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

Kai Hao Yang*

October 26, 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 buy by paying their values while the rest of consumers do not buy. Consequently, vertical integration between the data broker and the producer increases total surplus while leaving the consumer surplus unchanged and brokership improves total surplus compared with uniform pricing. Furthermore, the characterization of optimal mechanisms implies that market outcomes remain unchanged even if the data broker becomes more powerful—either by gaining the ability to sell access to consumers or by becoming a retailer who purchases the product and obtains the exclusive right to sell to the consumers directly.

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

*School of Management, 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, Joyee Deb, Alex Frankel, Andrew Gianou, Johannes Hörner, Andreas Kleiner, Jacob Leshno, Elliot Lipnowski, Alejandro Manelli, Roger Myerson, Barry Nalebuff, Michael Ostrovsky, Marek Pycia, Daniel Rappoport, Eric Rasmusen, Ilya Segal, Philipp Strack, Andy Skrzypacz, Wenji Xu, Weijie Zhong and Jidong Zhou. I also thank the participants of several conferences and seminars at which this paper was presented, and Stephan Waizmann for his research assistance. 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 a 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 including government records, financial activities, online activities, and medical records to retailers (Federal Trade Commission, 2014).\(^2\)

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 by a commonly known market demand. Into this environment, I introduce a data broker, who does not know the producer’s marginal cost of production but can sell any market segmentation to the producer via any selling mechanism.

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

\(^1\)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

\(^2\)More specifically, according to the Federal Trade Commission report (Federal Trade Commission, 2014), a major product of such data brokers is data append, where retailers submit identifying information to the data broker first. Then, the data broker offers a menu of additional variables (e.g., age, religious affiliation, technology interest, gender) from which retailers can select. After selecting variables, retailers then pay the data broker and the data broker appends these selected variables to the list submitted by retailers.
cost-dependent cutoff. That is, all consumers with values above the cutoff are separated from each other whereas 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. Furthermore, \textit{Theorem 2} provides a closed-form description of an optimal mechanism.

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 consumer surplus under any optimal mechanism is zero (\textit{Corollary 1}). In other words, in terms of consumer surplus, it is \textit{as if} all the information about the consumers’ values were revealed to the producer. Furthermore, \textit{Theorem 1} also allows a comparison between data brokership and uniform pricing, where no consumer data can be shared. It follows from \textit{Theorem 1} that data brokership always increases total surplus (\textit{Corollary 2}), 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 \textit{Proposition 1}).

Another set of relevant questions pertain to how different market regimes would affect market outcomes. More specifically, how would 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 \textit{Theorem 1} allows for comparisons across (i) data brokership; (ii) \textit{vertical integration}, where all the private information about production cost is revealed and the data broker merges with the producer; (iii) \textit{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) \textit{price-controlling data brokership}, where the data broker can contract with the producer on prices in addition to providing consumer data and, in particular, can sell \textit{access} to each segment of consumers in addition to selling segmentations. Using the main characterization, I show that vertical integration between the data broker and the producer increases total surplus while leaving consumer surplus unchanged (\textit{Corollary 3}). 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 (\textit{Theorem 3}).

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. Section 5 discusses the consequences of data brokership and Section 6 concludes.

\textsuperscript{3}In particular, while perfectly revealing consumers’ values is feasible, it is not optimal for the data broker, as the cutoff is higher than the marginal cost in general.
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 monopolist 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 mechanism of a data broker who sells consumer information to a producer to facilitate price discrimination.

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

Although both this paper and Wei and Green (2020) (and similarly, Esö and Szentes (2007)) have one-dimensional types 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 (purchasing decision in Wei and Green (2020) and whether to undertake a project in Esö and Szentes (2007)). As a result, the agent’s payoff, as a function of type and allocation, is non-multiplicative and non-single-crossing.

Relatedly, Acemoglu, Makhdoumi, Malekian, and Ozdaglar (forthcoming), Bergemann, Bonatti, and Gan (2021) and Ichihashi (2021) examine environments where a data broker buys data from the consumers and then sells the consumer data to downstream firms. Segura-Rodriguez (2021) studies 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 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 state space is large and the agent has a rich action space, as well as directly payoff-relevant private information. 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.

2 An Illustrative Example

To fix ideas, consider the following example. A publisher sells an advanced textbook for graduate study. Her (constant) marginal cost of production $c$ is her private information and takes two possible values, $1/4$ or $3/4$, with equal probability. There is a unit mass of consumers with three possible occupations: faculty, undergraduate, and graduate. Each of them makes up $1/3$ of the entire population. It is common knowledge that the textbook has 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 market demand $D_0$ (with $p$ on the vertical axis, adhering to the convention).

