{supp}(\sigma(c))$, by definition, for any $\delta > 0$, $\sigma(\mathbb{B}_\delta(D)|c) > 0$.[3] Furthermore, since $\sigma(c)$ has at most countably many atoms,

[2]See, for instance, Porter (2005) for a generalized version of Helly's selection theorem. To apply this theorem, notice that the family of functions $\{\sigma_n\}$ is of bounded p-variation due to the quasi-perfect structure. Furthermore, for any $c \in C$, the set $\text{cl}(\{\sigma_n(c)\})$ is closed in a compact metric space $\Delta(\mathcal{D})$ and hence is itself compact. As such, there exists a pointwise convergent subsequence of $\{\sigma_n\}$.

[3] $\mathbb{B}_\delta(D)$ is the δ-ball around D under the Lévy-Prokhorov metric on $\mathcal{D}$.

there exists a sequence $\{\delta_k\} \subset (0,1]$ such that $\{\delta_k\} \to 0$, $\sigma(\mathbb{B}_{\delta_k}(D)|c) > 0$ and $\sigma(\partial\mathbb{B}_{\delta_k}(D)|c) = 0$ for all $k \in \mathbb{N}$. As a result, since $\{\sigma_n(c)\} \to \sigma(c)$ under the weak-* topology, $\lim_{n\to\infty} \sigma_n(\mathbb{B}_{\delta_k}(D)|c) = \sigma(\mathbb{B}_{\delta_k}(D)|c) > 0$ for all $k \in \mathbb{N}$. Thus, for each $k \in \mathbb{N}$, there exists $n_k \in \mathbb{N}$ such that $\sigma_{n_k}(\mathbb{B}_{\delta_k}(D)|c) > 0$. Moreover, since $\sigma_n(c)$ has finite support as D_n is a step function and $\sigma_n(c) \in \mathcal{S}_{D_n}$, there must be some $D_{n_k} \in \mathbb{B}_{\delta_k}(D)$ such that $D_{n_k} \in \mathrm{supp}(\sigma_{n_k}(c))$. Notice that for the subsequence $\{n_k\}$, $\{D_{n_k}\} \to D$ and $D_{n_k} \in \mathrm{supp}(\sigma_{n_k}(c))$ for all $k \in \mathbb{N}$. As a result, since $D \mapsto \overline{\boldsymbol{p}}_D(c)$ is upper-semicontinuous (see Proposition 6 of Yang (2020a)) and since σ_{n_k} satisfies (10) for all $k \in \mathbb{N}$, for Lebesgue almost all $z \in [\underline{c}, c]$,

$$\psi(z) \leq \limsup_{k\to\infty} \overline{\boldsymbol{p}}_{D_{n_k}}(z) \leq \overline{\boldsymbol{p}}_D(z).$$

Since $c \in C$ and $D \in \mathrm{supp}(\sigma(c))$ are arbitrary, this completes the proof. ∎

Proof of Lemma 4

To prove Lemma 4, first notice that by the revelation principle, it is without loss to restrict attention to the collection of incentive feasible mechanisms $(\boldsymbol{q}, t)$, where $\boldsymbol{q}(c)$ stands for the quantity purchased for each report $c \in C$ and $t(c)$ stands for the amount of payment from the exclusive retailer to the producer for each report $c \in C$. $(\boldsymbol{q}, t)$ is incentive compatible if for any $c, c' \in C$,

$$t(c) - c\boldsymbol{q}(c) \geq t(c') - c\boldsymbol{q}(c') \tag{IC**}$$

and is individually rational if for any $c \in C$,

$$t(c) - c\boldsymbol{q}(c) \geq \pi_{D_0}(c). \tag{IR**}$$

Meanwhile, notice that given any quantity $q \in [0,1]$, it is optimal for the exclusive retailer to perfectly price discriminate the consumers with values above $D_0^{-1}(q)$.[4] Together, the exclusive retailer's problem is then to choose $(\boldsymbol{q}, t)$ to maximize

$$\int_C \left(\int_0^{\boldsymbol{q}(c)} D_0^{-1}(q)\,\mathrm{d}q - t(c) \right) G(\mathrm{d}c)$$

subject to (IC**) and (IR**).

Proof of Lemma 4. Consider the exclusive retailer's problem. First notice that by standard arguments, $(\boldsymbol{q}, t)$ is incentive compatible if and only if $\boldsymbol{q}$ is nonincreasing and there exists a constant $\bar{t}$ such that

$$t(c) = c\boldsymbol{q}(c) + \int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z - \bar{t},$$

for all $c \in C$. Moreover, any incentive compatible mechanism must give the producer indirect utility

$$\bar{t} + \int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z$$

when her cost is $c \in C$. Together, the exclusive retailer's profit maximization problem can be written as

$$\begin{aligned} &\max_{\boldsymbol{q}\in\mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} D_0^{-1}(q)\,\mathrm{d}q - \phi_G(c)\boldsymbol{q}(c) \right) G(\mathrm{d}c) \\ &\text{s.t. } \bar{\pi} + \int_c^{\bar{c}} \boldsymbol{q}(z)\,\mathrm{d}z \geq \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z. \end{aligned} \tag{A.14}$$

[4] See the formal proof of this argument in Lemma A.2 below.

Thus, by Lemma 7, the exclusive retailer's profit maximization problem is equivalent to the price-controlling data broker's revenue maximization problem. This completes the proof. ■

Proof of Lemma 5

For any nondecreasing function $\psi : C \to \mathbb{R}_+$ with $c \leq \psi(c)$ for all $c \in C$, since for any $c \in C$, $\sigma(c) \in \mathcal{S}$ is a $\psi(c)$-quasi-perfect segmentation for c, by definition,

$$\int_{\mathcal{D}} D(p)\sigma(\mathrm{d}D|c) = D_0(p), \tag{A.15}$$

for all $p \in V$, which proves assertion 1. Furthermore, since ψ is nondecreasing and is thus continuous except at countably many points, $\sigma : C \to \Delta(\mathcal{D})$ is measurable, which establishes assertion 2. For assertion 3, notice that for any $c \in C$, since $\sigma(c) \in \mathcal{S}$ is a $\psi(c)$-quasi-perfect segmentation for c, for any $D \in \operatorname{supp}(\sigma(c))$ such that $D(\overline{\boldsymbol{p}}_D(c)) > 0$,

$$D(\overline{\boldsymbol{p}}_D(c)) = D(\max(\operatorname{supp}(D))) = D(\psi(c)),$$

while for any $D \in \operatorname{supp}(\sigma(c))$ such that $D(\overline{\boldsymbol{p}}_D(c)) = 0$, it must be that $D(\psi(c)) = 0$ as well. Therefore,

$$\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) = \int_{\mathcal{D}} D(\psi(c))\sigma(\mathrm{d}D|c) = D_0(\psi(c)),$$

where the last equality follows from (A.15). This proves assertion 3. Finally, to prove assertion 4, consider any $c \in C$. First notice that if $D_0(c) = 0$, then assertion 4 clearly holds as both sides would be zero. Now suppose that $D_0(c) > 0$. The fact that $\sigma(c) \in \mathcal{S}$ is a $\psi(c)$-quasi-perfect segmentation for c ensures that $D_0(\psi(c)) > 0$. Then, for any $v \in [\psi(c), \overline{v}]$, let

$$H(v) := \sigma(\{D \in \mathcal{D} : \max(\operatorname{supp}(D)) \leq v\}|c).$$

Since $\sigma(c)$ is a probability measure, H is nondecreasing and right-continuous and hence induces a Borel measure μ_H on $[\psi(c), \overline{v}]$. Meanwhile, for any measurable sets $A, B \subseteq [\psi(c), \overline{v}]$, define

$$K(A|B) := \int_{\{D \in \mathcal{D} : \max(\operatorname{supp}(D)) \in A\}} m^D(B)\sigma(\mathrm{d}D|c).$$

