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

Kai Hao Yang*

April 21, 2021

Abstract

A data broker sells market segmentations to a producer with private cost who sells a product to a unit mass of consumers. This paper characterizes the revenue-maximizing mechanisms for the data broker. Every optimal mechanism induces quasi-perfect price discrimination—all the consumers with values above a cost-dependent cutoff buys by paying their values while the rest of consumers do not buy. The characterization of optimal mechanisms leads to several economic implications: (i) market outcomes remain unchanged even if the data broker becomes more active in the product market—either by gaining the ability to contract on prices or by becoming a retailer who purchases the product and obtains the exclusive right to sell to the consumers directly; (ii) vertical integration between the data broker and the producer increases total surplus while leaving consumer surplus unchanged and (iii) data brokership improves total surplus compared with uniform pricing.

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

JEL CLASSIFICATION: D42, D61, D82, 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 Manelli, 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. 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 a 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. Since the producer may rank the values of market segmentations differently when having different marginal costs, this leads to a screening problem with an infinite dimensional allocation space and a non-single-crossing agent. Moreover, as the data broker only affects the product market indirectly by selling consumer data to the producer and cannot contract on how the data are used (in particular, on prices), 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

\(^1\)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
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 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 brokering. 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. In those papers, market segmentation is treated as an exogenous object. In addition, Ali, Lewis, and Vasserman (2020) study the welfare effect of third-degree price discrimination when the consumers can disclose information about their values voluntarily, and thus market segmentation is formed endogenously by consumers’ equilibrium behavior. In contrast, market segmentation in this paper is determined endogenously by a data broker, who creates and sells market segmentations to a producer to facilitate price discrimination. Relatedly, Wei and Green (2020) also study price discrimination in a mechanism design framework. They consider a monopolist who can provide information about a product and designs selling mechanisms at the same time, while I consider a third party who sells only information about the consumers to a monopolist.

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. In this regard, I study the revenue-maximizing

2See also: Haghpanah and Siegel (2020), and Haghpanah and Siegel (2021) who further consider segmentations in environments that feature second-degree price discrimination.

3Although both Wei and Green (2020) and this paper have one-dimensional type and thus share some similarities in terms of methodology (e.g., the revenue equivalence formula, pointwise maximization), they are substantially different. In this paper, the agent solves a pricing problem using the information provided, as opposed to making a binary choice. As a result, the agent’s payoff, as a function of type and allocation, is non-multiplicative and non-single-crossing (see Section 3.5, Section 4.4, and Appendix D).
mechanism of a data broker who sells consumer information to a producer to facilitate price
discrimination.\footnote{Relatedly, Acemoglu, Makhdoumi, Malekian, and Ozdaglar (forthcoming), Bergemann, Bonatti, and Gan (2021) and Ichihashi (forthcoming) 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 parameterized family and the data-buying firm uses the purchased information to solve a (private) forecast problem.}

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
design problem with a high-dimensional allocation space and a non-single-crossing agent. In
particular, the characterization of incentive compatible mechanisms here resembles those that
appear in the dynamic mechanism design literature (e.g., Pavan, Segal, and Toikka (2014);
Bergemann and Valimaki (2019); Karsikov and Lamba (2020)), which in turn reflect the
\textit{integral monotonicity} condition in the literature (e.g., Berger, Muller, and Naeemi (2010);
Carbajal and Ely (2013)).

Among the aforementioned papers, Bergemann, Brooks, and Morris (2015), Bergemann,
Bonatti, and Smolin (2018) are the closest to this paper. Specifically, Bergemann, Brooks,
and Morris (2015) explore the welfare implications of different market segmentations, while I
introduce 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, I study 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. Nonetheless,
as in Bergemann, Bonatti, and Smolin (2018), agents with different types would also have
different rankings regarding the value of information in this paper.

\section{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
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 $D_0$.

Figure 1: Representation of the market

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? A natural guess would be 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 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 $\tau$ 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 $\tau$, 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

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

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.\(^5\)

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

\(^5\)This feature is specific to the parametric assumptions of the current example, and is the main reason (besides finiteness) why the example simplifies the main result. See more discussions in footnote 22.
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, \( M^* \) is an optimal mechanism of the data broker. The optimal mechanism \( M^* \) 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 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 about 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 $\sigma$-algebra. Endow $\Delta(X)$ with the weak-* topology and the Borel $\sigma$-algebra. When $X = [\underline{v}, \overline{v}] \subseteq \mathbb{R}$ is an interval, let $\mathcal{D}(X)$ denote the collection of nonincreasing and upper-semicontinuous functions $D : \mathbb{R}_+ \rightarrow [0, 1]$ such that $D(\underline{v}) = 1$, $D(\overline{v}^+) = 0$.

Since $\mathcal{D}(X)$ and $\Delta(X)$ are bijective, for any $D \in \mathcal{D}(X)$, let $m^D \in \Delta(X)$ be the probability measure associated with $D$ and define the integral

$$\int_{\mathcal{A}} h(x) D(dx) := \int_{\mathcal{A}} h(x)m^D(dx),$$

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

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 $V = [\underline{v}, \overline{v}] \subseteq \mathbb{R}_+$ 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 $\text{supp}(D_0)$ for all $c \geq 0$.

The producer has a constant marginal cost of production $c \in C = [\underline{c}, \overline{c}] \subseteq \mathbb{R}_+$ 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 $\phi_G$, defined as $\phi_G(c) := c + G(c)/g(c)$ for all $c \in C$. Henceforth, $G$ is said to be regular if $\phi_G$ 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

$$\int_{\mathcal{D}} D(p)s(dD) = D_0(p), \forall p \in V. \quad (1)$$

---

6As a convention, for any function $f$ defined on $\mathbb{R}_+$, $f(x^+)$ denotes the right limit of $f$ at $x$.

7This is because for any $D \in \mathcal{D}(X)$, the right limit of $1 - D$ is nondecreasing and right-continuous.

8If $D_0$ 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).
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.\(^9\) Let $\mathcal{S}$ denote the set of segmentations.

### 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 $D \in \text{supp}(s)$.\(^10\) 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 $\mathbf{P}_D(c)$ denote the set of optimal prices for the producer with marginal cost $c$ under market segment $D$. As a convention, regard $\mathbf{P}$ as a correspondence on $\mathcal{D} \times C$ and if $\mathbf{p}$ is a selection for $\mathbf{P}$, write $\mathbf{p} \in \mathbf{P}$.\(^11\) Furthermore, for any $c \in C$ and any $D \in \mathcal{D}$, let

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

\(^9\)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 $D \in \text{supp}(s)$ 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.

\(^10\)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.

\(^11\)For notational conveniences, I restrict the feasible prices for each producer to a large enough compact interval $\mathbb{V} \subset \mathbb{R}^+$ such that $\mathbb{V} \subset \mathbb{V}$. With this restriction, $\mathbf{P}_D(c)$ would be a subset of a compact interval for all $D \in \mathcal{D}$ and for all $c \in C$. Since $\mathbb{V}$ itself is bounded, this restriction is simply a notational convention and does not affect the model at all.
denote the maximized profit of the producer. Also, let

\[ p_D(c) := \max P_D(c) \]

be the largest optimal price for the producer with marginal cost \( c \) under market segment \( D \).\(^\text{12}\) For conciseness, let \( \bar{p}_0 := \bar{p}_{D_0} \).

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

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

Assumption 1 permits a wide class of \((D_0, G)\) and includes many common examples.\(^\text{13}\) 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.\(^\text{14}\)

Formally, a mechanism is a pair \((\sigma, \tau)\), where \( \sigma : C \to S \), \( \tau : C \to \mathbb{R} \) are measurable functions. Given a mechanism \((\sigma, \tau)\), for each report \( c \in C \), \( \sigma(c) \in S \) 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 \( \sigma : C \to S \) is called a *segmentation scheme* (or sometimes, a *scheme*).

A mechanism \((\sigma, \tau)\) is incentive compatible if for all \( c, c' \in C \),

\[
\int_D \pi_D(c)\sigma(dD|c) - \tau(c) \geq \int_D \pi_D(c)\sigma(dD|c') - \tau(c').
\]  

(\text{IC})

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

\[
\int_D \pi_D(c)\sigma(dD|c) - \tau(c) \geq \pi_{D_0}(c).
\]  

(\text{IR})

\(^\text{12}\)\( \bar{p} \) is well-defined under the notational convention stated in footnote 11, as \( P_D \) is a closed (implied by upper-semicontinuity of \( D \)) subset of a compact set \( \bar{V} \).

\(^\text{13}\)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.

\(^\text{14}\)Henceforth, unless otherwise noted, a *mechanism* stands for a *direct mechanism*. 
Henceforth, a mechanism \((\sigma, \tau)\) is said to be incentive feasible if it is incentive compatible and individually rational. A segmentation scheme \(\sigma\) is said to be implementable if there exists a measurable \(\tau : C \to \mathbb{R}\) such that \((\sigma, \tau)\) is incentive feasible. The goal of the data broker is to maximize expected revenue \(\mathbb{E}_C[\tau(c)]\) by choosing an incentive feasible mechanism.

### 3.5 Discussions about the Model

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 the space of market segmentations that would lead to the single-crossing property commonly assumed in low-dimensional screening problems. In Appendix D, I provide a counter example demonstrating that the producer’s profit, as a function of market segmentation and cost, is not single-crossing when market segmentations are ordered by the Blackwell order.\(^{15}\) A consequence of this feature is that, although local incentives of the producer can still be summarized by a revenue equivalence formula as in a one-dimensional screening problem, monotonicity of the “allocation rule” would not be sufficient for global incentives. As a result, more complicated constraints must be considered when solving for the optimal mechanisms (see Lemma 1 below).

Secondly, 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. Therefore, individual rationality constraints would not necessarily be satisfied even if there is no rent at the top. A continuum of individual rationality constraints must be kept track of when solving for the optimal mechanism.

Lastly, 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 \(s \in S\) 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.

\(^{15}\)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 (see Appendix D).
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 $s \in S$ is a $\kappa$-quasi-perfect segmentation for $c$ if for $s$-almost all $D \in D$, either $D(c) = 0$, or the set $\{v \in \text{supp}(D) : v \geq \kappa\}$ is a singleton and is a subset of $P_D(c)$.