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 $c$ buys the data, her net profit would be

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

(1)

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

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

(2)

Therefore, for any $\tau$, the publisher with cost $c$ would buy the value-revealing data if and only if her net profit given by (1) is at least the optimal uniform pricing profit given by (2). It then follows that a price $\tau = 5/12$ is optimal for the data broker and the optimal revenue is 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 \( D_H + D_A = D_0 \). Given this market segmentation, it can be shown that for both marginal costs \( c \in \{1/4, 3/4\} \), price 3 is optimal in segment \( D_H \) and price 2 is optimal in segment \( D_A \). As a result, regardless of her marginal cost, the publisher will sell to all consumers with values \( v = 3 \) and \( v = 2 \) by charging exactly their values upon receiving the residential data.

With this observation, it then follows that when \( c = 1/4 \), the publisher would prefer buying the value-revealing data whereas when \( c = 3/4 \), the publisher would prefer buying the residential data. Therefore, when menu \( \mathcal{M}^* \) is provided, the data broker’s revenue is \( 11/24 \), which is strictly greater than \( 5/12 \), the optimal revenue of selling the value-revealing data alone. The intuition behind such an improvement is simple. When selling the value-revealing data alone, the publisher with lower marginal cost retains more rents because the data broker would have to incentivize the high-cost publisher to purchase. However, by creating a menu containing both the value-revealing data and the residential data, the data broker can further screen the publisher. To see this, notice that even though the residential data becomes less informative than the value-revealing data, the only extra benefit of the value-revealing data is for the publisher to be able to price discriminate the consumers with \( v = 1 \). Thus, when the publisher’s marginal cost is high (i.e., \( c = 3/4 \), the additional information given by the value-revealing data is less useful to the publisher because the gains from selling to consumers with \( v = 1 \) are small. By contrast, when the publisher has a low marginal cost (i.e., \( c = 1/4 \), the value-revealing data is more valuable to the publisher since the gains from selling to consumers with \( v = 1 \) are larger. Therefore, by providing a
menu that contains two different datasets with different prices, the data broker can screen the publisher and extract more revenue from the publisher with lower marginal cost than by selling the value-revealing data alone.

In fact, as 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 perfectly 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 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). Consumers have different values for the product. Their values $v$ belong to a bounded interval $V := [\underline{v}, \overline{v}] \subseteq \mathbb{R}_+$ and are distributed according to a continuously differentiable and decreasing market demand $D_0 : V \rightarrow [0, 1]$, where $D_0(v) = 1$, $D_0(\overline{v}) = 0$, and $D_0(p)$ denotes the share of consumers whose values are above $p$. Let $\mathcal{D}$ denote the collection of all demand functions (i.e., nonincreasing and upper-semicontinuous functions $D : \mathbb{R}_+ \rightarrow [0, 1]$ with $D(\underline{v}) = 1$ and $D(\overline{v}^+) = 0$). Each consumer knows their own value. For the rest of the paper, I assume that $D_0$ induces a decreasing marginal revenue function.$^7$

The producer has a constant marginal cost of production $c \in C = [\underline{c}, \overline{c}] \subset \mathbb{R}_+$ for some $0 \leq \underline{c} < \overline{c} < \infty$. The marginal cost $c$ is private information to the producer and follows a cumulative distribution $G$, where $G$ has a density $g > 0$ and induces a virtual (marginal) cost

---

$^7$See more details about $\mathcal{D}$ in Appendix A

$^8$This is equivalent to saying that $1 - D_0$ is regular in the sense of Myerson (1981), and that the revenue function $q \mapsto qD_0^{-1}(q)$ is concave. This assumption, as well as the continuity and strict monotonicity of $D_0$, is merely for the ease of exposition and is not necessary for any of the results. Relaxations of these assumptions can be found in the Online Appendix.
function $\phi$, defined as
\[
\phi(c) := c + \frac{G(c)}{g(c)},
\]
for all $c \in C$. Henceforth, $G$ is assumed to be regular (i.e., $\phi$ is increasing).\(^9\)

The data broker can create any market segmentation, which is a probability measure $s \in \Delta(\mathcal{D})$ that satisfies the following condition
\[
\int_{\mathcal{D}} D(p)s(dD) = D_0(p), \forall p \in V.
\]
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.\(^10\) Let $\mathcal{S}$ denote the set of market segmentations.

### 3.2 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 $D \in \text{supp}(s)$. To maximize profit, for any segment $D \in \text{supp}(s)$, the producer with marginal cost $c$ solves
\[
\max_{p \geq 0} (p - c)D(p).
\]

For any $c \in C$ and any $D \in \mathcal{D}$, let $P_D(c)$ denote the set of optimal prices for the producer with marginal cost $c$ under market segment $D$. As a convention, regard $P$ as a correspondence on

\(^9\)Similar to the assumptions for $D_0$, this assumption is not necessary, and can be relaxed by the standard ironing technique. More details can be found in an earlier version of this paper (Yang, 2020b)

\(^10\)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. However, as the motivating example demonstrates, fully revealing consumers’ values is generally not optimal. In the Online Appendix, I consider an extension where the data broker has imperfect information about the consumers’ values.
For any \( c \in C \) and any \( D \in \mathcal{D} \), let
\[
\pi_D(c) := \max_{p \geq 0} (p - c) D(p)
\]
denote the maximized profit of the producer. Also, let
\[
p_D(c) := \max P_D(c)
\]
be the largest optimal price (in \( V \)) for the producer with marginal cost \( c \) under market segment \( D \). For conciseness, let \( p_0 := p_D(0) \).