Notice that for any measurable set $B \subseteq [\psi(c), \overline{v}]$, $K(\cdot|B)$ is a measure and is absolutely continuous with respect to μ_H and hence there exists a (essentially) unique Radon-Nikodym derivative $v \mapsto m^v(B)$ such that for any measurable $A \subseteq [\psi(c), \overline{v}]$,

$$K(A|B) = \int_{v \in A} m^v(B)H(\mathrm{d}v). \tag{A.16}$$

In particular, by definition of K and by (A.15), for any measurable set $B \subseteq [\psi(c), \overline{v}]$,

$$\int_{[\psi(c),\overline{v}]} m^v(B)H(\mathrm{d}v) = K([\psi(c), \overline{v}]|B) = \int_{\mathcal{D}} m^D(B)\sigma(\mathrm{d}D|c) = m^0(B). \tag{A.17}$$

Moreover, since for any measurable set $A \subseteq [\psi(c), \overline{v}]$, $K(A|\cdot)$ is a measure on $[\psi(c), \overline{v}]$ and thus m^v is also a measure on $[\psi(c), \overline{v}]$ for μ_H-almost all $v \in [\psi(c), \overline{v}]$. Furthermore, since $\sigma(c) \in \mathcal{S}$ is a $\psi(c)$-quasi-perfect segmentation for c, for any measurable sets $A, B \subseteq [\psi(c), \overline{c}]$,

$$K(A|B) = m^0(A \cap B) = K(B|A)$$

and hence, for any measurable sets $A, B \subseteq [\psi(c), \overline{v}]$,

$$\int_A m^v(B) H(\mathrm{d}v) = \int_B m^v(A) H(\mathrm{d}v). \tag{A.18}$$

As a result,

$$\begin{aligned}
\int_{\mathcal{D}} \overline{\boldsymbol{p}}_D(c) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) &= \int_{\mathcal{D}} \overline{\boldsymbol{p}}_D(c) D(\psi(c)) \sigma(\mathrm{d}D|c) \\
&= \int_{\mathcal{D}} \max(\mathrm{supp}(D)) m^D([\psi(c), \overline{v}]) \sigma(\mathrm{d}D|c) \\
&= \int_{[\psi(c),\overline{v}]} v K(\mathrm{d}v | [\psi(c), \overline{v}]) \\
&= \int_{[\psi(c),\overline{v}]} v m^v([\psi(c), \overline{v}]) H(\mathrm{d}v) \\
&= \int_{v \in [\psi(c),\overline{v}]} \int_{v' \in [\psi(c),\overline{v}]} v m^v(\mathrm{d}v') H(\mathrm{d}v) \\
&= \int_{v \in [\psi(c),\overline{v}]} v \left(\int_{v' \in [\psi(c),\overline{v}]} m^{v'}(\mathrm{d}v) H(\mathrm{d}v') \right) \\
&= \int_{[\psi(c),\overline{v}]} v D_0(\mathrm{d}v),
\end{aligned}$$

where the second equality follows from the fact that $\sigma(c)$ is a $\psi(c)$-quasi-perfect segmentation for c, the third equality follows from the definition of K, the fourth equality follows from (A.16), the sixth equality follows from (A.18), and the last equality follows from (A.17). This completes the proof. ■

Proof of Lemma 7

By Lemma 6, the producer's expected profit under an incentive compatible mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ of the price-controlling data broker can be written as

$$U(c) = U(\overline{c}) + \int_c^{\overline{c}} \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, z) \right) \sigma(\mathrm{d}D|z) \, \mathrm{d}z.$$

As such, an incentive compatible mechanism is individually rational if and only if

$$U(\overline{c}) + \int_c^{\overline{c}} \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, z) \right) \sigma(\mathrm{d}D|z) \, \mathrm{d}z \geq \bar{\pi} + \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z)) \, \mathrm{d}z.$$

Also, for any incentive compatible mechanism $(\sigma, \tau, \boldsymbol{\gamma})$, the price-controlling data broker's expected revenue can be written as

$$\mathbb{E}[\tau(c)] = \int_C \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \phi_G(c)) D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - U(\overline{c}).$$

Therefore, the price-controlling data broker's revenue maximization problem can be written as

$$\begin{aligned}
&\sup_{\sigma, \boldsymbol{\gamma}} \int_C \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \phi_G(c)) D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{\pi} \\
&\text{s.t. } c \mapsto \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}D|c) \text{ is nonincreasing,} \\
&\qquad \int_c^{\overline{c}} \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, z) \right) \sigma(\mathrm{d}D|z) \, \mathrm{d}z \geq \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z)) \, \mathrm{d}z, \ \forall c \in C,
\end{aligned}$$

where the supremum is taken over all segmentation schemes $\sigma : C \to \mathcal{S}$ and all measurable function $\boldsymbol{\gamma}$ that maps from C to the collection of transition kernels from $\mathcal{D}$ to $\Delta(\mathbb{R}_+)$.

In what follows, let Γ be the collection of transition kernels that maps from $\mathcal{D}$ to $\Delta(\mathbb{R}_+)$. Let $s^{\mathrm{VR}} \in \mathcal{S}$ denote the value-revealing segmentation and let $\sigma^{\mathrm{VR}} : C \to \mathcal{S}$ be the segmentation scheme such that $\sigma^{\mathrm{VR}}(c) = s^{\mathrm{VR}}$ for all $c \in C$. Furthermore, for any $q \in [0,1]$, let $\rho_q := D_0^{-1}(q)$, where D_0^{-1} is defined by (16). Notice that by definition of D_0^{-1},

$$q \in [D_0(\rho_q^+), D_0(\rho_q)].$$

If $D_0(\rho_q) = D_0(\rho_q^+)$, then let $\tilde{\gamma}^q : V \to \Delta(\mathbb{R}_+)$ be defined as

$$\tilde{\gamma}^q(\cdot|v) := \delta_{\{v\}},\ \forall v \in V.$$

Meanwhile, if $D_0(\rho_q) > D_0(\rho_q^+)$, then define $\tilde{\gamma}^q : V \to \Delta(\mathbb{R}_+)$ as

$$\tilde{\gamma}^q(\cdot|v) := \begin{cases} \delta_{\{v\}}, & \text{if } v \neq \rho_q \\ \frac{q - D_0(\rho_q^+)}{D_0(\rho_q) - D_0(\rho_q^+)}\delta_{\{v\}} + \frac{D_0(\rho_q) - q}{D_0(\rho_q) - D_0(\rho_q^+)}\delta_{\{\bar{v}\}}, & \text{if } v = \rho_q \end{cases},\ \forall v \in V.$$

Finally, let $\gamma^q \in \Gamma$ be defined as

$$\gamma^q(A|D) := \int_V \tilde{\gamma}^q(A|v) D(\mathrm{d}v),$$

for any measurable $A \subseteq V$ and for any $D \in \mathcal{D}$. By construction, under the segmentation s^{VR} and the randomized price γ^q, all the consumers with values above the $D_0^{-1}(q)$ buy the product by paying exactly their values while the other consumers do not buy, so that the traded quantity is exactly q (if the consumers with value $v = D_0^{-1}(q)$ has a mass, then some of them buy and some of them do not buy, so that the total quantity sold is exactly q). That is,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \gamma^q(\mathrm{d}p|D) \right) s^{\mathrm{VR}}(\mathrm{d}D) = q. \tag{A.19}$$

With this notation, I now introduce an auxiliary lemma.

Lemma A.2. *For any $q \in [0,1]$, let $\bar{R}(q)$ be the value of the maximization problem*

$$\begin{aligned} \sup_{s \in \mathcal{S}, \gamma \in \Gamma} & \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} pD(p)\gamma(\mathrm{d}p) \right) s(\mathrm{d}D) \\ \text{s.t.} & \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p)\gamma(\mathrm{d}p) \right) s(\mathrm{d}D) \leq q. \end{aligned} \tag{A.20}$$

Then

$$\bar{R}(q) = \int_0^q D_0^{-1}(y)\, \mathrm{d}y,$$

where D_0^{-1} is defined by (16). Moreover, $(s^{\mathrm{VR}}, \gamma^q)$ is a solution of (A.20).