A $\kappa$-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 that every market segment with positive trading volume\(^{16}\) 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 $\kappa$-quasi-perfect segmentation for $c$ induces $\kappa$-quasi-perfect price discrimination when the producer’s marginal cost is $c$ and 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 $\sigma$ is a $\psi$-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 $(\sigma, \tau)$ is a $\psi$-quasi-perfect mechanism if $\sigma$ is a $\psi$-quasi-perfect scheme and if the producer with marginal cost $\bar{c}$, when reporting truthfully, has net profit $\pi_{D_0}(\bar{c})$.

\(^{16}\)Notice that when the producer’s marginal cost is $c$, no trade occurs in market segment $D$ if and only if $D(c) = 0$. 
4.2 Characterization of the Optimal Mechanisms

With the definitions above, the main result of this paper can be stated. For any \( c \in C \), define \( \overline{\varphi}_G(c) := \min\{\varphi_G(c), p_0(c)\} \), where \( \varphi_G \) is the ironed virtual cost function.\(^{17}\)

**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 \( \overline{\varphi}_G \)-quasi-perfect 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. 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 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 \( v \geq \overline{\varphi}_G(c) \), define market segment \( D_{v_0}^{\overline{\varphi}_G(c)} \) as

\[
D_{v_0}^{\overline{\varphi}_G(c)}(p) := \begin{cases} 
D_0(p), & \text{if } p \in [v, \overline{\varphi}_G(c)] \\
D_0(\overline{\varphi}_G(c)), & \text{if } p \in (\overline{\varphi}_G(c), v] \\
0, & \text{if } p \in (v, \overline{v}] 
\end{cases}
\]  

(2)

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

\[
\sigma^*(\{D_{v_0}^{\overline{\varphi}_G(c)} : v \geq p\} | c) := \frac{D_0(p)}{D_0(\overline{\varphi}_G(c))}.
\]  

(3)

In other words, for any \( c \in C \), \( \sigma^*(c) \) induces market segments \( \{D_{v_0}^{\overline{\varphi}_G(c)}\}_{v \in [\overline{\varphi}_G(c), \overline{v}]} \), 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 \( \sigma^*(c) \in S \)).\(^{18}\) Figure 3a illustrates \( \sigma^* \) by plotting the (inverse) demands\(^{19}\) of generic market segments \( D_{v_0}^{\overline{\varphi}_G(c)} \), \( D_{v_0'}^{\overline{\varphi}_G(c)} \), and \( D_{v_0''}^{\overline{\varphi}_G(c)} \) induced by \( \sigma^*(c) \) (the dashed line represents the market demand

---

\(^{17}\)Ironing in the sense of Myerson (1981).

\(^{18}\)Notice that \( \sigma^* : C \to S \) is well-defined and measurable since for all \( c \in C \), \( v \mapsto D_{v_0}^{\overline{\varphi}_G(c)} \) is a measurable function from \( V \) to \( D \) and since \( D_0 \circ \overline{\varphi}_G : C \to [0, 1] \) is also measurable.

\(^{19}\)See Appendix A for the formal definition of inverse demands.
These inverse demands have a jump at $D_0(\varphi_G(c))$. To the left of $D_0(\varphi_G(c))$, all the consumer values are concentrated at $v$, $v'$, or $v''$, whereas the distributions of the consumer values to the right of $D_0(\varphi_G(c))$ remain the same as that under $D_0$.

With this definition, it turns out that when $D_0$ is regular, as will be shown in Section 4.3, there exists a unique transfer scheme $\tau^* : C \rightarrow \mathbb{R}$ such that $(\sigma^*, \tau^*)$ is an incentive feasible $\varphi_G$-quasi-perfect mechanism. Thus, by Theorem 1, $(\sigma^*, \tau^*)$ is optimal. Henceforth, I refer the mechanism $(\sigma^*, \tau^*)$ as the canonical $\varphi_G$-quasi-perfect mechanism.

**Theorem 2.** Suppose that $D_0$ is regular. Then the canonical $\varphi_G$-quasi-perfect mechanism $(\sigma^*, \tau^*)$ is optimal.

According to Theorem 1, under any optimal mechanism $(\sigma, \tau)$, a producer with cost $c$ pays the data broker $\tau(c)$ and purchases a $\varphi_G(c)$-quasi-perfect segmentation for $c$. The willingness to pay of a producer with cost $c$ for a $\varphi_G(c)$-quasi-perfect segmentation for $c$ is depicted in Figure 3b. As will be shown in Section 5.3, for a producer with cost $c < c^* := \inf\{c \in C : \overline{p}_0(c) \geq \varphi_G(c)\}$, her payment is strictly lower than her willingness to pay (i.e., (IR) is slack); while for a producer $c \geq c^*$, her payment equals to her willingness to pay (i.e., (IR) is binding). Furthermore, when the producer’s cost is $c$, all the consumers with values $v \geq \varphi_G(c)$ will be assigned to different market segments (i.e., $\{D_0^{\varphi_G(c)}(v)\}_{v \in [\varphi_G(c)\}}$), whereas all the consumers with values $v < \varphi_G(c)$ are (uniformly, under $(\sigma^*, \tau^*)$) distributed across each market segment. This allows the producer to distinguish consumers with $v \geq \varphi_G(c)$ among each other, but not from consumers with $v < \varphi_G(c)$. This type of segmentation can be interpreted as consumer data that differentiate high-value consumers but not the low-value ones.\(^{20}\)

As an example, notice that the 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 $5/4$,\(^{21}\) the menu $\mathcal{M}^*$ is indeed optimal.

\(^{20}\)For instance, aggregate purchase histories of related products (e.g., average purchase price throughout time) would be useful to differentiate high-value consumers among each other (as they tend to purchase similar products more frequently, and hence the average purchasing price would be more informative). Meanwhile, these histories are not very informative about low-value consumers (as there would be fewer transactions, and hence the average price would be more noisy), nor could they completely differentiate low-value consumers from the high-values (as consumers with a given average purchase price could be someone who has purchased many similar products or someone who has purchased only one similar product for other reasons).

\(^{21}\)Although the characterization is stated for cost distributions that admit densities, as in standard mech-
As another example, consider the case where $D_0$ is linear and $G$ is uniform. Suppose that $V = C = [0, 1]$, $D_0(v) = (1 - v)$ for all $v \in V$ and $G(c) = c$ for all $c \in C$. In this case, $\varphi_G(c) = \phi_G(c) = 2c$ and $\overline{\pi}_0(c) = (1 + c)/2$. Thus, $\overline{\varphi}_G(c) = 2c$ for all $c \in [0, 1/3]$ and $\overline{\varphi}_G(c) = (1 + c)/2$ for all $c \in (1/3, 1]$. The canonical quasi-perfect mechanism $(\sigma^*, \tau^*)$ is as follows: For each $c$, market segments $\{D_{v_G}^G(c)\}_{v \in [\varphi_G(c), 1]}$ (as defined by (2)) are uniformly distributed under $\sigma^*(c)$. Moreover, for the producer with cost $c \in [0, 1]$, her payment and net profit are:

$$\tau^*(c) = \begin{cases} 
\frac{1}{6} - c^2, & \text{if } c \in [0, \frac{1}{3}] \\
\frac{1}{3} - c^2, & \text{if } c \in \left(\frac{1}{3}, 1\right],
\end{cases}$$

and

$$\int_D \pi_D(c)\sigma^*(dD|c) - \tau^*(c) = \begin{cases} 
\frac{(1-2c)^2}{4} + \frac{1}{12}, & \text{if } c \in [0, \frac{1}{3}] \\
\frac{(1-c)^2}{4}, & \text{if } c \in \left(\frac{1}{3}, 1\right],
\end{cases}$$

respectively, while the data broker’s expected revenue is $5/54$ and the prices charged by the producer with cost $c$ are uniformly distributed on $[\varphi_G(c), 1]$.
4.3 Outline of the Proof

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 $\varphi_G$-quasi-perfect mechanism is optimal. Finally, I argue that any mechanism that gives revenue $\bar{R}$ must be $\varphi_G$-quasi-perfect.

Before outlining the proof, recall that the data broker’s revenue maximization problem differs from standard one-dimensional screening problems in two aspects: (i) the allocation space is infinite-dimensional and (ii) the producer’s outside option is type dependent. 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 4.3, I assume that $D_0$ and $G$ are regular and that

$$\phi_G(c) \leq \overline{p}_0(c), \forall c \in C.$$  \hspace{1cm} (4)

Note that (4) is a sufficient condition for Assumption 1. Also note that all the lemmas stated in this section do not rely on any of these additional assumptions, nor on Assumption 1.

With these additional conditions, $\varphi_G(c) = \phi_G(c)$ for all $c \in C$ and hence $\varphi_G$ can be replaced by the virtual cost function $\phi_G$. Among these assumptions, regularity of $G$ is purely for conciseness and can be relaxed by ironing $\phi_G$. 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) ensures that individual rationality of the constructed mechanism is implied by incentive compatibility, and by the fact that there is no rent at the top, effectively circumventing the complication caused by feature (ii) above.

Nonetheless, even with these simplifying assumptions, the data broker’s problem is still noticeably different from a standard one-dimensional screening problem. Specifically, as discussed in Section 3.5, the producer’s profit, as a function of market segmentation and cost, is not single-crossing in general. In fact, even when restricting attention to the class of quasi-perfect segmentations (so that they can be ranked by a one-dimensional cutoff $\kappa$), the producer’s profit can still be non-single-crossing (see Appendix D for an example). As a result, feature (i) above does not only lead to a more challenging pointwise maximization

\footnote{Although the intuition provided in Section 2 might seem to suggest that the producer’s profit as a function of segmentation and cost exhibits the single-crossing property when restricting to quasi-perfect segmentations, this is not generally true. After all, the intuition in Section 2 relies on the fact that both the high-cost producer ($c = 3/4$) and the low-cost producer ($c = 1/4$) would optimally sell to $v = 2$ and $v = 3$ by charging their values under the segmentation created by the residential data. In general, although the definition of quasi-perfect segmentations requires the high-cost producer to do so, the low-cost producer would not necessarily behave in the same way. As a result, it is not necessary that the producer with a}
problem (as it is infinite-dimensional rather than one-dimensional), it also means that a simple monotonicity condition—even when restricting to quasi-perfect mechanisms—would not be sufficient for incentive compatibility, and thus requires more sophisticated arguments.