For the rest of the paper, I impose the following technical assumption on the market demand \( D_0 \) and the distribution \( G \).

**Assumption 1.** The function \( c \mapsto \max\{g(c)(\phi(c) - p_0(c)), 0\} \) is nondecreasing.

### 3.3 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.

Formally, a mechanism is a pair of (measurable) functions \((\sigma, \tau)\). 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, a function \( \sigma : C \rightarrow S \) 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})

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 \( \tau : C \rightarrow \mathbb{R} \) such that \((\sigma, \tau)\) is incentive feasible. The data broker seeks to maximize expected revenue \( \mathbb{E}[\tau(c)] \) by choosing an incentive feasible mechanism.

**11** Assumption 1 permits a wide class of \((D_0, G)\). For instance, Assumption 1 holds 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. Further discussions about and relaxations of Assumption 1 can be found in an earlier version of this paper (Yang, 2020b).

**12** Henceforth, a mechanism stands for a direct mechanism.
3.4 Discussion of 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.\(^{13}\)

Secondly, the producer’s outside option is type-dependent. This is because the producer has direct access to the consumers, and only buys 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 \(1 - D_0\), 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.

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.

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

\(^{13}\)As in standard screening problems, the producer’s profit can be regarded as a function of her type \(c \in C\) and the “allocation” \(s \in S\). A counterexample in the Online Appendix shows that this function is not single-crossing in general. Meanwhile, it is noteworthy that although Sinander (2021) 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 (2021)), a key assumption of Sinander (2021) 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 Online Appendix.
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 to 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 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 (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 $\tau$, when reporting truthfully, has net profit $\pi_{D_0}(\tau)$.

### 4.2 Characterization of the Optimal Mechanisms

With the definitions above, the main result of this paper can be stated. For any $c \in C$, let $\overline{\phi}(c) := \min\{\phi(c), p_0(c)\}$.

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

According to to Theorem 1, while there might be multiple optimal mechanisms, 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 $D(c) = 0$.

---

14 Notice that when the producer’s marginal cost is $c$, no trade occurs in market segment $D$ if and only if $D(c) = 0$. 
$v \geq \bar{\phi}(c)$ and all the purchasing consumers pay their values. In other words, the multiplicity only accounts for the off-path incentives.

Furthermore, among all the optimal mechanisms, there is a particularly straightforward one, which is described as follows: For any $c \in C$ and for any $v \geq \phi(c)$, let $D_v^{\phi(c)} \in \mathcal{D}$ be defined as

$$D_v^{\phi(c)}(p) := \begin{cases} D_0(p), & \text{if } p \in [v, \bar{\phi}(c)] \\ D_0(\bar{\phi}(c)), & \text{if } p \in (\bar{\phi}(c), v] \\ 0, & \text{if } p \in (v, \bar{\nu}] \end{cases},$$

(4)

for all $p \in V$. Then, for any $c \in C$ and for any $p \in [\bar{\phi}(c), \bar{\nu}]$, let

$$\sigma^*\left(\left\{D_v^{\phi(c)} : v \geq p\right\} \mid c \right) := \frac{D_0(p)}{D_0(\bar{\phi}(c))}.$$  

(5)

Figure 3a illustrates $\sigma^*$ by plotting generic market segments $D_v^{\phi(c)}$, $D_v'^{\phi(c)}$, and $D_v''^{\phi(c)}$ induced by $\sigma^*(c)$ (the dashed line represents the market demand).

Lastly, let $\tau^* : C \to \mathbb{R}$ be the transfer implied by the revenue-equivalence formula (see Lemma 1 below). As will be shown in Section 4.3, the mechanism $(\sigma^*, \tau^*)$ is an incentive feasible $\bar{\phi}$-quasi-perfect mechanism (and therefore is optimal, according to Theorem 1). Henceforth, I refer to the mechanism $(\sigma^*, \tau^*)$ as the canonical $\bar{\phi}$-quasi-perfect mechanism.