Proof. Consider the dual problem of (A.20). That is, for any $\nu \geq 0$, let

$$\begin{aligned} d(\nu) :=& \sup_{s \in \mathcal{S}, \gamma \in \Gamma} \left[\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} pD(p)\gamma(\mathrm{d}p|D) \right) s(\mathrm{d}D) + \nu \left(q - \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} pD(p)\gamma(\mathrm{d}p|D) \right) s(\mathrm{d}D) \right) \right] \\ =& \sup_{s \in \mathcal{S}, \gamma \in \Gamma} \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \nu) D(p)\gamma(\mathrm{d}p|D) \right) s(\mathrm{d}D) + \nu q. \end{aligned}$$

Clearly, $d(\nu) \geq \bar{R}(q)$ for any $\nu \geq 0$. Thus, by weak duality, to solve (A.20), it suffices to find ν^* and (s^*, γ^*) such that (s^*, γ^*) is feasible in the primal problem (A.20), (s^*, γ^*) solves the dual problem

$$\sup_{s \in \mathcal{S}, \gamma \in \Gamma} \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \nu^*) D(p) \gamma(\mathrm{d}p|D) \right) s(\mathrm{d}D) \tag{A.21}$$

and that the complementary slackness condition

$$\nu^* \left[q - \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} p D(p) \gamma^*(\mathrm{d}p|D) \right) s^*(\mathrm{d}D) \right] = 0 \tag{A.22}$$

holds. Since this would imply that

$$\bar{R}(q) \leq d^* = \inf_{\lambda \geq 0} d(\lambda) \leq d^*(\nu^*) = \bar{R}(q)$$

and hence (s^*, γ^*) must be a solution to (A.20).

To this end, let

$$\nu^* := D_0^{-1}(q).$$

and consider the pair $(s^{\mathrm{VR}}, \gamma^q)$. Notice that by definition, $(s^{\mathrm{VR}}, \gamma^q)$ perfectly price-discriminates all the consumers with $v > \nu^*$ and does not sell to any consumers with $v < \nu^*$. Therefore,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \nu^*) D(p) \gamma(\mathrm{d}p|D) \right) s^{\mathrm{VR}}(\mathrm{d}D) = \int_V (v - \nu^*)^+ D_0(\mathrm{d}v)$$

Furthermore, notice that for any $s \in \mathcal{S}$ and any $\gamma \in \Gamma$

$$\begin{aligned} &\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \nu^*) D(p) \gamma(\mathrm{d}p|D) \right) s(\mathrm{d}D) \\ \leq &\int_{\mathcal{D}} \max_{p \in \mathbb{R}_+} (p - \nu^*) D(p) s(\mathrm{d}D) \\ \leq &\int_V (v - \nu^*)^+ D_0(\mathrm{d}v). \end{aligned}$$

Therefore, $(s^{\mathrm{VR}}, \gamma^q)$ solves the dual problem (A.21). Meanwhile, by (A.19),

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \gamma^q(\mathrm{d}p|D) \right) s^{\mathrm{VR}}(\mathrm{d}D) = q.$$

Thus, the complementary slackness condition (A.22) holds and $(s^{\mathrm{VR}}, \gamma^q)$ is feasible in the primal problem (A.20). Together, $(s^{\mathrm{VR}}, \gamma^q)$ is a solution to the primal problem (A.20).

Finally, notice that by the definition of D_0^{-1} and $(s^{\mathrm{VR}}, \gamma^q)$,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} p D(p) \gamma^q(\mathrm{d}p|D) \right) s^{\mathrm{VR}}(\mathrm{d}D) = \int_0^q D_0^{-1}(y) \, \mathrm{d}y.$$

This completes the proof. ■

With Lemma A.2, we can now begin the proof of Lemma 7.

Proof of Lemma 7. Consider any incentive feasible mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ for the price-controlling data broker, I will first show that there exists $\boldsymbol{q} : C \to [0,1]$ such that the mechanism $(\sigma^{\mathrm{VR}}, \tau^{\boldsymbol{q}}, \boldsymbol{\gamma}^{\boldsymbol{q}})$ generates weakly higher revenue for price-controlling data broker and is incentive feasible, where

$$\boldsymbol{\gamma}^{\boldsymbol{q}}(c) := \gamma^{\boldsymbol{q}(c)}, \, \forall\, c \in C$$

and $\tau^{\boldsymbol{q}}$ is the transfer determined by $(\sigma^{\mathrm{VR}}, \boldsymbol{\gamma}^{\boldsymbol{q}})$ according to Lemma 6, with the constant being chosen so that the producer with cost $\bar{c}$ obtains profit $\bar{\pi}$ when reporting truthfully. Next, I will show that maximizing revenue across the family of incentive feasible mechanisms $(\sigma^{\mathrm{VR}}, \tau^{\boldsymbol{q}}, \boldsymbol{\gamma}^{\boldsymbol{q}})$ is equivalent to solving (17). Finally, the existence of the optimal mechanism can then be ensured by the existence of the solution of (17), which will be proved at the end.

To this end, for any $c \in C$, let

$$\boldsymbol{q}(c) := \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}D | c).$$

By Lemma 6, incentive compatibility of $(\sigma, \tau, \boldsymbol{\gamma})$ implies that $\boldsymbol{q} : C \to [0,1]$ is nonincreasing and, by (A.19), for any $c \in C$,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}^{\boldsymbol{q}}(\mathrm{d}p | D, c) \right) \sigma^{\mathrm{VR}}(\mathrm{d}D | c) = \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \gamma^{\boldsymbol{q}(c)}(\mathrm{d}p | D) \right) s^{\mathrm{VR}}(\mathrm{d}D) = \boldsymbol{q}(c).$$

Thus, by Lemma A.2, $(\sigma^{\mathrm{VR}}(c), \boldsymbol{\gamma}^{\boldsymbol{q}}(c))$ solves the problem (A.20) with the quantity constraint being $\boldsymbol{q}(c)$ and hence, since $(\sigma(c), \boldsymbol{\gamma}(c))$ is also feasible in this problem,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} p D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}D | c) \leq \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} p D(p) \boldsymbol{\gamma}^{\boldsymbol{q}}(\mathrm{d}p | D, c) \right) \sigma^{\mathrm{VR}}(\mathrm{d}D | c) = \bar{R}(\boldsymbol{q}(c)). \tag{A.23}$$

As a result,

$$\begin{aligned}
&\int_C \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \phi_G(c)) D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}x | c) \right) G(\mathrm{d}c) \\
&= \int_C \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} p D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}x | c) \right) G(\mathrm{d}c) - \int_C \phi_G(c) \boldsymbol{q}(c) G(\mathrm{d}c) \\
&\leq \int_C (\bar{R}(\boldsymbol{q}(c)) - \phi_G(c) \boldsymbol{q}(c)) G(\mathrm{d}c) \\
&= \int_C \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} p D(p) \boldsymbol{\gamma}^{\boldsymbol{q}}(\mathrm{d}p | D, c) \right) \sigma^{\mathrm{VR}}(\mathrm{d}x | c) \right) G(\mathrm{d}c) - \int_C \phi_G(c) \boldsymbol{q}(c) G(\mathrm{d}c) \\
&= \int_C \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \phi_G(c)) D(p) \boldsymbol{\gamma}^{\boldsymbol{q}}(\mathrm{d}p | D, c) \right) \sigma^{\mathrm{VR}}(\mathrm{d}x | c) \right) G(\mathrm{d}c),
\end{aligned}$$

where the first and the third equalities follows from the definition of $\boldsymbol{q}(c)$ and from (A.19), and the inequality and the second equality follows from (A.23). Moreover, by (A.19), since $\boldsymbol{q}$ is nonincreasing, the function