Characterization of IC Mechanisms and an Upper Bound for Revenue

Despite the high-dimensionality of the date broker’s problem, a revenue-equivalence formula can still be derived by properly invoking the envelope theorem. To see this, notice that for any incentive compatible mechanism \((\sigma, \tau)\), the indirect utility of a producer with marginal cost \(c\) is

\[
U(c) := \int_D \pi_D(c) \sigma(dD|c) - \tau(c) = \max_{c' \in C} \left[ \int_D \pi_D(c) \sigma(dD|c') - \tau(c') \right].
\]

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

\[
U'(c) = \int_D \pi'_D(c) \sigma(dD|c).
\]

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 \(c \in C\),

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

Together,

\[
U(c) = U(\tau) + \int_c^\tau \left( \int_D D(\overline{p}_D(z)) \sigma(dD|z) \right) dz, \forall c \in C.
\]

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

\[
U(c) - \left( \int_D \pi_D(c) \sigma(dD|c') - \tau(c') \right)
\]

\[
= \int_D [\pi_D(c) - \pi_D(c')] \sigma(dD|c') - (U(c) - U(c'))
\]

\[
= \int_c^{c'} \left[ \int_D -\pi'_D(z) \sigma(dD|c') - \int_D D(\overline{p}_D(z)) \sigma(dD|z) \right] dz
\]

\[
= \int_c^{c'} \left[ \int_D D(\overline{p}_D(z)) \sigma(dD|c') - \int_D D(\overline{p}_D(z)) \sigma(dD|z) \right] dz.
\]

Together, these lead to Lemma 1 below.

lower cost would gain (strictly) more from the value-revealing data relative to the residential data than the producer with a higher cost.
Lemma 1. A mechanism $(\sigma, \tau)$ is incentive compatible if and only if:

1. For all $c \in C$,
   \[
   \tau(c) = \int_D \pi_D(c)(dD|c) - \int_c^c \left( \int_D D(\bar{p}_D(z))\sigma(dD|z) \right) dz - U(\bar{c}).
   \]

2. For all $c, c' \in C$,
   \[
   \int_c^{c'} \left( \int_D D(\bar{p}_D(z))(\sigma(dD|z) - \sigma(dD|c')) \right) dz \geq 0.
   \]

Furthermore, $\bar{p}$ can be replaced by any $p \in P$ 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 revenue-equivalence formula stating that the transfer $\tau$ must be determined by $\sigma$ 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), Karsikov and Lamba (2020)). This condition, rather than the usual monotonicity condition, is needed because the allocation space is infinite dimensional and the producer’s profit is not single-crossing in general.

From Lemma 1, for any incentive compatible mechanism $(\sigma, \tau)$, the data broker’s expected revenue can be written as

\[
E_G[\tau(c)] = \int_C \left( \int_D (\bar{p}_D(c) - \phi_G(c)) D(\bar{p}_D(c))\sigma(dD|c) \right) G(dc) - U(\bar{c}), \tag{6}
\]

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 $(\sigma, \tau)$ 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 $\sigma$.

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 $(\sigma, \tau)$ to be individually rational, it must be that $U(\bar{c}) \geq \bar{\pi} := \pi_{D_0}(\bar{c})$. Moreover, for any $c \in C$,

\[
\int_D (\bar{p}_D(c) - \phi_G(c)) D(\bar{p}_D(c))\sigma(dD|c) \leq \int_{p \in \mathbb{R}_+} \max\{(p - \phi_G(c))D(p)\} \sigma(dD|c) \\
\leq \int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(dv),
\]
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

$$\bar{R} := \int_C \left( \int_{\{v \geq \phi G(c)\}} (v - \phi G(c)) D_0(dv) \right) G(dc) - \bar{\pi}$$

$$\geq \int_C \left( \int_D (\bar{p}_D(c) - \phi G(c)) D(\bar{p}_D(c)) \sigma(dD|c) \right) G(dc) - U(\bar{c})$$

$$= \mathbb{E}_G[\tau(c)].$$

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 $\psi : C \rightarrow \mathbb{R}_+$ and for any $\psi$-quasi-perfect scheme $\sigma$, 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 $v \geq \psi(c)$.

Namely, for all $c \in C$,

$$\int_D \bar{p}_D(c) D(\bar{p}_D(c)) \sigma(dD|c) = \int_{\{v \geq \psi(c)\}} v D_0(dv)$$

and

$$\int_D D(\bar{p}_D(c)) \sigma(dD|c) = D_0(\psi(c)).$$

Therefore, if there is an incentive feasible $\phi G$-quasi-perfect mechanism $(\sigma, \tau)$, then by Lemma 1, the data broker can attain revenue

$$\mathbb{E}[\tau(c)] = \int_C \left( \int_D (\bar{p}_D(c) - \phi G(c)) D(\bar{p}_D(c)) \sigma(dD|c) \right) G(dc) - \bar{\pi}$$

$$= \int_C \left( \int_{\{v \geq \phi G(c)\}} (v - \phi G(c)) D_0(dv) \right) G(dc) - \bar{\pi}$$

$$= \bar{R}.$$ 

However, not every $\phi G$-quasi-perfect scheme is implementable (even if $\phi G$ is nondecreasing, see Appendix D). To ensure incentive compatibility, the integral monotonicity condition

---

Formal arguments are in the proof of Lemma 5, which can be found in the Online Appendix.
(i.e., condition 2 of 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 \( \psi \)-quasi-perfect scheme \( \sigma \), there exists a transfer scheme \( \tau: C \to \mathbb{R} \) such that \((\sigma, \tau)\) is incentive compatible if for any \( c \in C \),

\[
\psi(z) \leq \overline{p}_D(z),
\]

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

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 \( \psi \)-quasi-perfect scheme, there is always no downward-deviation incentive (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 \( \psi \)-quasi-perfect mechanism.

**Lemma 3.** For any nondecreasing function \( \psi: C \to \mathbb{R}_+ \) such that that \( c \leq \psi(c) \leq p_0(c) \) for all \( c \in C \), there exists a \( \psi \)-quasi-perfect scheme \( \sigma \) that satisfies (10).

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

\[
\int_c^\bar{c} D_0(\phi_G(z)) \, dz \geq \int_c^\bar{c} D_0(\overline{p}_0(z)) \, dz.
\]

Together, by Lemma 1 and (5), after possibly adding a constant to \( \tau \) so that the indirect utility of the producer with cost \( \tau \) equals to \( \overline{\pi} \), \((\sigma, \tau)\) is an incentive feasible \( \phi_G \)-quasi-perfect mechanism, which in turn implies that \((\sigma, \tau)\) is optimal. Combined with (9), it then follows that any incentive feasible \( \phi_G \)-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 \( \psi \)-quasi-perfect scheme \( \sigma_n \) of each \( D_n \) through recursion. Together with a continuity property of quasi-perfect mechanisms and optimal prices, the limit of \( \{\sigma_n\} \) is then a desired \( \psi \)-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 $v \in [\psi(c), \bar{v}]$, let $D_v^{\psi(c)} \in \mathcal{D}$ be defined as (2) with $\overline{\nu}_G(c)$ replaced by $\psi(c)$. Also, let $\sigma^* : C \rightarrow \Delta(\mathcal{D})$ be defined as (3) with $\overline{\nu}_G$ replaced by $\psi$. By construction, $\sigma^*(c) \in \mathcal{S}$ for all $c \in C$. Furthermore, $\sigma^*$ is a $\psi$-quasi-perfect scheme satisfying (10). To see this, for any $c \in C$, let $p^{\psi(c)} := \min\{v \in \text{supp}(D_0) : v \geq \psi(c)\}$. By the hypothesis that $\psi(c) \leq \overline{p}_0(c)$, it must be that $p^{\psi(c)} \leq \overline{p}_0(c)$. This in turn implies that, by regularity of $D_0$ (i.e., single-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 $v \in [\psi(c), \bar{v}] \cap \text{supp}(D_0)$, 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 $v \in [\psi(c), \bar{v}] \cap \text{supp}(D_0)$. Therefore, since $(p^{\psi(c)}, v) \cap \text{supp}(D_0^{\psi(c)}) = \emptyset$, it follows that $\overline{p}_{D_v^{\psi(c)}}(c) = v$ 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 $\overline{p}_{D_v^{\psi(c)}}$ is nonincreasing, it must be that either $\overline{p}_{D_v^{\psi(c)}}(z) = v$ or $\overline{p}_{D_v^{\psi(c)}}(z) < \psi(c)$. In the former case, since $\psi$ is nondecreasing, it then follows that $\overline{p}_{D_v^{\psi(c)}}(z) = v \geq \psi(c) \geq \psi(z)$, as desired. In the latter case, since $D_v^{\psi(c)}(z) = D_0(z)$ for all $p \leq \psi(c)$ and since $p \mapsto (p-z)D_0(p)$ is single-peaked, $\overline{p}_{D_v^{\psi(c)}}(z)$ must be the largest optimal price for the producer under $D_0$ as well. That is, $\overline{p}_{D_v^{\psi(c)}}(z) = \overline{p}_0(z)$. Combined with the hypothesis that $\psi(z) \leq \overline{p}_0(z)$, this then implies that $\psi(z) \leq \overline{p}_{D_v^{\psi(c)}}(z)$, as desired. As a result, $\sigma^*$ is indeed a $\psi$-quasi-perfect scheme satisfying (10).

Combining Lemma 1, Lemma 2 and Lemma 3, it then follows that there exists an incentive feasible $\phi_G$-quasi-perfect mechanism and hence the data broker can attain revenue $\tilde{R}$, 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 $\overline{\nu}_G$-quasi-perfect mechanism $(\sigma^*, \tau^*)$ defined by (3) and Lemma 1 is incentive feasible, which, together with Theorem 1, proves Theorem 2.