**Theorem 2.** The canonical $\bar{\phi}$-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 $\bar{\phi}(c)$-quasi-perfect segmentation for $c$. The willingness to pay of a producer with cost $c$ for a $\bar{\phi}(c)$-quasi-perfect segmentation for $c$ is depicted in Figure 3b. As will be shown below, for a producer with cost $c < c^* := \inf\{c \in C : p_0(c) \geq \phi(c)\}$, her payment is strictly lower than her willingness to pay (i.e., (IR) is slack); while for a producer with $c \geq c^*$, her payment equals her willingness to pay (i.e., (IR) is binding). Furthermore, when the producer’s cost is $c$, all the consumers with values $v \geq \bar{\phi}(c)$ will be assigned to different market segments (i.e., $\{D_v^{\bar{\phi}(c)}\}_{v \in [\bar{\phi}(c), \bar{\nu}]}$ under $(\sigma^*, \tau^*)$), whereas all the consumers with values $v < \bar{\phi}(c)$ are (uniformly, under $(\sigma^*, \tau^*)$) distributed across each market segment. This allows the producer to distinguish consumers with $v \geq \bar{\phi}(c)$ among each other, but not from consumers with $v < \bar{\phi}(c)$. This type of segmentation can be interpreted as consumer data that differentiate high-value consumers but not the low-value ones.\(^{15}\)

\(^{15}\)In particular, for the producer with cost $c$, the segmentation she receives under the optimal mechanism is equivalent to the value-revealing segmentation. Together with the definition of quasi-perfect mechanisms, the optimal mechanism exhibits both the no-distortion-at-the-bottom and the no-rent-at-the-top properties.
Figure 3: Market segmentation $\sigma^*(c)$

![Diagram of market segmentation](image)

Note: Panel (a) plots three (out of a continuum) of the market segments induced by $\sigma^*(c)$, while panel (b) plots the difference between the producer's profit when operating under uniform pricing (i.e., $\pi_{D_0}(c)$) and under $\sigma^*(c)$ (i.e., selling to all consumers with $v \geq \phi(c)$ by charging them their values).

As an example, notice that the menu $M^*$ 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 $5/4$, the menu $M^*$ is indeed optimal.

As another example, 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, $\phi(c) = 2c$ and $p_0(c) = (1 + c)/2$. Thus, $\phi(c) = 2c$ for all $c \in [0, 1/3]$ and $\phi(c) = (1 + c)/2$ for all $c \in (1/3, 1]$. The canonical quasi-perfect mechanism $(\sigma^*, \tau^*)$ is as follows: For each $c$, market segments $\{D_v(c)\}_{v \in [\phi(c), 1]}$ (as defined by (4)) are uniformly distributed under $\sigma^*(c)$. Moreover, for the producer with cost $c \in [0, 1]$,

\[\text{WTP for } \sigma^*(c)\]

---

\[16\text{Although consumers’ values and the producer’s cost are distributed according to measures with finite supports in this example, the economic intuitions are entirely the same as in the continuous model introduced above. For the cost distribution, there is a straightforward analogous notion of virtual cost function when the cost distribution has atoms as in standard mechanism design problems. For the market demand, see the Online Appendix for more details.}\]
her payment and net profit are:
\[
\tau^*(c) = \begin{cases} 
\frac{1}{6} - c^2, & \text{if } c \in [0, \frac{1}{3}] \\
\frac{(1-c)^2}{8}, & \text{if } c \in \left(\frac{1}{3}, 1\right]
\end{cases},
\]
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 \([\phi(c), 1]\).

### 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 \(\bar{\phi}\)-quasi-perfect mechanism is
optimal. Finally, I argue that any mechanism that gives revenue \(\bar{R}\) must be \(\bar{\phi}\)-quasi-perfect.

To highlight the main insights and avoid unnecessary complications, in this subsection, I
further impose an assumption stronger than Assumption 1. That is, throughout the remain-
ing part of Section 4.3, I assume that
\[
\phi(c) \leq p_0(c), \quad \forall c \in C.
\]
Note that (6) is a sufficient condition for Assumption 1. Also note that all the lemmas stated
in this section do not rely on this assumption, nor on Assumption 1.

With (6), \(\bar{\phi}(c) = \phi(c)\) for all \(c \in C\). In essence, (6) ensures that individual rationality
of the constructed mechanism is implied by incentive compatibility and the fact that there
is no rent at the top, effectively circumventing the complication caused by type-dependent
outside options. The proof without (6) will be explained in Section 5.3.

### Characterization of IC Mechanisms and an Upper Bound for Revenue

To begin with, first notice that a revenue-equivalence formula can be derived by properly
invoking the envelope theorem in this setting. Specifically, 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, it follows that
\[
U'(c) = \int_{D} \pi'_D(c)\sigma(dD|c).
\]
Moreover, since $\pi_D(c)$ is the optimal profit of the producer with marginal cost $c$ under segment $D$, again by the envelope theorem,

$$\pi'_D(c) = -D(p_D(c)).$$

(7)

Together,

$$U(c) = U(\bar{c}) + \int_c^\bar{c} \left( \int_D D(p_D(z))\sigma(dD|z) \right) dz, \forall c \in C.$$

This leads to Lemma 1, as stated below.

**Lemma 1.** A mechanism $(\sigma, \tau)$ is incentive compatible if and only if:

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

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

Furthermore, $p$ can be replaced by any selection of $P$ for the “only if” part.