$$c \mapsto \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}^{\boldsymbol{q}}(\mathrm{d}p | D, c) \right) \sigma^{\mathrm{VR}}(\mathrm{d}D | c)$$

is nonincreasing. Together with Lemma 6 and individual rationality of $(\sigma,\tau,\boldsymbol{\gamma})$, for any $c\in C$,

$$\begin{aligned}\int_c^{\overline{c}}\left(\int_{\mathcal{D}}\left(\int_{\mathbb{R}_+}D(p)\boldsymbol{\gamma}^{\boldsymbol{q}}(\mathrm{d}p|D,z)\right)\sigma^{\mathrm{VR}}(\mathrm{d}D|z)\right)\mathrm{d}z &= \int_c^{c'}\boldsymbol{q}(z)\,\mathrm{d}z\\ &= \int_c^{\overline{c}}\left(\int_{\mathcal{D}}\left(\int_{\mathbb{R}_+}D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,z)\right)\sigma(\mathrm{d}D|z)\right)\mathrm{d}z\\ &\geq \int_c^{\overline{c}}D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\end{aligned}$$

these imply that $(\sigma^{\mathrm{VR}},\tau^q,\boldsymbol{\gamma}^{\boldsymbol{q}})$ is incentive feasible.

Now notice that by (A.19) and Lemma A.2, for any $\boldsymbol{q}:C\to[0,1]$ and for any $c\in C$,

$$\int_{\mathcal{D}}\left(\int_{\mathbb{R}_+}(p-\phi_G(c))D(p)\boldsymbol{\gamma}^{\boldsymbol{q}}(\mathrm{d}p|D,c)\right)\sigma^{\mathrm{VR}}(\mathrm{d}D|c)=\bar{R}(\boldsymbol{q}(c))=\int_0^{\boldsymbol{q}(c)}D_0^{-1}(q)\,\mathrm{d}q.$$

Meanwhile, by (A.19) and by Lemma A.2, $(\sigma^{\mathrm{VR}},\tau^{\boldsymbol{q}},\boldsymbol{\gamma}^{\boldsymbol{q}})$ is incentive feasible if and only if $\boldsymbol{q}$ is nonincreasing and

$$\int_c^{\overline{c}}\boldsymbol{q}(z)\,\mathrm{d}z\geq\int_c^{\overline{c}}D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c\in C.$$

Therefore, maximizing revenue among all incentive feasible mechanism is equivalent to solving (17).

Finally, notice that for the maximization problem (17), endow the set of nonincreasing functions with the L^1 norm. Helly's selection theorem and the Lebesgue dominated convergence theorem then imply that this set is compact. Moreover, for any sequence $\{\boldsymbol{q}_n\}\subset\mathcal{Q}$ such that $\{\boldsymbol{q}_n\}\to\boldsymbol{q}$, consider any subsequence $\{\boldsymbol{q}_{n_k}\}$ of $\{\boldsymbol{q}_n\}$, by the Riesz-Fischer theorem, there exists a further subsequence $\{\boldsymbol{q}_{n_{k,l}}\}$ such that $\{\boldsymbol{q}_{n_{k,l}}\}\to\boldsymbol{q}$ pointwise. By the dominated convergence theorem,

$$\lim_{l\to\infty}\int_C\left(\int_0^{\boldsymbol{q}_{n_{k,l}}(c)}D_0^{-1}(q)-\phi_G(c)\,\mathrm{d}q\right)G(\mathrm{d}c)=\int_C\left(\int_0^{\boldsymbol{q}(c)}D_0^{-1}(q)-\phi_G(c)\,\mathrm{d}q\right)G(\mathrm{d}c)$$

and

$$\lim_{l\to\infty}\int_c^{\overline{c}}\boldsymbol{q}_{n_{k,l}}(z)\,\mathrm{d}z=\int_c^{\overline{c}}\boldsymbol{q}(z)\,\mathrm{d}z,\ \forall c\in C.$$

Therefore,

$$\lim_{n\to\infty}\int_C\left(\int_0^{\boldsymbol{q}_n(c)}D_0^{-1}(q)-\phi_G(c)\,\mathrm{d}q\right)G(\mathrm{d}c)=\int_C\left(\int_0^{\boldsymbol{q}(c)}D_0^{-1}(q)-\phi_G(c)\,\mathrm{d}q\right)G(\mathrm{d}c)$$

and

$$\lim_{n\to\infty}\int_c^{\overline{c}}\boldsymbol{q}_n(z)\,\mathrm{d}z=\int_c^{\overline{c}}\boldsymbol{q}(z)\,\mathrm{d}z,\ \forall c\in C.$$

Together, the feasible set of (17) is compact and the objective is continuous (under the L^1 norm) and hence the solution must exist. This completes the proof. ■

Proof of Theorem 8

For each $\theta \in \Theta$, write $\text{supp}(D_\theta)$ as $[l(\theta), u(\theta)]$. Also, for any $p \in V$, let θ_p be the unique θ such that $p \in (l(\theta), u(\theta)]$. Notice that since $\{(l(\theta), u(\theta))\}_{\theta\in\Theta}$ is disjoint, for any $\beta \in \Delta(\Theta)$, any $\theta \in \Theta$, and any $p \in \text{supp}(D_\beta)$,

$$D_\beta(p) = \sum_{\{\theta' : u(\theta') \geq u(\theta_p)\}} D_{\theta'}(p)\beta(\theta') = D_\theta(p)\beta(\theta) + \sum_{\{\theta' : u(\theta') > u(\theta_p)\}} \beta(\theta').$$

In particular, different prices set in $\text{supp}(D_\theta)$ does not affect the probability of trade through $\theta' \in \Theta$ such that $u(\theta') > u(\theta)$.

As a result, the construction in the proof of Lemma 3 is still valid, with the demands being replaced by D_β. Specifically, for any $\beta \in \Delta(\Theta)$ and any $c \in C$, there exists $\{\beta_i\}_{i=1}^n \subseteq \Delta(\Theta)$ such that:

1. $\beta \in \text{co}(\{\beta_i\}_{i=1}^n)$.

2. For each $i \in \{1, \dots, n\}$, the set
$$\{\theta \in \text{supp}(\beta_i) | u(\theta) \geq \overline{\boldsymbol{p}}_{D_{\beta_i}}(c)\}$$
is nonempty and is a singleton.

3. For each $i \in \{1, \dots, n\}$,
$$\boldsymbol{P}_{D_{\beta_i}}(c) \bigcap \text{supp}(D_{\bar{\theta}_{\beta_i}}) \neq \emptyset,$$
where $\bar{\theta}_{\beta_i} := \max\{u(\theta) : \theta \in \text{supp}(\beta_i)\}$.

4. For each $i \in \{1, \dots, n\}$ and any $z \in [\underline{c}, c]$,
$$\overline{\boldsymbol{p}}_{D_{\beta_i}}(z) \geq \overline{\boldsymbol{p}}_{D_\beta}(z).$$

This further implies that, by Lemma 5, and by the same argument as in the proof of Lemma 2, for any $\beta \in \Delta(\Theta)$, there exists $\sigma^\beta : C \to \Delta(\Delta(\Theta))$ such that

5. For any $c \in C$,
$$\begin{aligned}
&\sum_{\beta' \in \text{supp}(\sigma^\beta(c))} (\overline{\boldsymbol{p}}_{D_{\beta'}}(c) - \overline{\boldsymbol{p}}_{D_\beta}(c)) D_{\beta'}(\overline{\boldsymbol{p}}_{D'_\beta}(c))\sigma^\beta(\beta'|c) \\
=& \sum_{\{\theta : u(\theta) \geq \theta(\overline{\boldsymbol{p}}_{D_\beta}(c))\}} (\overline{\boldsymbol{p}}_{D_\theta}(c) - \overline{\boldsymbol{p}}_{D_\beta}(c)) D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c))\beta(\theta).
\end{aligned}$$

6. For any $c \in C$, $\sum_{\beta' \in \text{supp}(\sigma^\beta(c))} D_{\beta'}(\overline{\boldsymbol{p}}_{D_{\beta'}}(c))\sigma^\beta(\beta'|c) = D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c))$.