Uniqueness
To see why any optimal mechanism of the data broker is $\phi_G$-quasi-perfect, suppose that $(\sigma, \tau)$ is optimal. Then,
\[
\tilde{R} = \int_C \left( \int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c))D_0(dv) \right) G(dc) - \bar{\pi} = \int_C \left( \int_{D} (\overline{p}_D(c) - \phi_G(c))D(\overline{p}_D(c))\sigma(dD|c) \right) G(dc) - \bar{\pi},
\]
which in turn implies that for (almost) all $c \in C$,
\[
\int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c))D_0(dv) = \int_{D} (\overline{p}_D(c) - \phi_G(c))D(\overline{p}_D(c))\sigma(dD|c),
\]
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 \( \sigma \) must be a \( \phi_G \)-quasi-perfect mechanism. Indeed, if \( \sigma \) is not a \( \phi_G \)-quasi-perfect scheme, then there must be a positive \( G \)-measure of \( c \in C \) and a positive \( \sigma(c) \)-measure of \( D \in \text{supp}(\sigma(c)) \) such that either \( D(v) > 0 \) for some \( v > \bar{p}_D(c) \), or \( D(\phi_G(c)) \neq D(\bar{p}_D(c)) \). 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, \( (\sigma, \tau) \) must be a \( \phi_G \)-quasi-perfect mechanism. Moreover, \( (\sigma, \tau) \) must also induce quasi-perfect price discrimination since \( \bar{p} \) can be replaced with any \( p \in P \) according to Lemma 1.

### 4.4 Further Remarks for the Proof

Although the proof of Theorem 1 resembles standard methods for one-dimensional problems in some aspects (i.e., the revenue equivalence formula (condition 1 of Lemma 1) and the fact that \( \varphi_G(c) \)-quasi-perfect segmentations solve a pointwise maximization problem which ignores the global incentives constraints), it is substantially different from standard methods due to the technical challenges posed by the infinite dimensional allocation space. Specifically, since the allocation space is the entire set of market segmentations, even pointwise maximization is infinite dimensional. The proof of Theorem 1 solves this problem by finding segmentations that attain the solution of a relaxed problem where the data broker can control prices. Furthermore, since the producer’s profit is not single-crossing in general—even when restricting attention to quasi-perfect segmentations—global incentive constraints cannot be ensured by a simple monotonicity condition (in particular, monotonicity of \( \varphi_G \)). The proof of Theorem 1 keeps track of the incentive constraints through the integral monotonicity condition (condition 2 of Lemma 1) and its sufficient conditions (Lemma 2 and Lemma 3).

Theorem 1 underlines a noteworthy feature of the optimal mechanisms. According to Theorem 1, for any optimal mechanism \( (\sigma, \tau) \), the segmentation scheme \( \sigma \) does not generate value-revealing segmentations in general. Specifically, for any report \( c \) such that \( \varphi_G(c) > \nu \), 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 not to sell to any consumers with values \( v \in [c, \varphi_G(c)] \). Consumers with values above the desirable threshold \( \varphi_G(c) \) must be assigned to segments where some consumers with values below \( \varphi_G(c) \) are also assigned to. 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 \( \varphi_G(c) \)-quasi-perfect price discrimination for all \( c \).
In addition to the pricing incentives, the ways low-value consumers are pooled with the high-values are crucial for the entire mechanism to be incentive compatible. Although the revenue equivalence formula accounts for the local incentives and hence there would be no incentives to deviate locally as long as each cost $c$ is assigned with a $\varphi_G(c)$-quasi-perfect segmentation for $c$, this does not guarantee global incentives in general (even if $\varphi_G$ is nondecreasing), since the producer’s profit is not necessarily single-crossing.\footnote{See the example in Appendix D, which demonstrates that it is possible to have a $\varphi_G$-quasi-perfect scheme that is not implementable.} In other words, the ways low-value consumers are pooled with the high-values serve two purposes at the same time. On one hand, they incentivize the producer to only sell to consumers with $v \geq \varphi_G(c)$ when reporting truthfully. On the other hand, they discourage non-local deviations by ensuring that Lemma 2 (and hence condition 2 of Lemma 1) is satisfied.

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 $(\sigma^*, \tau^*)$. 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.

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 $\varphi_G(c)$-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 compared 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.

The reason behind Theorem 4 is that while all the purchasing consumers pay their values, data brokership induces larger trade volume compared with uniform pricing. As a result, 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. Proposition 1 immediately

25For 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.)
implies that, if the data broker has to purchase data by compensating the consumers with monetary transfers before they learn their values, 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.

**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.*

### 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.

---

26It 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.

27Jones 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.
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 \((\sigma, \tau, \gamma)\) such that for all \( c, c' \in C \),

\[
\int_{\mathcal{D} \times \mathbb{R}^+} (p - c)D(p)\gamma(dp|D, c)\sigma(dD|c) - \tau(c) \geq \int_{\mathcal{D} \times \mathbb{R}^+} (p - c)D(p)\gamma(dp|D, c')\sigma(dD|c') - \tau(c')
\]

and for all \( c \in C \),

\[
\int_{\mathcal{D} \times \mathbb{R}^+} (p - c)D(p)\gamma(dp|D, c)\sigma(dD|c) - \tau(c) \geq \pi_{D_0}(c),
\]

where for each \( c \in C \), \( \sigma(c) \in S \) is the market segmentation provided to the producer, \( \tau(c) \in \mathbb{R} \) is the payment from the producer to the data broker, and \( \gamma(c) : \mathcal{D} \to \Delta(\mathbb{R}^+) \) is a transition kernel so that \( \gamma(\cdot|D, c) \) 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 brokering and vertical integration. To see this, recall that any optimal mechanism \((\sigma, \tau)\) 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 \( \text{supp}(D_0) = V \) 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 \( \varphi < \bar{c} \).

For other market regimes, it is noteworthy that since prices are contractable under price-controlling data brokership, for any mechanism \((\sigma, \tau, \gamma)\), the producer’s private marginal cost affects her profit only through the quantity produced and sold to the consumers induced by \((\sigma, \gamma)\). This effectively reduces the 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 \( \varphi_G(c) \)-quasi-perfect price discrimination for \( G \)-almost all \( c \in C \). In particular, the optimal revenue is

\[
R^* = \int \int_c \left( \int_{\{v \geq \varphi_G(c)\}} (v - \phi_G(c)) D_0(dv) \right) G(dc) - \bar{\pi}.
\]

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 brokershlp and data brokershlp 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, regardless of the practicality of the exclusive retail regime and the price-controlling data brokershlp regime: 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 $\varphi_G$-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 $\varphi_G$-quasi-perfect mechanism. Meanwhile, by the definition of $\varphi_G$, $c \leq \varphi_G(c) \leq p_0(c)$ for all $c \in C$, and hence $\varphi_G$ satisfies the condition required by **Lemma 3**. As a result, combining **Lemma 2** and **Lemma 3**, there is indeed an incentive feasible $\varphi_G$-quasi-perfect mechanism, which, by definition, generates revenue $R^*$, and hence is optimal. As noted at the end of the previous section, while $\phi_G$-quasi-perfect mechanisms may not be individually rational when (4) fails, $\varphi_G$-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 (i.e., those with $c \geq c^* = \inf\{c \in C : p_0(c) \geq \varphi_G(c)\}$) 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 θ ∈ Θ are distributed according to β₀ ∈ Δ(Θ). 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 θ ∈ Θ, suppose that among the consumers who share characteristic θ, their values are distributed according a demand Dₜₜ ∈ ℰ (i.e., Dₜₜ(p) denotes the share of consumers with values above p among those with characteristic θ). Moreover, suppose that \{supp(Dₜₜ)\}_θ∈Θ forms a partition of V and that supp(Dₜₜ) is an interval for all θ ∈ Θ. 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 ∈ V, let
\[
D₀(p) := \sum_{θ ∈ Θ} Dₜₜ(p)β₀(θ).
\]
D₀ ∈ ℰ then describes the market demand in this environment.

In this environment, a market segmentation is defined by s ∈ Δ(Δ(Θ)) such that
\[
\int_{Δ(Θ)} β(θ)s(dβ) = β₀(θ),
\]
for all θ ∈ Θ. A market segmentation s induces market segments \{Dₜₜ\}_β∈supp(s) and
\[
\int_{Δ(Θ)} Dₜₜ(p)s(dβ) = D₀(p),
\]
for all p ∈ V, where Dₜₜ(p) := \sum_{θ ∈ Θ} Dₜₜ(p)β(θ) for any β ∈ Δ(Θ) and any p ∈ 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 Dₜₜ 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 $\theta$ 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 $\{D_\theta\}_{\theta \in \Theta}$ 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 $\theta$ is fully disclosed.