The proof of Lemma 1 can be found in Appendix B. In essence, condition 1 is a revenue-equivalence formula stating that the transfer $\tau$ must be determined by $\sigma$ up to a constant, whereas condition 2 is reminiscent of Lemma 1 of Pavan, Segal, and Toikka (2014), and is sometimes referred to as the integral monotonicity condition that guarantees global incentive compatibility.\textsuperscript{17}

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

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

(8)

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

\textsuperscript{17}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. Similar conditions can be found 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)).
marginal cost $c$ is replaced by the virtual marginal cost $\phi(c)$ while she still prices optimally according to marginal cost $c$—by choosing an implementable scheme $\sigma$.

With (8), there is an immediate upper bound for the data broker’s revenue. First notice that since the producer’s outside option is $\pi_D(c)$ when her cost is $c$, for an incentive compatible mechanism $(\sigma, \tau)$ to be individually rational, it must be that $U(c) \geq \bar{\pi} := \pi_D(\bar{c})$. Moreover, for any $c \in C$,

$$\int_D (p_D(c) - \phi(c)) D(p_D(c)) \sigma(dD|c) \leq \int_D \max_{p \in \mathbb{R}_+} [(p - \phi(c)) D(p)] \sigma(dD|c) \leq \int_{\{v \geq \phi(c)\}} (v - \phi(c)) D_0(dv),$$

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 $\phi(c)$.

Together with (8), it then follows that

$$\bar{R} := \int_C \left( \int_{\{v \geq \phi(c)\}} (v - \phi(c)) D_0(dv) \right) G(dc) - \bar{\pi} \geq \int_C \left( \int_D (p_D(c) - \phi(c)) D(p_D(c)) \sigma(dD|c) \right) G(dc) - U(\tau) = \mathbb{E}[\tau(c)].$$

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 $\psi : C \to \mathbb{R}_+$ 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 $c \in C$,

$$\int_D p_D(c) D(p_D(c)) \sigma(dD|c) = \int_{\{v \geq \psi(c)\}} v D_0(dv) \quad (9)$$

and

$$\int_D D(p_D(c)) \sigma(dD|c) = D_0(\psi(c)). \quad (10)$$

\[^{18}\text{As a notational convention, } D_0(dv) \text{ denotes the measure uniquely associated with the CDF } 1 - D_0. \text{ See Appendix A for more details.}\]

\[^{19}\text{Formal arguments are in the proof of Lemma 6, which can be found in the Online Appendix.}\]
Therefore, if there is an incentive feasible \( \phi \)-quasi-perfect mechanism \((\sigma, \tau)\), then by Lemma 1, the data broker can attain revenue

\[
\mathbb{E}[\tau(c)] = \int_{C} \left( \int_{D} (p_D(c) - \phi(c))D(p_D(c))\sigma(dD(c)) \right) G(dc) - \bar{\tau} \\
= \int_{C} \left( \int_{\{v \geq \phi(c)\}} (v - \phi(c))D_0(dv) \right) G(dc) - \bar{\tau} \\
= \bar{R}.
\]  

(11)

However, not every \( \phi \)-quasi-perfect scheme is implementable. To ensure incentive compatibility, the integral monotonicity condition (i.e., condition 2 of Lemma 1) must be satisfied, to which the following lemma provides a simple 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 \( \tau : C \to \mathbb{R} \) such that \((\sigma, \tau)\) is incentive compatible if for any \( c \in C \),

\[
\psi(z) \leq p_D(z),
\]  

(12)

for almost all \( z \in [\underline{c}, \overline{c}] \) 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. 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 (12).

**Proof.** For any \( c \in C \) and for any \( v \in [\psi(c), \bar{\psi}] \), let \( D^\psi(c) \in \mathcal{D} \) be defined as in (4) with \( \bar{\phi}(c) \) replaced by \( \psi(c) \). Also, let \( \sigma^* : C \to \Delta(\mathcal{D}) \) be defined as (5) with \( \bar{\phi} \) replaced by \( \psi \).

By construction, \( \sigma^*(c) \in \mathcal{S} \) for all \( c \in C \). Furthermore, \( \sigma^* \) is a \( \psi \)-quasi-perfect scheme satisfying (12). To see this, for any \( c \in C \), since \( D_0 \) induces decreasing marginal revenue, the function \( p \mapsto (p - c)D_0(p) \) is single-peaked. Therefore, \((p - c)D_0(p) \leq (\psi(c) - c)D_0(\psi(c)) \) for all \( p \leq \psi(c) \). As a result, for any \( v \in [\psi(c), \bar{\psi}] \), since \( D_0(\psi(c)) = D^\psi(\psi(c)) \) and since \( D^\psi(\psi(c)) = p \) for all \( p \leq \psi(c) \), it must be that

\[
(p - c)D^\psi(c)(p) = (p - c)D_0(p) \leq (\psi(c) - c)D_0(\psi(c)) \leq (v - c)D_0(\psi(c)) = (v - c)D^\psi(c)(v),
\]