7. For any $c \in C$, $\sum_{\beta' \in \text{supp}(\sigma^\beta(c))} \beta' \sigma^\beta(\beta'|c) = \beta$.

8. For any $\beta' \in \text{supp}(\sigma^\beta(c'))$,
$$\overline{\boldsymbol{p}}_{D_{\beta'}}(c) \geq \overline{\boldsymbol{p}}_{D_\beta}(c),$$
for any $c, c' \in C$ such that $c < c'$ and
$$\sum_{\beta' \in \text{supp}(\sigma^\beta(c))} D_{\beta'}(\overline{\boldsymbol{p}}_{D_{\beta'}}(c))\sigma^\beta(\beta'|c) \geq D(\overline{\boldsymbol{p}}_{D_\beta}(c)),$$
for any $c, c' \in C$ such that $c > c'$.

Now consider any mechanism (σ,τ). Suppose that there is a selection $\boldsymbol{p}\in\boldsymbol{P}$ and a positive G-measure of c such that there exists some $\beta\in\operatorname{supp}(\sigma(c))$ and with

$$\{\theta\in\operatorname{supp}(\beta):u(\theta)>u(\theta_{\boldsymbol{p}_{D_\beta}(c)})\}\neq\emptyset. \tag{A.24}$$

Then, for such $\boldsymbol{p}\in\boldsymbol{P}$, $c\in C$ and $\beta\in\operatorname{supp}(\sigma(c))$, there exists $\sigma^\beta(c)\in\Delta(\Delta(\Theta))$ such that assertions 5 through 8 above hold. In particular, assertions 5 and 6 imply that

$$\begin{aligned}
&\sum_{\beta'\in\operatorname{supp}(\sigma^\beta(c))}(\overline{\boldsymbol{p}}_{D_{\beta'}}(c)-\phi_G(c))D_{\beta'}(\overline{\boldsymbol{p}}_{D_{\beta'}}(c))\sigma^\beta(\beta'|c)\\
=&\sum_{\beta'\in\operatorname{supp}(\sigma^\beta(c))}(\overline{\boldsymbol{p}}_{D_{\beta'}}(c)-\overline{\boldsymbol{p}}_{D_\beta}(c))D_{\beta'}(\overline{\boldsymbol{p}}_{D_{\beta'}}(c))\sigma^\beta(\beta'|c)+(\overline{\boldsymbol{p}}_{D_\beta}(c)-\phi_G(c))D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c))\\
\geq&\sum_{\{\theta:u(\theta)\geq u(\theta_{\overline{\boldsymbol{p}}_{D_\beta}(c)})\}}(\overline{\boldsymbol{p}}_{D_\theta}(c)-\overline{\boldsymbol{p}}_{D_\beta}(c))D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c))\beta(\theta)+(\boldsymbol{p}_{D_\beta}(c)-\phi_G(c))D_\beta(\boldsymbol{p}_{D_\beta}(c))\\
>&(\boldsymbol{p}_{D_\beta}(c)-\phi_G(c))D_\beta(\boldsymbol{p}_{D_\beta}(c)),
\end{aligned}$$

where the second equality follows from 5 and 6 and the inequality is strict due to (A.24).

As such, together with assertion 7, $\sigma^\beta(c)$ induces another segmentation $\hat{\sigma}(c)$ through

$$\hat{\sigma}(\hat{\beta}|c):=\sum_{\beta\in\operatorname{supp}(\sigma(c))}\sigma^\beta(\hat{\beta}|c)\sigma(\beta|c),\ \forall\hat{\beta}\in\bigcup_{\beta\in\operatorname{supp}(\sigma(c))}\operatorname{supp}(\sigma^\beta(c))$$

As (A.24) holds with positive G-measure of $c\in C$, the induced segmentation scheme $\hat{\sigma}:C\to\Delta(\Delta(\Theta))$ strictly improves the data broker's revenue. Finally, by the revenue equivalence formula,[5] assertions 6 and 8 above and Lemma 1 ensure that there exists a transfer $\hat{\tau}$ such that $(\hat{\sigma},\hat{\tau})$ is incentive compatible and individually rational and strictly improves the data broker's revenue.

Together, any optimal mechanism (σ,τ) must be such that for G-almost all $c\in C$ and for all $\beta\in\operatorname{supp}(\sigma(c))$,

$$\{\theta\in\operatorname{supp}(\beta):u(\theta)>u(\theta_{\boldsymbol{p}_{D_\beta}(c)})\}=\emptyset.$$

which, together with the fact that $\sum_{\beta\in\operatorname{supp}(\sigma(c))}\sigma(\beta|c)=\beta_0$ for all $c\in C$, implies that for G-almost all $c\in C$, the consumer surplus must be lower than the case when all the information about θ is revealed. ■

Proof of Theorem 9

To prove Theorem 9, I first introduce two useful lemmas.

Lemma A.3. *For any $c\in C$, any $\nu\geq c$ and any segmentation $s\in\Delta(\Delta(\Theta))$,*

$$\int_{\Delta(\Theta)}(\overline{\boldsymbol{p}}_{D_\beta}(c)-\nu)D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c))s(\mathrm{d}\beta)\leq\int_{\{\theta:\overline{\boldsymbol{p}}_{D_\theta}(c)\geq\nu\}}(\overline{\boldsymbol{p}}_{D_\theta}(c)-\nu)D_\theta(\overline{\boldsymbol{p}}_{D_\theta})\beta_0(\mathrm{d}\theta),$$

Proof. I first show that for any segmentation $s\in\Delta(\Delta(\Theta))$, there must exist another segmentation $\hat{s}$ such that for any $\beta\in\operatorname{supp}(\hat{s})$, either $\beta(\{\theta:u(\theta)<c\})=1$ or $\overline{\boldsymbol{p}}_{D_\beta}(c)=\overline{\boldsymbol{p}}_{D_{\bar{\theta}_\beta}}(c)$ and

$$\int_{\Delta(\Theta)}(\overline{\boldsymbol{p}}_{D_\beta}-\nu)D(\overline{\boldsymbol{p}}_\beta(c))s(\mathrm{d}\beta)\leq\int_{\Delta(\Theta)}(\overline{\boldsymbol{p}}_{D_\beta}-\nu)D(\overline{\boldsymbol{p}}_\beta(c))\hat{s}(\mathrm{d}\beta),$$

[5] See more detailed arguments in the proof of Theorem S1 of an earlier version of this paper (Yang, 2020b)

where $\bar{\theta}_\beta := \max(\text{supp}(\beta))$. Indeed, consider any segmentation $s \in \Delta(\Delta(\Theta))$. For any $\beta \in \text{supp}(s)$, by definition, it must be that $\text{supp}(\beta) \cap \{\theta \in \Theta : u(\theta) \geq \overline{\boldsymbol{p}}_{D_\beta}(c)\} \neq \emptyset$. Now define $\hat{\beta}^\theta$ as

$$\hat{\beta}^\theta(\theta') := \begin{cases} \beta(\theta), & \text{if } \theta' \leq \overline{\boldsymbol{p}}_{D_\beta}(c) \\ \sum_{\{\hat{\theta}: u(\hat{\theta}) \geq \overline{\boldsymbol{p}}_{D_\beta}(c)\}} \beta(\hat{\theta}), & \text{if } \theta' = \theta \\ 0, & \text{otherwise} \end{cases},$$