The intuition behind Theorem 8 is simple. Since there are only finitely many characteris-
tics and since $\{\text{supp}(D_\theta)\}_{\theta \in \Theta}$ forms a partition of $V$, identifying the consumers’ characteristic
$\theta$ 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 $\theta$ is fully-revealed.28

In addition to the surplus extraction result, the characterization of the optimal mechan-
isms can be generalized as well. With proper regularity conditions, there is an optimal
mechanism analogous to the canonical $\lambda_G$-quasi-perfect mechanism introduced in Section 4.
To state this result, given any $\{D_\theta\}_{\theta \in \Theta}$, 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 $\theta$ such that $p \in (l(\theta), u(\theta)]$. For any $c \in C$,
let $\hat{p}_0(c)$ be the largest optimal price for the producer with marginal cost $c \in C$ under
the demand whose support contains $\bar{p}_0(c)$.29 Also, let $\hat{\varphi}_G(c) := \min\{\varphi_G(c), \hat{p}_0(c)\}$ for all $c \in C$.
Furthermore, given any function $\psi : C \to \mathbb{R}_+$, say that a mechanism $(\sigma, \tau)$ is a canonical
$\psi$-quasi-perfect segmentation if the producer with marginal cost $\tau$, when reporting truthfully,
receives $\hat{\pi}$, and if for any $c \in C$, and for any $\beta \in \text{supp}(\sigma(c))$, either

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

(13)

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

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

(14)

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

$$
\beta(\theta') = \beta_0(\theta').
$$

28A more detailed argument can be found in the proof, which is provided in the Online Appendix
29That is, $\hat{p}_0(c) := \min\{\varphi_G(c), \hat{p}_0(c)\}$. Notice that $\hat{p}_0(c) \leq \bar{p}_0(c)$ for all $c \in C$. Moreover, in the case where the
data broker can disclose all the information about the value $v$, $\hat{p}_0(c) = \bar{p}_0(c)$ for all $c \in C$. 
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 $c \mapsto \max\{g(c)(\phi_G(c) - \tilde{p}_0(c)), 0\}$ is nondecreasing and that $D_0$ is regular, there is a canonical $\hat{\varphi}_G$-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, $p_0(c)$ can be written as $p_0(c) = c + \xi_0(c)$, where $\xi_0(c) := p_0(c) - c$ 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.\(^\footnote{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.}

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 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.\(^\footnote{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 $\varphi^*$ (may not necessarily be of a closed form) such that every optimal mechanism must be a $\varphi^*$-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.}
7.2 Creating Market Segmentations by Partitioning Underlying Characteristics

Throughout the paper, a market segmentation is formalized as a probability measure $s \in S$ that splits the market demand $D_0$ into several segments $D \in D$, 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 $s \in S$ 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 assumption captures certain 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 $v \in [v, \bar{v}] = V \subseteq \mathbb{R}_+$ is heterogeneous across consumers and distributed according to $D_0 \in D$, while $\xi \in [0, v]$ is the same across consumers. All the consumers and the producer (with a
commonly known marginal cost that is normalized to zero) know $\xi$, while the data broker only knows that $\xi$ 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 $\xi$) than the data broker does. In this setting, market segmentations are defined the same way as before: A market segmentation is a probability measure $s \in S \subseteq \Delta(D)$. It then follows that under market condition $\xi$, 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 $\xi$). Thus, given a demand shifter $\xi$, under any market segment $D \in \mathcal{D}$, the producer’s pricing problem is given by

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

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 $\xi$, 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, as well as a channel for the data broker to compensate

---

32For 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.
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 facilitate 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.

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 brokering 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 $\mathcal{D} = \mathcal{D}([v, \overline{v}])$ is the collection of nonincreasing and left-continuous functions $D$ on $[v, \overline{v}]$ such that $D(v) = 1$ 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 $f : V \to \mathbb{R}$,

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

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,

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

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. $\int_D D(p)\sigma(dD|c) = D_0(p)$ for all $p \in V$ and for all $c \in C$.
2. $\sigma : C \to \Delta(\mathcal{D})$ is measurable.
3. $\int_D D(\overline{p}_D(c))\sigma(dD|c) = D_0(\psi(c))$ for all $c \in C$.
4. $\int_D \overline{p}_D(c)D(\overline{p}_D(c))\sigma(dD|c) = \int_{\{v \geq \psi(c)\}} vD_0(dv)$ 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 \((\sigma, \tau, \gamma)\) is incentive compatible if and only if:

1. There exists some \(\bar{\tau} \in \mathbb{R}\) such that for any \(c \in C\),
   \[
   \tau(c) = \int_D \int_{\mathbb{R}_+} (p - c) D(p) \gamma(dp|D,c) \sigma(dD|c) - \int_c^{\bar{\tau}} \int_D \int_{\mathbb{R}_+} D(p) \gamma(dp|D,z) \sigma(dD|z) \, dz - \bar{\tau}.
   \]

2. The function \(c \mapsto \int_D \int_{\mathbb{R}_+} D(p) \gamma(dp|D,c) \sigma(dD|c)\) 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:

\[
\sup_{q \in Q} \int_C \left( \int_0^{q(c)} \left( D_0^{-1}(q) - \phi_G(c) \right) \, dq \right) G(dc) - \bar{\pi} \tag{17}
\]

s.t. \(\bar{\pi} + \int_c^{\bar{\tau}} q(z) \, dz \geq \bar{\pi} + \int_c^{\bar{\tau}} D_0(p_0)(z) \, dz\),

where \(Q\) 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 \(\sigma\) and \(\gamma\) by

\[
q(c) = \int_{D \times \mathbb{R}_+} D(p) \gamma(dp|D,c) \sigma(dD|c),
\]

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 \(q\) and then prescribe perfect price discrimination subject to a capacity constraint \(q(c)\), 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 \(q\); the monotonicity condition \(q \in Q\) 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 \( \mu^* \) on \( C \) and a feasible \( \mathbf{q}^* \in \mathbf{Q} \) such that \( \mathbf{q}^* \) is a solution of

\[
\sup_{\mathbf{q} \in \mathbf{Q}} \left[ \int_C \left( \int_0^{q(c)} (D_0^{-1}(q) - \phi_G(c)) \, dq \right) G(dc) - \bar{\pi} + \int_C \left( \int_{c}^{q} (q(z) - D_0(\bar{\mathbf{p}}_0(z))) \, dz \right) \mu^*(dc) \right]
\]

and that

\[
\int_C \left( \int_{c}^{q} (q(z) - D_0(\bar{\mathbf{p}}_0(z))) \, dz \right) \mu^*(dc) = 0.
\]

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

\[
M^*(c) := \lim_{z \downarrow c} \mu(z)(\phi_G(z) - \bar{p}_0(z))^+, \forall c \in C.
\]

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

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

\[
\sup_{\mathbf{q} \in \mathbf{Q}} \left[ \int_C \left( \int_0^{q(c)} (D_0^{-1}(q) - \phi_G(c)) \, dq \right) G(dc) - \bar{\pi} - \int_C M^*(c)D_0(\bar{\mathbf{p}}_0(c)) \, dc \right],
\]

where \( \bar{\phi}_G := \min\{\phi_G, \bar{p}_0\} \).

To solve (21), let \( \varphi_G \) be the ironed virtual cost. That is, \( \varphi_G \) 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 \( H(q) := \int_0^q h(s) \, ds \) and \( K := \text{conv}(H) \). Then, for every \( q \in [0,1] \) let \( k(q) := K'(q) \) and define \( \varphi_G \) as \( \varphi_G(c) := k(G(c)) \). Also, let \( \overline{\varphi}_G := \min\{\varphi_G, \bar{p}_0\} \). Now notice that for any \( \mathbf{q} \in \mathbf{Q} \), and for any \( c \in C \),

\[
\int_0^{q(c)} (D_0^{-1}(q) - \overline{\varphi}_G(c)) \, dq = \int_0^{q(c)} (D_0^{-1}(q) - \varphi_G(c)) \, dq + (\varphi_G(c) - \overline{\varphi}_G(c))q(c).
\]

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

\[
\int_C (\overline{\varphi}_G(c) - \varphi_G(c)) \mathbf{q}(c)G(dc) = \int_{c^*}^c (\varphi_G(c) - \phi_G(c)) \mathbf{q}(c)G(dc) = - \int_{c^*}^c (K(G(c)) - H(G(c))) \mathbf{q}(dc) \leq 0,
\]

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 \( \overline{\varphi}_G(c) = \overline{\phi}_G(c) = \bar{p}_0(c) \) for all \( c > c^* \), which is due to Assumption 1, and the inequality follows from the fact that \( K = \text{conv}(H) \) and that \( \mathbf{q} \) is nonincreasing for any \( \mathbf{q} \in \mathbf{Q} \).

Meanwhile, notice that

\[
\int_C \left( \int_0^{q(c)} (D_0^{-1}(q) - \overline{\varphi}_G(c)) \, dq \right) G(dc) \leq \int_C \left( \int_0^{D_0(\overline{\varphi}_G(c))} (D_0^{-1}(q) - \overline{\varphi}_G(c)) \, dq \right), \forall \mathbf{q} \in \mathbf{Q}.
\]

In addition, since \( \overline{\varphi}_G(c) = \overline{\phi}_G(c) = \bar{p}_0(c) \) for all \( c \in (c^*, \bar{c}] \) and since \( K(G(c)) < H(G(c)) \) on an interval \([c_1, c_2] \subseteq [c, c^*] \) 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

\[
\int_C (\overline{\varphi}_G(c) - \phi_G(c))D_0(\overline{\varphi}_G(c))G(dc) = - \int_{c^*}^c (K(G(c)) - H(G(c)))D_0 \circ \overline{\varphi}_G(dc) = 0.
\]
Together with (22), and (23), (24) for any $q \in Q$,
\[
\int_C \left( \int_0^q (D_0^{-1}(q) - \phi_G(c)) dq \right) G(dc) \leq \int_C \left( \int_0^{D_0(\phi_G(c))} (D_0^{-1}(q) - \phi_G(c)) dq \right) G(dc).
\]

Also, since $\phi_G$ is nondecreasing by definition, $D_0 \circ \phi_G$ is indeed a solution of (21) and hence a solution of (18).