for all \( p \leq \psi(c) \). Therefore, since \( \text{supp}(D^\psi(c)) \cap [\psi(c), \bar{\psi}] = \{v\} \), it follows that \( p_{D^\psi(c)}(c) = v \) 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 \( p_{D^\psi(c)} \) is nonincreasing, it must be that either \( p_{D^\psi(c)}(z) = v \) or \( p_{D^\psi(c)}(z) < \psi(c) \). In the former case, since \( \psi \) is nondecreasing,
it then follows that \( p_{D^{\psi(c)}}(z) = v \geq \psi(c) \geq \psi(z) \), as desired. In the latter case, since \( D^{\psi(c)}(p) = D_0(p) \) for all \( p \leq \psi(c) \) and since \( p \mapsto (p-z)D_0(p) \) is singled-peaked, \( p_{D^{\psi(c)}}(z) \) must be the largest optimal price for the producer under \( D_0 \) as well. That is, \( p_{D^{\psi(c)}}(z) = p_0(z) \).

Combined with the hypothesis that \( \psi(z) \leq p_0(z) \), this then implies that \( \psi(z) \leq p_{D^{\psi(c)}}(z) \), as desired. As a result, \( \sigma^* \) is indeed a \( \psi \)-quasi-perfect scheme satisfying (12). ■

A direct consequence of Lemma 2 and Lemma 3 is that there exists an incentive compatible \( \phi \)-quasi-perfect mechanism \((\sigma, \tau)\), provided that (6) holds. Furthermore, for any \( c \in C \), (6) also implies that

\[
\int_c^{\bar{c}} D_0(\phi(z)) \, dz \geq \int_c^{\bar{c}} D_0(\phi_1(z)) \, dz.
\]

Together, by Lemma 1 and (7), after possibly adding a constant to \( \tau \) so that the indirect utility of the producer with cost \( \bar{c} \) equals to \( \bar{\pi} \), \((\sigma, \tau)\) is an incentive feasible \( \phi \)-quasi-perfect mechanism, which in turn implies that \((\sigma, \tau)\) is optimal. Combined with (11), it then follows that any incentive feasible \( \phi \)-quasi-perfect mechanism is optimal.

Combining Lemma 1, Lemma 2 and Lemma 3, it then follows that there exists an incentive feasible \( \phi \)-quasi-perfect mechanism and hence the data broker can attain revenue \( \bar{R} \), proving the first part of Theorem 1 (under (6)). In fact, even without (6), the proof above still implies the canonical \( \phi \)-quasi-perfect mechanism \((\sigma^*, \tau^*)\) defined by (5) 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 \( \phi \)-quasi-perfect, suppose that \((\sigma, \tau)\) is optimal. Then,

\[
\bar{R} = \int_C \left( \int \left( (v - \phi(c))D_0(dv) \right) G(dc) - \bar{\pi} \right) = \int_C \left( \int_D (p_D(c) - \phi(c))D(p_D(c))\sigma(dD|c) \right) G(dc) - \bar{\pi},
\]

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

\[
\int \{v \geq \phi(c)\} (v - \phi(c))D_0(dv) = \int_D (p_D(c) - \phi(c))D(p_D(c))\sigma(dD|c),
\]

since the left-hand side is the efficient surplus in an economy where the producer’s cost is \( \phi(c) \) and hence must be an upper-bound of the right-hand side. (13) then implies that the right-hand side of (14) must attain this upper bound for (almost) all \( c \in C \).

It then follows that \( \sigma \) must be a \( \phi \)-quasi-perfect mechanism. Indeed, if \( \sigma \) is not a \( \phi \)-quasi-perfect scheme, then there must exist (a positive measure of) \( c \in C \) and (a positive measure of) \( D \in \text{supp}(\sigma(c)) \) such that either \( D(v) > 0 \) for some \( v > p_D(c) \), or \( D(\phi(c)) \neq D(p_D(c)) \).
That is, either there are some consumers with \( v \geq \phi(c) \) who do not buy the product or buy the product at a price below \( v \), or there are some consumers with \( v < \phi(c) \) who end up buying the product. This contradicts (14). As a result, \((\sigma, \tau)\) must be a \( \phi \)-quasi-perfect mechanism. Moreover, \((\sigma, \tau)\) must also induce quasi-perfect price discrimination since \( p \) can be replaced with any selection of \( P \) according to Lemma 1.

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 to 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.

**Corollary 1 (Surplus Extraction).** *Consumer surplus is zero under any optimal mechanism.*

Corollary 1 follows directly from the characterization given by Theorem 1. According to Theorem 1, any optimal mechanism must induce \( \phi(c) \)-quasi-perfect price discrimination for (almost) all \( c \in C \), which means that every purchasing consumer must be paying their values. Notably, Corollary 1 provides an unambiguous assertion about the consumer surplus under data brokership. According to Corollary 1, even though the data broker does not sell the product to 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 Corollary 1 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 Lemma 4, allows for such a comparison.