for any $\theta' \in \text{supp}(\beta)$ and for any $\theta \in \text{supp}(\beta)$ with $u(\theta) \geq \overline{\boldsymbol{p}}_{D_\beta}(c)$. Notice that by construction, $\beta \in \text{co}(\{\hat{\beta}^\theta\}_{\theta \geq \overline{\boldsymbol{p}}_{D_\beta}(c)})$ and hence there exists $K^\beta \in \Delta(\Delta(\Theta))$ such that $\beta = \sum_{\hat{\beta}} K^\beta(\hat{\beta})$. Therefore, by splitting every β according to K^β, and by the same arguments as in the proof of Lemma 3, the resulting segmentation $\hat{s} \in \Delta(\Theta)$ must be such that for any $\hat{\beta} \in \text{supp}(\hat{s})$, $\overline{\boldsymbol{p}}_{D_{\hat{\beta}}}(c)$ is in the interval described by $\max(\text{supp}(\hat{\beta}))$. Moreover, since $\{(l(\theta), u(\theta))\}_{\theta \in \Theta}$ is disjoint, it follows that $\overline{\boldsymbol{p}}_{D_{\bar{\theta}_{\hat{\beta}}}}(c) = \overline{\boldsymbol{p}}_{D_{\hat{\beta}}}(c)$. Furthermore, since for any $\beta \in \text{supp}(s)$,

$$\begin{aligned}(\overline{\boldsymbol{p}}_{D_\beta}(c) - \nu) D_\beta(\overline{\boldsymbol{p}}_\beta(c)) =& (\overline{\boldsymbol{p}}_{D_\beta}(c) - \nu) \sum_{\{\theta: u(\theta) \geq \overline{\boldsymbol{p}}_{D_\beta}(c)\}} D_\theta(\overline{\boldsymbol{p}}_\beta(c)) \beta(\theta) \\ \leq & \sum_{\{\theta: u(\theta) \geq \overline{\boldsymbol{p}}_{D_\beta}(c)\}} (\overline{\boldsymbol{p}}_{D_\theta}(c) - \nu) D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) \beta(\theta) \\ = & \sum_{\hat{\beta} \in \text{supp}(K^\beta)} (\overline{\boldsymbol{p}}_{D_{\hat{\beta}}}(c) - \nu) D_{\hat{\beta}}(\overline{\boldsymbol{p}}_{D_{\hat{\beta}}}(c)) K^\beta(\hat{\beta}).\end{aligned}$$

As a result, since $\hat{s}(\hat{\beta}) = \sum_\beta K^\beta(\hat{\beta}) s(\beta)$, it then follows that

$$\int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \nu) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c)) s(\mathrm{d}\beta) \leq \int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \nu) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c)) \hat{s}(\mathrm{d}\beta).$$

Finally, since for any $\beta \in \text{supp}(\hat{s})$, either $\beta(\{\theta : u(\theta) < c\}) = 1$ or $\overline{\boldsymbol{p}}_{D_\beta}(c) = \overline{\boldsymbol{p}}_{D_{\bar{\theta}_\beta}}(c)$, it must be that

$$\int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \nu) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c)) \hat{s}(\mathrm{d}\beta) \leq \int_{\{\theta: \overline{\boldsymbol{p}}_{D_\theta}(c) \geq \nu\}} (\overline{\boldsymbol{p}}_{D_\theta}(c) - \nu) D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) \beta_0(\mathrm{d}\theta),$$

as desired. ■

Lemma A.4. *Suppose that D_0 is regular. For any $c \in C$ and for any $\nu \in [c, \overline{\boldsymbol{p}}_0(c)]$,*

$$D_0(\overline{\boldsymbol{p}}_0(c)) \leq \sum_{\{\theta: u(\theta) \geq \nu\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) \beta_0(\theta) \tag{A.25}$$

and

$$D_0(\overline{\boldsymbol{p}}_0(c)) \geq \sum_{\{\theta: l(\theta) \geq \overline{\boldsymbol{p}}_0(c)\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) \beta_0(\theta). \tag{A.26}$$

Proof. Consider any $c \in C$. I first show that for any $\theta \in \Theta$ such that $l(\theta) \geq \overline{\boldsymbol{p}}_0(c)$, $\overline{\boldsymbol{p}}_{D_\theta}(c) = l(\theta)$. Indeed, since D_0 is regular, for any $\theta \in \Theta$ such that $l(\theta) \geq \overline{\boldsymbol{p}}_0(c)$ and for any $p \in (l(\theta), u(\theta)]$,

$$\begin{aligned}
&(p-c)\left[D_{\theta_p}(p)\beta_0(\theta_p) + \sum_{\{\theta':l(\theta')\geq p\}} \beta_0(\theta')\right] \\
=&(p-c)\sum_{\{\theta':u(\theta')\geq p\}} D_{\theta'}(p)\beta_0(\theta') \\
=&(p-c)D_0(p) \\
\leq&(l(\theta)-c)D_0(l(\theta)) \\
=&(l(\theta)-c)\left[\sum_{\{\theta':u(\theta')\geq l(\theta)\}} D_{\theta'}(l(\theta))\beta_0(\theta')\right] \\
=&(l(\theta)-c)\left[D_\theta(l(\theta))\beta_0(\theta) + \sum_{\{\theta':l(\theta')\geq l(\theta)\}} \beta_0(\theta')\right].
\end{aligned}$$

As such, since $p \in (l(\theta), u(\theta)]$ and $u(\theta_p) = u(\theta)$, it must be that

$$(p-c)D_\theta(p) < (l(\theta)-c)D_\theta(l(\theta)),$$

which then implies that $\overline{\boldsymbol{p}}_{D_\theta}(c) = l(\theta)$.

Now, I show that $\overline{\boldsymbol{p}}_0(c) \geq \widehat{\boldsymbol{p}}_0(c) := \overline{\boldsymbol{p}}_{D_{\theta_{\overline{\boldsymbol{p}}_0(c)}}}(c)$. Indeed, by definitions,

$$\begin{aligned}
=&(\widehat{\boldsymbol{p}}_0(c)-c)\left[D_{\theta_{\overline{\boldsymbol{p}}_0(c)}}(\widehat{\boldsymbol{p}}_0(c))\beta_0(\theta_{\overline{\boldsymbol{p}}_0(c)}) + \sum_{\{\theta':l(\theta')\geq \widehat{\boldsymbol{p}}_0(c)\}} \beta_0(\theta')\right] \\
=&(\widehat{\boldsymbol{p}}_0(c)-c)D_0(\widehat{\boldsymbol{p}}_0(c)) \\
\leq&(\overline{\boldsymbol{p}}_0(c)-c)D_0(\overline{\boldsymbol{p}}_0(c)) \\
=&(\overline{\boldsymbol{p}}_0(c)-c)\left[D_{\theta_{\overline{\boldsymbol{p}}_0(c)}}(\overline{\boldsymbol{p}}_0(c)) + \sum_{\{\theta':l(\theta')\geq \overline{\boldsymbol{p}}_0(c)\}} \beta_0(\theta')\right],
\end{aligned}$$

and

$$(\overline{\boldsymbol{p}}_0(c)-c)D_{\theta_{\overline{\boldsymbol{p}}_0(c)}}(\overline{\boldsymbol{p}}_0(c)) \leq (\widehat{\boldsymbol{p}}_0(c)-c)D_{\theta_{\overline{\boldsymbol{p}}(c)}}(\widehat{\boldsymbol{p}}_0(c)).$$

As a result, it must be that $\widehat{\boldsymbol{p}}_0(c) \leq \overline{\boldsymbol{p}}_0(c)$.