Moreover, since $\phi_G \leq p_0$, for all $c \in C$, $\int_C D_0(\phi_G(z)) dz \geq \int_C D_0(p_0(z)) dz$. Therefore, $D_0 \circ \phi_G \in Q$ is feasible in the primal problem (17). Meanwhile, since $M^*(c) = 0$ for all $c \in [\underline{c}, \overline{c}]$ and since $\phi_G(c) = \overline{p_0}(c)$ for all $c \in (c^*, \overline{c}]$, the complementary slackness condition (19) is also satisfied. Together, $D_0 \circ \phi_G$ is indeed a solution of (17). Finally, by definition of $D_0^{-1}$, it then follows that
\[
R^* = \int_C \left( \int_0^{D_0(\phi_G(c))} (D_0^{-1}(q) - \phi_G(c)) dq \right) G(dc) - \bar{\pi} = \int_C \left( \int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(dv) \right) G(dc) - \bar{\pi}.
\]

The see that any solution of the price-controlling data broker’s problem must induce $\phi_G(c)$-quasi-perfect price discrimination for $G$ almost all $c \in C$, consider any optimal mechanism $(\sigma, \tau, \gamma)$ of the price-controlling data broker. By optimality, it must be that $E_G[\tau(c)] = R^*$ and that the indirect utility of the producer with marginal cost $\bar{c}$ is $\bar{\pi}$. Thus, by Lemma 7, it must be that
\[
\int_C \left( \int_{\mathbb{R}_+} (p - \phi_G(c)) D(p) \gamma(dp|D, c) \right) \sigma(dD|c) G(dc) = \int_C \left( \int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(dv) \right) G(dc),
\]
which is equivalent to
\[
\int_C \left( \int_{\mathbb{R}_+} (p - \phi_G(c)) D(p) \gamma(dp|D, c) \sigma(dD|c) \right) G(dc) + \int_C (\phi_G(c) - \phi_G(c)) \eta_G(c) G(dc)
\]
\[
= \int_C \left( \int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(dv) \right) G(dc) + \int_C (\phi_G(c) - \phi_G(c)) D_0(\phi_G(c)) G(dc),
\]
where $\eta_G(c) := \int_{D \times \mathbb{R}_+} D(p) \gamma(dp|D, c) \sigma(dD|c)$ for all $c \in C$. Moreover, since for any $c \in C$,
\[
\int_{D \times \mathbb{R}_+} (p - \phi_G(c)) D(p) \gamma(dp|D, c) \sigma(dD|c) \leq \int_{\mathbb{R}_+} \max\{p - \phi_G(c)\} |D(p)| \sigma(dD|c) \leq \int_V (v - \phi_G(c))^+ D_0(dv),
\]
it must be that
\[
\int_C (\phi_G(c) - \phi_G(c)) \eta_G(c) G(dc) \geq \int_C (\phi_G(c) - \phi_G(c)) D_0(\phi_G(c)) G(dc).
\]
Meanwhile, since $(\sigma, \tau, \gamma)$ is incentive compatible, Lemma 6 implies that $\eta_G$ is nonincreasing in $c$. Together with (23) and (25), we have
\[
\int_C (\phi_G(c) - \phi_G(c)) \eta_G(c) G(dc) \geq \int_C (\phi_G(c) - \phi_G(c)) D_0(\phi_G(c)) G(dc).
\]
Furthermore, since $\bar{\phi}_G(c) = \overline{p_0}(c) \leq \phi_G(c)$ for all $c \in (c^*, \overline{c}]$ and $\phi_G(c) = \phi_G(c)$, for all $c \in [\underline{c}, c^*]$, by the definition of $M^*$ given by (20), together with integration by parts, (29) is equivalent to
\[
\int_C \left( \int_{c} \eta_G(z) - D_0(\overline{p_0}(c)) dz \right) M^*(dc) \leq 0
\]
Lastly, since \((\sigma, \tau, \gamma)\) is individually rational, for any \(c \in C\),
\[
\int \sigma' \left( q_c^*_\sigma(z) - D_0(\bar{p}_0(z)) \right) \, dz \geq 0.
\]

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, \((\sigma, \tau, \gamma)\) must induce \(\varphi_G(c)\)-quasi-perfect price discrimination for \(G\)-almost all \(c \in C\), as desired. \(\blacksquare\)

### B.3 Proof of Lemma 1

**Proof of Lemma 1.** For necessity, consider any incentive compatible mechanism \((\sigma, \tau)\). 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 \(\pi'_D(c) = -D(p_D(c))\) for all \(p \in P\) and for almost all \(c \in C\). Moreover, since for any \(D \in \mathcal{D}\) and for any \(p \in P\), \(|\pi'_D(c)| = |D(p_D(c))| \leq 1\), for almost all \(c \in C\), the order of integral and differential can be interchanged. That is, for any \(c, c' \in C\),
\[
\frac{d}{dc} \int_D \pi_D(c) \sigma(dD|c') = \int_D \pi'_D(c) \sigma(dD|c') = -\int_D (p_D(c)) \sigma(dD|c'). \tag{31}
\]

As such, for any \(c' \in C\), the function \(c \mapsto \int_D \pi_D(c) \sigma(dD|c')\) is convex and, by (31), has an almost-everywhere derivative \(-\int_D (p_D(c)) \sigma(dD|c')\), for any \(p \in P\). Now let \(u(c, c') := \int_D \pi_D(c) \sigma(dD|c') - \tau(c')\) 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 \(\pi_D\) is convex and continuous for all \(D \in \mathcal{D}\). Furthermore, since the mechanism \((\sigma, \tau)\) is incentive compatible, by the envelope theorem (Milgrom and Segal, 2002), let \(U(c) := u(c, c)\), we then have
\[
U(c) = U(\bar{c}) - \int_c^\bar{c} \frac{\partial}{\partial c} u(z, z) \, dz = U(\bar{c}) + \int_c^\bar{c} \left( \int_D (p_D(z)) \sigma(dD|z) \right) \, dz. \tag{32}
\]

Assertion 1 then follows after rearranging.

Furthermore, for any mechanism \((\sigma, \tau)\) satisfying assertion 1 (and hence (32)) with any \(p \in P\), we have
\[
U(c) - u(c, c') = (U(c) - U(c')) + \int_D (\pi_D(c) - p_D(c')) \sigma(dD|c')
\]
\[
= \int_c^{c'} \left( \int_D (p_D(z)) \sigma(dD|z) - \int_D (p_D(z)) \sigma(dD|c') \right) \, dz
\]
\[
= \int_c^{c'} \left( \int_D (p_D(z)) (\sigma(dD|z) - \sigma(dD|c')) \right) \, dz,
\]
where the second equality follows from the fundamental theorem of calculus and (31). Therefore, for any mechanism \((\sigma, \tau)\) satisfying assertion 1 with any \(p \in P\), \(U(c) \geq u(c, c')\) for all \(c, c' \in C\) if and only if assertion 2 holds. This completes the proof. \(\blacksquare\)

### B.4 Proof of Lemma 2

**Proof of Lemma 2.** Given any nondecreasing function \(\psi : C \to \mathbb{R}_+\), and any \(\psi\)-quasi-perfect scheme \(\sigma : C \to \mathcal{S}\), suppose that for any \(c \in C\), \(\psi(z) \leq \bar{p}_D(z)\), for Lebesgue almost all \(z \in [c, c]\) and for all \(D \in \text{supp}(\sigma(c))\).
Then, for any \( c, c' \in C \) with \( c < c' \),
\[
\int_c^{c'} \left( \int_D D(\overline{p}_D(z))(\sigma(dD|z) - \sigma(dD|c')) \right) \, dz = \int_c^{c'} \left( D_0(\psi(z)) - \int_D D(\overline{p}_D(z))\sigma(dD|c') \right) \, dz \\
\geq \int_c^{c'} \left( D_0(\psi(z)) - \int_D D(\psi(z))\sigma(dD|c') \right) \, dz \\
= \int_c^{c'} (D_0(\psi(z)) - D_0(\psi(z))) \, dz \\
= 0,
\]
where the first equality follows from assertion 3 of Lemma 5, the inequality follows from the hypothesis, and the second equality follows from \( \sigma(z) \in S \) for all \( z \in [c, c'] \). Meanwhile, for any \( c, c' \in C \) with \( c > c' \),
\[
\int_c^{c'} \left( \int_D D(\overline{p}_D(z))(\sigma(dD|c) - \sigma(dD|z)) \right) \, dz = \int_c^{c'} \left( \int_D D(\overline{p}_D(z))\sigma(dD|c) - D_0(\psi(z)) \right) \, dz \\
= \int_c^{c'} (\min\{D_0(\psi(c)), D_0(z)\} - D_0(\psi(z))) \, dz \\
\geq 0,
\]
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.\(^{33}\) Therefore, by Lemma 1, there exists a transfer \( \tau \) such that \((\sigma, \tau)\) is incentive compatible, as desired.

\[\boxed{\text{B.5 Proof of Theorem 1}}\]

**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 \( \overline{R}^* \). Since \( \overline{R}^* \) 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 \( \overline{R}^* \). To this end, notice that since \( c \leq \overline{\varphi}_G(c) \leq \overline{p}_0(c) \) 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 \( \overline{\sigma} : C \to \mathcal{S} \) that satisfies (10). Together with Lemma 2, there exists a transfer \( \overline{\tau} \) such that \((\overline{\sigma}, \overline{\tau})\) is incentive compatible. Meanwhile, by possibly adding a constant to the transfer \( \overline{\tau} \) so that the indirect utility of the producer with cost \( \overline{\tau}, U(\overline{\tau}) \), equals to \( \overline{\pi} \) under the mechanism \((\overline{\sigma}, \overline{\tau})\), it must be that, for any \( c \in C \),
\[
\int_D \pi_D(c)\overline{\sigma}(dD|c) - \overline{\tau}(c) \, dz = U(\overline{\tau}) + \int_c^{\overline{\tau}} \left( \int_D D(\overline{p}_D(z))\overline{\sigma}(dD|z) \right) \, dz \\
= \overline{\pi} + \int_c^{\overline{\tau}} D_0(\overline{\varphi}_G(z)) \, dz \\
\geq \overline{\pi} + \int_c^{\overline{\tau}} D_0(\overline{p}_0(z)) \, dz \\
= \pi_{D_0}(c),
\]

\(^{33}\)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 \( D \in \text{supp}(\sigma(c)) \), if \( D(c) > 0 \) and \( \max(\text{supp}(D)) \geq z \), then \( \overline{p}_D(z) = \overline{p}_D(c) \) and hence \( D(\overline{p}_D(z)) = D_0(\psi(c)) = D_0(z) \); if \( D(c) > 0 \) and \( \max(\text{supp}(D)) < z \), then \( D(\overline{p}_D(z)) = 0 \); 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 $\varphi_G \leq \overline{p}_0$ and the last equality follows from (5). As a result, $(\overline{\sigma}, \overline{\tau})$ is individually rational.

Furthermore, since $\overline{\sigma} : C \to S$ is a $\varphi_G$-quasi-perfect scheme, by assertion 3 and assertion 4 of Lemma 5, for any $c \in C$,
\[
\int_{D} (\overline{p}_D(c) - \phi_G(c)) D(\overline{p}_D(c)) \overline{\sigma}(dD|c) = \int_{\{v \geq \varphi_G(c)\}} (v - \phi_G(c)) D_0(dv). \quad (33)
\]
and therefore, together with Lemma 1,
\[
\mathbb{E}\overline{\tau}(c) = \int_{C} \left( \int_{D} (\overline{p}_D(c) - \phi_G(c)) D(\overline{p}_D(c)) \overline{\sigma}(dD|c) \right) G(dc) - \overline{\pi} \\
= \int_{C} \left( \int_{\{v \geq \varphi_G(c)\}} (v - \phi_G(c)) D_0(dv) \right) G(dc) - \overline{\pi} \\
= R^*,
\]
as desired.