**Lemma 4.** *The data broker’s optimal revenue is no less than the consumer surplus under uniform pricing.*
An immediate consequence of Lemma 4 is that total surplus under data brokership is greater compared with uniform pricing, as summarized below.

**Corollary 2 (Total Surplus Improvement).** *Data brokership always increases total (expected) surplus compared with uniform pricing.*

The reason behind Lemma 4 and Corollary 2 is that while all the purchasing consumers pay their values, more consumers are buying the product under data brokership compared with uniform pricing (similar to the reason why perfect price discrimination induces higher total surplus than uniform pricing does). 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 Lemma 4 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.

Specifically, consider the following augmented model: (i) the data broker makes a take-it-or-leave-it offer to all consumers; (ii) each consumer chooses whether to accept the offer simultaneously; (iii) Nature then draws values from $D_0$ independently for each consumer;\(^{20}\) (iv) the data broker gains access to complete information about the value of every consumer who accepts the offer, and has no information about the consumers who reject the offer; (v) the data broker then sells market segmentations among those who accept the offer to the producer. In this setting, when deciding whether to accept the data broker’s offer, if a consumer accepts, then from Corollary 1, their expected surplus would be zero. In contrast, for all consumers who reject the offer, the producer will not be able to price discriminate them and thus their expected surplus would be the same as that under uniform pricing.\(^{21}\)

---

\(^{20}\)It is crucial here the data broker purchases before the consumers learn their value, since otherwise he would also have to screen the consumers to elicit their private information. This corresponds to several business models in practice. For example, Nielsen often reaches out to potential consumers through mail and offer them to fill out surveys in exchange of monetary payments. This assumption is particularly 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. For instance, when creating an account for Google, Amazon, or Facebook, users have to agree to their terms of services before using the applications. Regardless of whether payments are monetary or non-monetary (e.g., compensate consumers through services), these type of transactions occur before consumers know their value for a particular product offered by a data buyer. See Bergemann, Bonatti, and Gan (2021) for more examples.

\(^{21}\)Since values are independently drawn, each consumer’s payoff will not be affected by other consumers’ decisions.
As a result, there is an equilibrium in which the intermediary offers all consumers their expected surplus under uniform pricing and every consumer accepts the offer, followed by the data broker selling market segmentations using the optimal mechanisms characterized by Theorem 1. In this equilibrium, since the data broker’s revenue is greater than the consumer surplus under uniform pricing according to Lemma 4, and since the producer always has an outside option of uniform pricing, the induced outcome is Pareto improving compared with uniform pricing in the ex-ante sense, as stated below.\footnote{Jones and Tonetti (2020) also conclude that granting consumers ownership of their own data is welfare-improving. However, their results are derived in a monopolistic competition setting and the main driving force is the non-rival property of data, whereas Proposition 1 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.}

**Proposition 1** (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.

**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 \(c, c' \in C\),

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

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

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

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 \(\gamma(\cdot|D, c) \in \Delta(\mathbb{R}_+)\) specifies the distribution from which prices charged in segment \(D\) must be drawn. It is noteworthy that price-controlling data brokership effectively allows the data broker to sell both consumer data and access to each individual consumers. After all, for any segment \(D\) and for any report \(c\), the data broker can set a price that is higher than all consumers’ values in segment \(D\). This would be equivalent to excluding all the consumers in segment \(D\) from the producer. As a result, price-controlling data brokership can also be interpreted as selling consumer data for the purpose of both price discrimination and targeting.\(^{23}\)

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

An immediate consequence of Theorem 1 is the comparison between data brokership and vertical integration. To see this, recall that any optimal mechanism \((\sigma, \tau)\) of the data broker must induce \(\overline{\phi}\)-quasi-perfect price discrimination but not perfect price discrimination in general, as \(\overline{\phi}(c) > c\) for all \(c > c\). Thus, 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 Corollary 1, this means that vertical integration between the data broker and producer strictly increases total profit.

\(^{23}\)I thank Barry Nalebuff and an anonymous referee for highlighting this interpretation.
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.

**Corollary 3** (Vertical Integration). Compared with data brokership, vertical integration strictly increases total surplus and leaves the consumer surplus unchanged if \( \bar{v} < \bar{c} \).

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 5** below, price-controlling data brokership is always equivalent to exclusive retail.

**Lemma 5.** Exclusive retail and price-controlling data brokership are outcome-equivalent.

With **Lemma 5**, to compare exclusive retail and price-controlling data brokership with data brokership, it suffices to compare only price-controlling data brokership with data brokership. This comparison is particularly convenient since the price-controlling data broker’s revenue maximization problem is a relaxation of the data broker’s. After all, with the extra ability to contract on prices, the constraints in the price-controlling data broker’s problem are clearly weaker. Nevertheless, as an implication of **Theorem 1** and **Proposition 2** below, it turns out that the data broker’s optimal revenue is in fact the same as the price-controlling data broker’s optimal revenue.