Consequently,

$$\begin{aligned}
\sum_{\{\theta:l(\theta)\geq \overline{\boldsymbol{p}}_0(c)\}} D_\theta(\overline{\boldsymbol{p}}_\theta(c))\beta_0(\theta) =& \sum_{\{\theta:l(\theta)\geq \overline{\boldsymbol{p}}_0(c)\}} \beta_0(\theta) \\
\leq& \sum_{\{\theta:l(\theta)\geq \overline{\boldsymbol{p}}_0(c)\}} \beta_0(\theta) + D_{\theta_{\overline{\boldsymbol{p}}_0(c)}}(\overline{\boldsymbol{p}}_0(c))\beta_0(\theta_{\overline{\boldsymbol{p}}_0(c)}) \\
\leq& D_0(\overline{\boldsymbol{p}}_0(c)),
\end{aligned}$$

which proves (A.26). On the other hand, for any $\nu \in [c, \overline{\boldsymbol{p}}_0(c)]$

$$
\begin{aligned}
&\sum_{\{\theta: u(\theta)\geq \nu\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c))\beta_0(\theta)\\
&= \sum_{\{\theta: \nu\leq u(\theta)<\overline{\boldsymbol{p}}_0(c)\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c))\beta_0(\theta) + \sum_{\{\theta: u(\theta)\geq \overline{\boldsymbol{p}}_0(c)\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c))\beta_0(\theta)\\
&\geq \sum_{\{\theta: u(\theta)\geq \overline{\boldsymbol{p}}_0(c)\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c))\beta_0(\theta)\\
=&D_{\theta_{\overline{\boldsymbol{p}}_0(c)}}(\widehat{\boldsymbol{p}}_0(c)) + \sum_{\{\theta': l(\theta')\geq \overline{\boldsymbol{p}}_0(c)\}} D_{\theta'}(l(\theta'))\beta_0(\theta')\\
\geq&D_{\theta_{\overline{\boldsymbol{p}}_0(c)}}(\overline{\boldsymbol{p}}_0(c)) + \sum_{\{\theta': l(\theta')\geq \overline{\boldsymbol{p}}_0(c)\}} \beta_0(\theta')\\
=&D_0(\overline{\boldsymbol{p}}_0(c)),
\end{aligned}
$$

which proves (A.25) ■

With Lemma A.3 and Lemma A.4, the proof of Theorem 9 is as below.

Proof of Theorem 9. To prove Theorem 9, first notice that Lemma 1 still applies and hence the data broker's maximization problem can be written as

$$
\begin{aligned}
&\max_\sigma \int_C \left(\int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \phi_G(c)) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c))\sigma(\mathrm{d}\beta|c)\right) G(\mathrm{d}c)\\
&\text{s.t. } \int_c^{c'} \left(D_\beta(\overline{\boldsymbol{p}}_\beta(z))(\sigma(\mathrm{d}\beta|z) - \sigma(\mathrm{d}\beta|c'))\right) \mathrm{d}z \geq 0,\ \forall c, c' \in C \qquad \text{(A.27)}\\
&\quad \bar{\pi} + \int_c^{\bar{c}} \left(\int_{\Delta(\Theta)} D_\beta(\overline{\boldsymbol{p}}_\beta(z))\sigma(\mathrm{d}\beta|z)\right) \mathrm{d}z \geq \bar{\pi} + \int_c^{\bar{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in C,
\end{aligned}
$$

where the maximum is taken over all $\sigma : C \to \Delta(\Delta(\Theta))$ such that $\sigma(c)$ is a segmentation for all $c \in C$.

Consider first a relaxed problem of (A.27) where the first constraint is relaxed to $\boldsymbol{D}_\sigma : C \to [0, 1]$ being nonincreasing, where

$$
\boldsymbol{D}_\sigma(c) := \int_{\Delta(\Theta)} D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c))\sigma(\mathrm{d}\beta|c),
$$

for all $c \in C$. By the same duality argument as in the proof of Lemma 7, it suffices to find a feasible σ^* and a Borel measure μ^* on C such that

$$
\begin{aligned}
\sigma^* \in \operatorname*{argmax}_{\sigma\in\Sigma} \Bigg[&\int_C \left(\int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \phi_G(c)) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c))\sigma(\mathrm{d}\beta|c)\right) G(\mathrm{d}c)\\
&+ \int_C \left(\int_c^{\bar{c}} \left(\int_{\Delta(\Theta)} D_\beta(\overline{\boldsymbol{p}}_\beta(z))\sigma(\mathrm{d}\beta|z) - D_0(\overline{\boldsymbol{p}}_0(z))\right) \mathrm{d}z\right) \mu^*(\mathrm{d}c)\Bigg],
\end{aligned}
$$

where Σ is the collection of segmentation schemes such that $\boldsymbol{D}_\sigma$ is nonincreasing, and that

$$
\int_C \left(\int_c^{\bar{c}} \left(\int_{\Delta(\Theta)} D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(z))\sigma^*(\mathrm{d}\beta|z) - D_0(\overline{\boldsymbol{p}}_0(z))\right) \mathrm{d}z\right) \mu^*(\mathrm{d}c) = 0.
$$

To this end, let M^* be defined as

$$M^*(c) := \lim_{c' \downarrow c} g(c)(\phi_G(c) - \widehat{\boldsymbol{p}}_0(c))^+.$$

Since $c \mapsto g(c)(\phi_G(c) - \widehat{\boldsymbol{p}}_0(c))^+$ in nondecreasing, M^* is nondecreasing and right-continuous and hence induced a Borel measure μ^* with $\mathrm{supp}(\mu^*) = [c^*, \overline{c}]$ for some $c^* \leq \overline{c}$. Then, by the same arguments as in the proof of Lemma 7 (in particular, the definition of $\widehat{\varphi}_G$, (23) and (24)),

$$\begin{aligned}\max_{\sigma \in \Sigma} \Bigg[\int_C \left(\int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \phi_G(c)) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c)) \sigma(\mathrm{d}\beta|c) \right) G(\mathrm{d}c) \\ + \int_C \left(\int_c^{\overline{c}} \left(\int_{\Delta(\Theta)} D_\beta(\overline{\boldsymbol{p}}_\beta(z)) \sigma(\mathrm{d}\beta|z) - D_0(\overline{\boldsymbol{p}}_0(z)) \right) \mathrm{d}z \right) \mu^*(\mathrm{d}c) \Bigg]\end{aligned}$$

is equivalent to

$$\max_{\sigma \in \Sigma} \int_C \left(\int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \widehat{\varphi}_G(c)) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c)) \sigma(\mathrm{d}\beta|c) \right) G(\mathrm{d}c). \tag{A.28}$$

To solve (A.28), notice that for any $c \in [\underline{c}, c^*)$,

$$\sum_{\{\theta : u(\theta) \geq \widehat{\varphi}_G(c)\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) > D_0(\overline{\boldsymbol{p}}_0(c)),$$

which is due to $\widehat{\varphi}_G(c) = \varphi_G(c) \leq \widehat{\boldsymbol{p}}_0(c) \leq \overline{\boldsymbol{p}}_0(c)$ and (A.26). Meanwhile, for any $c \in (c^*, \overline{c}]$, there exists a unique $\lambda(c)$ such that

$$\lambda(c) D_{\theta_{\widehat{\varphi}_G(c)}}(\widehat{\boldsymbol{p}}_0(c)) + \sum_{\{\theta : l(\theta) \geq \widehat{\varphi}_G(c)\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) = D_0(\overline{\boldsymbol{p}}_0(c)),$$

which is due to the fact that $\widehat{\varphi}_G(c) = \widehat{\boldsymbol{p}}_0(c)$ for all $c \in (c^*, \overline{c}]$ and (A.25). Furthermore, Since D_0 is regular, for any $\theta \in \Theta$ such that $u(\theta) \geq \widehat{\varphi}_G(c)$ and for any $p \leq l(\theta_{\widehat{\varphi}_G(c)})$,