Since the data broker’s optimal revenue is $R^*$ and since (33) holds for any $\varphi_G$-quasi-perfect scheme $\sigma$, by Lemma 1, any incentive feasible $\varphi_G$-quasi-perfect mechanism must give revenue $R^*$ and hence is optimal.

Conversely, to see why any optimal mechanism must be a $\varphi_G$-quasi-perfect mechanism, consider any optimal mechanism $(\sigma, \tau)$. As it is optimal and incentive compatible, by Lemma 1,
\[
R^* = \mathbb{E}\overline{\tau}(c) = \int_{C} \left( \int_{D} (p_D(c) - \phi_G(c)) D(p_D(c)) \sigma(dD|c) \right) G(dc) - \overline{\pi}, \quad (34)
\]
for any $p \in P$. Also, since $(\sigma, \tau)$ is incentive compatible, for any $p \in P$, the function
\[
c \mapsto \int_{D} D(p_D(c)) \sigma(dD|c)
\]
is nonincreasing on $C$.\textsuperscript{34} Thus, by (23),
\[
\int_{C} \overline{\phi}_G(c) \left( \int_{D} D(p_D(c)) \sigma(dD|c) \right) G(dc) \geq \int_{C} \overline{\varphi}_G(c) \left( \int_{D} D(p_D(c)) \sigma(dD|c) \right) G(dc). \quad (35)
\]
Moreover, since $(\sigma, \tau)$ is individually rational, by Lemma 1, it must be that
\[
\int_{c} \overline{\pi} \left( \int_{D} D(p_D(z)) \sigma(dD|z) \right) dz \geq \int_{c} \overline{\pi} D_0(\overline{p}_0(z)) dz, \forall c \in C. \quad (36)
\]

Now suppose that $(\sigma, \tau)$ is not a $\varphi_G$-quasi-perfect mechanism or it does not induce $\varphi_G(c)$-quasi-perfect price discrimination for a positive $G$-measure of $c$, then there exists $p \in P$, a positive $G$-measure of $c$ and a positive $\sigma(c)$-measure of $D \in D$ such that either $p_D(c) < \overline{p}_D(c)$, or $D(c) > 0$ and either $\#\{v \in \text{supp}(D) :$
\textsuperscript{34}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 $c \mapsto \int_{D} D(p_D(c)) \sigma(dD|c)$ must be nonincreasing.
\( v \geq \overline{\varphi}_G(c) \neq 1 \) or \( \max(\text{supp}(D)) \notin P_D(c) \), which imply that there is a positive \( G \)-measure of \( c \) and a positive \( \sigma(c) \)-measure of \( D \) such that

\[
\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \overline{\varphi}_G(c)) D(dv) \geq \int_{\{v \geq p_D(c)\}} (v - \overline{\varphi}_G(c)) D(dv)
\]

\[
= (p_D(c) - \overline{\varphi}_G(c)) D(p_D(c)) + \int_{\{v \geq p_D(c)\}} (v - p_D(c)) D(dv)
\]

\[
geq (p_D(c) - \overline{\varphi}_G(c)) D(p_D(c)),
\]

with at least one inequality being strict. Therefore,

\[
\int_C \left( \int_D (p_D(c) - \overline{\varphi}_G(c)) D(p_D(c)) \sigma(dD|c) \right) G(dc) < \int_C \left( \int_V (v - \overline{\varphi}_G(c))^+ D_0(dv) \right) G(dc).
\] (37)

Meanwhile, since by (34)

\[
\int_C \left( \int_D (p_D(c) - \overline{\varphi}_G(c)) D(p_D(c)) \sigma(dD|c) \right) G(dc) + \int_C (\overline{\varphi}_G(c) - \phi_G(c)) \left( \int_D D(p_D(c)) \sigma(dD|c) \right) G(dc)
\]

\[
= \int_D \left( \int_D (p_D(c) - \phi_G(c)) D(p_D(c)) \sigma(dD|c) \right) G(dc)
\]

\[
= \int_C \left( \int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(dv) \right) G(dc)
\]

\[
= \int_C \left( \int_V (v - \overline{\varphi}_G(c))^+ D_0(dv) \right) G(dc) + \int_C (\overline{\varphi}_G(c) - \phi_G(c)) D_0(\overline{\varphi}_G(c)) G(dc),
\]

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

\[
\int_C (\overline{\phi}_G(c) - \phi_G(c)) \left( \int_D D(p_D(c)) \sigma(dD|c) \right) G(dc) \geq \int_C (\overline{\varphi}_G(c) - \phi_G(c)) \left( \int_D D(p_D(c)) \sigma(dD|c) \right) G(dc)
\]

\[
> \int_C (\overline{\varphi}_G(c) - \phi_G(c)) D_0(\overline{\varphi}_G(c)) G(dc)
\]

\[
= \int_C (\overline{\phi}_G(c) - \phi_G(c)) D_0(\overline{\phi}_G(c)) G(dc),
\]

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

\[
\int_{c^*}^{\overline{c}} (\phi_G(c) - \overline{p}_0(c)) \left( \int_D D(p_D(c)) \sigma(dD|c) \right) G(dc) < \int_{c^*}^{\overline{c}} (\phi_G(c) - \overline{p}_0(c)) D_0(\overline{p}_0(c)) G(dc),
\]

Using integration by parts, this is equivalent to

\[
\int_{c^*}^{\overline{c}} \left( \int_{c^*}^{\overline{c}} \left( \int_D D(p_D(z)) \sigma(dD|z) \right) dz \right) M^*(dc) < \int_{c^*}^{\overline{c}} \left( \int_{c^*}^{\overline{c}} D_0(\overline{p}_0(z)) dz \right) M^*(dc),
\]

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,

\[
\int_{c^*}^{\overline{c}} \left( \int_{c^*}^{\overline{c}} D(p_D(z)) \sigma(dD|z) \right) dz \right) M^*(dc) \geq \int_{c^*}^{\overline{c}} \left( \int_{c^*}^{\overline{c}} D_0(\overline{p}_0(z)) dz \right) M^*(dc),
\]

a contradiction. Therefore, \( \sigma \) 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 \( U(\overline{\sigma}) = \overline{\pi} \) under any optimal mechanism, \((\sigma, \tau)\) 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 \( c \leq \varphi_G(c) \leq \bar{p}_0(c) \) for all \( c \in \mathcal{C} \), the canonical \( \varphi_G \)-quasi-perfect scheme \( \sigma^* \) defined in (3) is implementable. Therefore, there exists \( \tau^* \) such that \( (\sigma^*, \tau^*) \) is an incentive feasible \( \varphi_G \)-quasi-perfect mechanism. By Theorem 1, \( (\sigma^*, \tau^*) \) is optimal.

C Proofs of Other Main Results

C.1 Proof of Theorem 3

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

\[
\int_C \left( \int_D \left( \int_{\{v \geq p_D(c)\}} (v - p_D(c)) D(dv) \right) \sigma^*(dD|c) \right) G(dc) = \int_C \left( \int_D \left( \int_{p_D(c)}^\infty D(z) dz \right) \sigma^*(dD|c) \right) G(dc) = 0,
\]
as desired.

C.2 Proof of Proposition 1

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

\[
\int_C \left( \int_D \left( \int_{\{v \geq \varphi_G(c)\}} (v - \varphi_G(c)) D_0(dv) \right) G(dc) - \bar{\pi} - \int_C \left( \int_D \left( \int_{\{v \geq \bar{p}_0(c)\}} (v - \bar{p}_0(c)) D_0(dv) \right) G(dC) \right)
\]

\[
= \int_C \left( \int_{\{v \geq \varphi_G(c)\}} (\bar{p}_0(c) - \varphi_G(c)) D_0(\bar{p}_0(c)) + \int_{\{v \geq \varphi_G(c)\}} (v - \varphi_G(c)) D_0(dv) \right) - \bar{\pi}
\]

\[
= \int_C \left( \int_{\varphi_G(c)}^\infty D_0(\bar{p}_0(z)) dz - \frac{G(c)}{g(c)} D_0(\bar{p}_0(c)) \right) G(dc)
\]

\[
+ \int_C \left( \int_{\{v \geq \varphi_G(c)\}} (v - \varphi_G(c)) D_0(dv) + \int_C (\varphi_G(c) - \varphi_G(c))(D_0(\varphi_G(c)) - D_0(\bar{p}_0(c))) G(dc) \right)
\]

\[
\geq \int_C G(c)(D_0(\bar{p}_0(c)) - D_0(\bar{p}_0(c))) dc + \int_C (\varphi_G(c) - \varphi_G(c)) D_0(\varphi_G(c)) G(dc)
\]

\[
\geq 0.
\]

where the second equality follows from Lemma 1, the first inequality follows from the fact that \( \varphi_G(c) < \bar{p}_0(c) \) if and only if \( \varphi_G(c) < \bar{p}_0(c) \), 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 $R^\ast$. Furthermore, for any optimal mechanism $(\sigma, \tau)$ of the data broker and any optimal mechanism $(\hat{\sigma}, \hat{\tau}, \hat{\gamma})$ of the price-controlling data broker, both of them must induce $\varphi_G(c)$-quasi-perfect price discrimination for almost all $c \in C$. In particular, for $G$-almost all $c \in C$, all the consumers with $v \geq \varphi_G(c)$ buys the product by paying their values and all the consumers with $v < \varphi_G(c)$ do not buy the product. Thus, the consumer surplus and the allocation of the product induced by $(\sigma, \tau)$ and $(\hat{\sigma}, \hat{\tau}, \hat{\gamma})$ are the same.