$$\begin{aligned}(p - c) D_{\beta^\theta_{\widehat{\varphi}_G(c)}}(p) = & \sum_{\{\theta' : u(\theta') \geq u(\theta_p)\}} (p - c) D_{\theta'}(p) \beta^\theta_{\widehat{\varphi}_G(c)}(\theta') \\ = & (p - c) D_0(p) \\ \leq & (l(\theta_{\widehat{\varphi}_G(c)}) - c) D_0(l(\theta_{\widehat{\varphi}_G(c)})) \\ \leq & (l(\theta) - c) D_0(l(\theta_{\widehat{\varphi}_G(c)})) \\ = & (l(\theta) - c) \sum_{\{\theta' : u(\theta') \geq \widehat{\varphi}_G(c)\}} \beta^\theta_{\widehat{\varphi}_G(c)}(\theta') \\ = & (l(\theta) - c) D_{\beta^\theta_{\widehat{\varphi}_G(c)}}(l(\theta)) \\ = & (\overline{\boldsymbol{p}}_{D_\theta}(c) - c) D_{\beta^\theta_{\widehat{\varphi}_G(c)}}(\overline{\boldsymbol{p}}_{D_\theta}(c)),\end{aligned} \tag{A.29}$$

where $\beta^\theta_{\widehat{\varphi}_G(c)}$ is defined in (13). In addition, by the same construction as in the proof of Lemma 3, for any $c \in (c^*, \overline{c}]$, there exists a segmentation $\tilde{\sigma}(c) \in \Delta(\Delta(\Theta))$ such that $\mathrm{supp}(\tilde{\sigma}(c)) = \{\tilde{\beta}^\theta_{\widehat{\boldsymbol{p}}_0(c)} : l(\theta) \geq \widehat{\boldsymbol{p}}_0(c)\}$, with $\tilde{\beta}^\theta_{\widehat{\boldsymbol{p}}_0(c)}$ satisfying (14) and (15) and that

$$(p - c) D_{\tilde{\beta}^\theta_{\widehat{\boldsymbol{p}}_0(c)}}(p) \leq (l(\theta) - c) D_\theta(l(\theta)) = (\overline{\boldsymbol{p}}_{D_\theta}(c) - c) D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) \tag{A.30}$$

for all $\theta \in \Theta$ such that $l(\theta) \geq \overline{\boldsymbol{p}}_0(c)$, as well as

$$\overline{\boldsymbol{p}}_{D_{\tilde{\beta}^{\theta}_{\widehat{\boldsymbol{p}}_0(c)}}}(z) \geq \overline{\boldsymbol{p}}_{D_0}(z) \geq \widehat{\boldsymbol{p}}_0(z) \tag{A.31}$$

for all $z \in [\underline{c}, c]$ and for all $\theta \in \Theta$ such that $l(\theta) \geq \overline{\boldsymbol{p}}_0(c)$.

Now define σ^* as follows.

$$\sigma^*(c) := \begin{cases} \sigma_1(c), & \text{if } c \in [\underline{c}, c^*] \\ \sigma_2(c), & \text{if } c \in (c^*, \overline{c}] \end{cases},$$

where

$$\sigma_1(\beta^{\theta}_{\varphi_G(c)}|c) := \frac{\beta_0(\theta)}{\sum_{\{\theta': u(\theta') \geq \varphi_G(c)\}} \beta_0(\theta')}$$

for all $c \in [\underline{c}, c^*]$ and for all $\theta \in \Theta$ such that $u(\theta) \geq \varphi_G(c)$, whereas

$$\sigma_2(\beta|c) := \begin{cases} \lambda(c) \frac{\beta_0(\theta)}{\sum_{\{\theta': u(\theta') \geq \widehat{\boldsymbol{p}}_0(c)\}} \beta_0(\theta')}, & \text{if } \beta = \beta^{\theta}_{\widehat{\boldsymbol{p}}_0(c)},\ u(\theta) \geq \widehat{\boldsymbol{p}}_0(c) \\ (1-\lambda(c))\tilde{\sigma}(\tilde{\beta}^{\theta}_{\widehat{\boldsymbol{p}}_0(c)}|c), & \text{if } \beta = \tilde{\beta}^{\theta}_{\widehat{\boldsymbol{p}}_0(c)},\ l(\theta) \geq \widehat{\boldsymbol{p}}_0(c) \\ 0, & \text{otherwise} \end{cases},$$

for all $c \in (c^*, \overline{c}]$. It then follows that, by (A.29) and (A.30),

$$\begin{aligned} &\int_C \left(\int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \widehat{\varphi}_G(c)) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c)) \sigma^*(\mathrm{d}\beta|c) \right) G(\mathrm{d}c) \\ =& \int_C \left(\sum_{\{\theta: \overline{\boldsymbol{p}}_{D_\theta}(c) \geq \hat{\varphi}_G(c)\}} (\overline{\boldsymbol{p}}_{D_\theta}(c) - \widehat{\varphi}_G(c)) D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) \beta_0(\theta) \right) G(\mathrm{d}c), \end{aligned}$$

which, together with Lemma A.3, implies that σ^* is a solution of (A.28).

Furthermore, for any $c > c^*$, by the definition of $\sigma_2(c)$ and $\lambda(c)$, by (A.29) and (A.30), and by the fact that $\widehat{\varphi}_G(c) = \widehat{\boldsymbol{p}}_0(c)$,

$$\int_{\Delta(\Theta)} D_\beta(\overline{\boldsymbol{p}}_\beta(c)) \sigma^*(\mathrm{d}\beta|c) = D_0(\overline{\boldsymbol{p}}_0(c)).$$

Therefore,

$$\int_C \left(\int_c^{\overline{c}} \left(\int_{\Delta(\Theta)} D_\beta(\overline{\boldsymbol{p}}_\beta(z)) \sigma^*(\mathrm{d}\beta|z) - D_0(\overline{\boldsymbol{p}}_0(z)) \right) \mathrm{d}z \right) \mu^*(\mathrm{d}c) = 0.$$

Finally, by definition of $\widehat{\varphi}_G$ and by Lemma A.4,

$$\int_{\Delta(\Theta)} D_\beta(\overline{\boldsymbol{p}}_\beta(c)) \sigma^*(\mathrm{d}\beta|c) \geq D_0(\overline{\boldsymbol{p}}_0(c))$$

for all $c \in [\underline{c}, c^*)$. Together with monotonicity of $\widehat{\varphi}_G$, $\sigma^* \in \Sigma$ and is a solution of the relaxed problem of (A.27).

It then suffices to show that σ^* is implementable. Notice that for any $c \in C$ and for any $z \in [\underline{c}, c]$ and for any $\beta^{\theta}_{\widehat{\varphi}_G(c)} \in \operatorname{supp}(\sigma^*(c))$, if

$$\boldsymbol{P}_{D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}}(z) \cap \operatorname{supp}(D_\theta) = \emptyset,$$

then it must be that

$$\begin{aligned}(p-z)D_0(p) =&(p-z)D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}(p)\\ \leq&(\overline{\boldsymbol{p}}_{D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}}(z)-z)D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}(\overline{\boldsymbol{p}}_{D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}}(z))\\ =&(\overline{\boldsymbol{p}}_{D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}}(z)-z)D_0(\overline{\boldsymbol{p}}_{D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}}(z)),\end{aligned}$$

for all $p \leq \overline{\boldsymbol{p}}_{D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}}(z)$. Therefore,

$$\overline{\boldsymbol{p}}_{D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}}(z) \geq \overline{\boldsymbol{p}}_0(z) \geq \widehat{\boldsymbol{p}}_0(z) \geq \widehat{\varphi}_G(z),$$

for all $z \in [\underline{c}, c]$. Together with (A.31), by the same argument as the proof of Lemma 2, σ^* is indeed implementable. This completes the proof. ■

References

Hart, S. and P. Reny (2019): "The Better Half of Selling Separately," *ACM Transactions on Economics and Computation*, 1, Article 18.

Porter, J. E. (2005): "Helly's Selection Principle for Functions with Bounded p-Variation," *Rocky Mountain Journal of Mathematics*, 35, 675–679.

Yang, K. H. (2020a): "A Note on Topological Properties of Outcomes in a Monopoly Market," Working Paper.

——— (2020b): "Selling Consumer Data for Profit: Optimal Market-Segmentation Design and its Consequences," Discussion Paper 2258, Cowles Foundation for Research in Economics, Yale University.