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

$$\int_D \pi_D(c) \sigma(dD|c) - \tau(c) = \bar{\pi} + \int_c^\pi \left( \int_D (\varphi_D(z)) \sigma(dD|z) \right) dz = \bar{\pi} + \int_c^\pi D_0(\varphi_G(z)) dz. \quad (38)$$

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

$$\int_D \left( \int_{\mathbb{R}_+} (p - c) D(p) \hat{\gamma}(dp|D, c) \right) \hat{\sigma}(dD|c) - \hat{\tau}(c) = \bar{\pi} + \int_c^\pi q_{\gamma}^c(z) dz = \bar{\pi} + \int_c^\pi D_0(\varphi_G(z)) dz. \quad (39)$$

Thus, the producer’s profit under both $(\sigma, \tau)$ and $(\hat{\sigma}, \hat{\tau}, \hat{\gamma})$ are the same. This completes the proof. \blacksquare

D Counterexample: Producer’s Profit Is Not Single-Crossing

This example demonstrates the fact that the producer’s profit, as a function of market segmentation and marginal cost, does not exhibit the single-crossing property—even when restricting to the set of quasi-perfect segmentations and ordering them by the cutoff $\kappa$. Formally, let $\geq_B$ denote the Blackwell order on $\mathcal{S}$.\(^{35}\) Meanwhile, define the following two orders over the family of quasi-perfect segmentations. Let $s$ be a $\kappa$-quasi-perfect segmentation for $c \geq 0$, and let $s'$ be a $\kappa'$-quasi-perfect segmentation for $c' \geq 0$. Say that $s \geq_{QP} s'$ if $\kappa \leq \kappa'$, and that $s \geq_{QP} s'$ if $\kappa \leq \kappa'$ and $c \leq c'$. That is, $\geq_{QP}$ is a (total) order on the family of quasi-perfect segmentations (regardless of cost, and hence regardless of pricing incentives) implied by their cutoffs $\kappa$; whereas $\geq_{QP}^+$ is a (partial) order on the same family when costs (and hence pricing incentives) are further taken into account. Note that for any nondecreasing function $\psi : C \rightarrow \mathbb{R}_+$ with $\psi(c) \geq c$ for all $c$, a $\psi$-quasi-perfect scheme $\sigma$ is monotone in both $\geq_{QP}$ and $\geq_{QP}^+$.

Below, I show that there exists a (regular) market demand $D_0$, two costs $c_L < c_H$, and two market segmentations $s_L$ and $s_H$ such that $s_L \geq_B s_H$, $s_L \geq_{QP} s_H$, $s_L \geq_{QP}^+ s_H$,

$$\int_D \pi_D(c_H) s_L(dD) > \int_D \pi_D(c_H) s_H(dD)$$

\(^{35}\)That is, $s \geq_B s'$ if and only if $s$ is a mean preserving spread of $s'$. 
and yet 
\[ \int_D \pi_D(c_L)s_L(dD) = \int_D \pi_D(c_L)s_H(dD). \]

This means that the producer’s profit is not single-crossing in general, either under the the Blackwell order, or when restricting attention to quasi-perfect segmentations (even when the pricing incentives are correct so that the producer induces quasi-perfect price discrimination on path).

Let the market demand \( D_0 \) be defined as
\[
D_0(p) := \begin{cases} 
1, & \text{if } p \in [0, 1] \\
1, & \text{if } p \in (1, 2] \\
\frac{1}{2}, & \text{if } p \in (2, 3] \\
0, & \text{if } p > 3 
\end{cases}
\]

Notice that \( D_0 \) is regular as the function \( p \mapsto (p - c)D_0(p) \) is single-peaked on \( \text{supp}(D_0) = \{1, 2, 3\} \) for all \( c \geq 0 \).

Now consider two costs, \( c_L = 1/2 \) and \( c_H = 3/2 \), and consider two market segmentations \( s_L \) and \( s_H \), where \( s_H = \delta_{\{D_0\}} \) is the degenerate segmentation that does not segment \( D_0 \); and \( s_L \) induces two segments, \( D_L^1 \) and \( D_L^3 \), where
\[
D_L^2(p) := \begin{cases} 
1, & \text{if } p \in [0, 1] \\
\frac{1}{2}, & \text{if } p \in (1, 2] \\
0, & \text{if } p > 2 
\end{cases} \quad \text{and} \quad D_L^3(p) := \begin{cases} 
1, & \text{if } p \in [0, 1] \\
\frac{1}{3}, & \text{if } p \in (1, 3] \\
0, & \text{if } p > 3 
\end{cases}
\]

and \( s_L(\{D_L^2\}) = 3/8 \); \( s_L(D_L^3) = 5/8 \). Clearly \( s_L \geq_B s_H \).

Direct calculation shows \( P_0(c_H) = \{3\} \), \( P_0(c_L) = \{1\} \), \( P_{D_L^2}(c_L) = \{1, 2\} \), and \( P_{D_L^3}(c_L) = \{1, 3\} \), which in turn implies \( P_{D_L^2}(c_H) = \{2\} \) and \( P_{D_L^3}(c_H) = \{3\} \). Together, it follows that for any \( \kappa_L \) and \( \kappa_H \) such that \( 1 \leq \kappa_L \leq 2 < \kappa_H \leq 3 \), \( s_L \) is a \( \kappa_L \)-quasi-perfect segmentation for \( c_L \), and \( s_H \) is a \( \kappa_H \)-quasi-perfect segmentation for \( c_H \). Therefore, \( s_L \geq_{QP} s_H \) and \( s_L \geq_{QP}^* s_H \).

However,
\[
\int_D \pi_D(c_L)s_L(dD) = \frac{3}{8} \cdot \left( 1 - \frac{1}{2} \right) \cdot D_L^2(1) + \frac{5}{8} \cdot \left( 1 - \frac{1}{2} \right) \cdot D_L^3(1) = \left( 1 - \frac{1}{2} \right) \cdot D_0(1) = \int_D \pi_D(c_L)s_H(dD),
\]
where the first equality follows from \( 1 \in P_{D_L^2}(c_L) \cap P_{D_L^3}(c_L) \), and the third equality follows from \( P_0(c_L) = \{1\} \). Meanwhile,
\[
\int_D \pi_D(c_H)s_L(dD) = \frac{3}{8} \cdot \left( 2 - \frac{3}{2} \right) \cdot D_L^2(2) + \frac{5}{8} \cdot \left( 3 - \frac{3}{2} \right) \cdot D_L^3(3) = \frac{3}{4} > \frac{3}{16} = \left( 3 - \frac{3}{2} \right) \cdot D_0(3) = \int_D \pi_D(c_H)s_H(dD),
\]
where the first equality follows from \( P_{D_L^2}(c_H) = \{2\} \) and \( P_{D_L^3}(c_H) = \{3\} \), and the third equality follows from \( P_0(c_H) = \{3\} \). Thus, the producer’s profit, as a function of market segmentation and cost, is not single-crossing in general.

In fact, this example implies that the producer’s profit function does not satisfy monotone difference in general. To see this, let \( c_M := 3/4 \). Then \( P_0(c_M) = \{2\} \), \( P_{D_L^2}(c_M) = \{2\} \), and \( P_{D_L^3}(c_M) = \{3\} \) and thus
\[
\int_D \pi_D(c_M)s_L(dD) = \frac{3}{8} \cdot \left( 2 - \frac{3}{4} \right) \cdot D_L^2(2) + \frac{5}{8} \cdot \left( 3 - \frac{3}{4} \right) \cdot D_L^3(3) = \frac{7}{16},
\]
and
\[ \int_D \pi_D(c_M)s_H(dD) = \left(2 - \frac{3}{4}\right) D_0(2) = \frac{5}{16}. \]
Together, it follows that \( c_L < c_M < c_H \), and yet
\[ \int_D \pi_D(c_L)s_L(dD) - \int_D \pi_D(c_L)s_H(dD) = 0 < \frac{1}{8} = \int_D \pi_D(c_M)s_L(dD) - \int_D \pi_D(c_M)s_H(dD) \]
while
\[ \int_D \pi_D(c_M)s_L(dD) - \int_D \pi_D(c_M)s_H(dD) = \frac{1}{8} > \frac{1}{16} = \int_D \pi_D(c_H)s_L(dD) - \int_D \pi_D(c_H)s_H(dD). \]

Furthermore, this example also implies that any segmentation scheme \( \sigma : C \to S \) with \( \sigma(c_L) = s_L \) and \( \sigma(c_H) = s_H \) is not implementable, even if it is monotone under \( \geq_B, \geq_{QP} \), and \( \geq_{QP} \). Indeed, if \( \sigma \) can be implemented by \( \tau \), then the incentive constraint for \( c_L \),
\[ \int_D \pi_D(c_L)\sigma(dD|c_L) - \tau(c_L) \geq \int_D \pi_D(c_L)\sigma(dD|c_H) - \tau(c_H), \]
implies \( \tau(c_L) \leq \tau(c_H) \). However, from the incentive constraint for \( c_H \),
\[ \int_D \pi_D(c_H)\sigma(dD|c_H) - \tau(c_H) \geq \int_D \pi_D(c_H)\sigma(dD|c_L) - \tau(c_L), \]
it follows that
\[ 0 < \int_D \pi_D(c_H)\sigma(dD|c_L) - \int_D \pi_D(c_H)\sigma(dD|c_H) \leq \tau(c_L) - \tau(c_H), \]
a contradiction. In particular, for any nondecreasing function \( \psi \) on \( C = [0, c_H] \) such that \( \psi(c) \geq c \) for all \( c \), and that \( 1 < \psi(c_L) \leq 2 < \psi(c_H) \leq 3 \), any \( \psi \)-quasi-perfect scheme \( \sigma \) with \( \sigma(c_L) = s_L \) and \( \sigma(c_H) = s_H \) is not implementable. This demonstrates that monotonicity of the cutoff function \( \psi \) is not sufficient for implementability of a \( \psi \)-quasi-perfect scheme.

Finally, it is noteworthy that the exact values of \( D_0, c_L, c_H, s_L \) and \( s_H \) are not essential for this counterexample, the crucial part is the fact that \( c_L \) has multiple optimal prices under both segments \( D_L^2 \) and \( D_L^3 \). This suggests the example here is generic.

References