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

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

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 and access to consumers does not benefit the data broker at all. In fact, as stated by **Theorem 3** below, this ability is entirely irrelevant in terms of market outcomes.

**Theorem 3** (Outcome-Equivalence). Exclusive retail, price-controlling data brokership and data brokership are outcome-equivalent.
In other words, Theorem 3 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 3 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 brokership regime: According to Theorem 3, 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 (6) imposed in Section 4.2. To see this, since the price-controlling data broker’s optimal mechanisms always induce $\phi$-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 $\phi$-quasi-perfect mechanism. Meanwhile, by the definition of $\phi$, $c \leq \phi(c) \leq p_0(c)$ for all $c \in C$, and hence $\phi$ satisfies the condition required by Lemma 3. As a result, combining Lemma 2 and Lemma 3, there is indeed an incentive feasible $\phi$-quasi-perfect mechanism, which, by definition, generates revenue $R^*$, and hence is optimal.

5.4 Discussions and Policy Implications

The results above have several broader policy implications. First, in terms of welfare, although Corollary 1 implies that data brokership is undesirable for the consumers, Corollary 2 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 redistributational policy tools available. If the policymaker’s objective is to simply maximize total surplus, or if redistributational 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 redistributational policies are accessible, then the presence of a data broker can be
fairly unfavorable. However, Proposition 1 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,\(^{24}\) as well as a channel for the data broker to compensate the consumers, then not only the consumers can secure their surplus as if their data is not used for price discrimination (via compensation), but also the entire economy can benefit from data brokership, because less deadweight loss will be generated.

Furthermore, the discussions in Section 5.3 facilitate the evaluation of whether a certain market regime is desirable than another. According to Corollary 3, 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 3 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 allowing 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.

6 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.

Several topics remain to be explored by future studies. First, instead of having private

\(^{24}\)For instance, just as what is stipulated by the recent regulation of the European Union, General Data Protection Regulation (GDPR, Art. 7), consumers’ property right for their own data can be better protected by prohibiting all the processing of personal data unless the data subject has consented to the use.
marginal production cost, a model where the producer’s private information is about how consumer data can be used to predict their values is worth exploring. After all, while a one-dimensional private cost brings a great amount of tractability and allows for a closed form characterization, it is not the only main source of private information in reality. In many settings, the data broker and the producers could have very different knowledge about the correlation between consumers characteristics and their values. Second, while the model can be extended so that the data broker can only create a limited set of market segmentations (see Online Appendix), such an extension is restricted to a partitional environment. A natural direction is then to study a setting where the feasible market segmentation is restricted by an arbitrary Blackwell upper bound. Lastly, while both the data broker and the producer are assumed to be monopolists in this paper, it would be economically relevant to explore the consequences of consumer-data brokership under different market structures.

Furthermore, as a well-known feature of one-dimensional private types, the optimal mechanism in this setting can be quite specific compared to a model with multi-dimensional types.
Appendix

A Details for $D$

Below I first discuss more formally about the properties of the set $D$. Recall that $D$ is the collection of nonincreasing and left-continuous functions $D$ on $\mathbb{R}_+$ such that $D(v) = 1$ and $D(\bar{v}^+) = 0$. Since for every $D \in D$, 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 integral, and thus a topology on $D$ by the following notions: For any $D \in D$, and for any measurable function $f : V \to \mathbb{R}$, let $m^D$ be the unique probability associated with $D$ and let

$$
\int_V f(v)D(dv) := \int_V f(v)m^D(dv).
$$

Moreover, for any $\{D_n\} \subseteq D$ and any $D \in D$, say that $\{D_n\} \to D$ if and only if for any bounded continuous function $f : V \to \mathbb{R}$,

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

This corresponds to the weak-* topology on $\Delta(V)$ and hence this topology on $D$ is also called the weak-* topology. Using this topology, endow $D$ with the Borel $\sigma$-algebra so that $\Delta(D)$ denotes the collection of probability measures on $D$. Finally, for any $D \in D$, let $S_D$ denote the collection of $s \in \Delta(D)$ such that (3) holds with $D_0$ replaced by $D$ (so that $S_{D_0} = S$). 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 the main results, and therefore are relegated to the Online Appendix.

**Lemma 6.** Consider any nondecreasing function $\psi : C \to \mathbb{R}_+$ with $c \leq \psi(c)$ for all $c \in C$. Suppose that for any $c \in C$, $\sigma(c) \in S$ 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. $\sigma : C \to \Delta(D)$ is measurable.
3. $\int_D D(p_D(c))\sigma(dD|c) = D_0(\psi(c))$ for all $c \in C$.
4. $\int_D p_D(c)D(p_D(c))\sigma(dD|c) = \int_{\{v \geq \psi(c)\}} vD_0(dv)$ 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 7.** 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 7 follows directly from the standard envelope arguments and therefore is omitted. In addition to Lemma 7, 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 8 below.

**Lemma 8.** 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:

\[
\begin{align*}
\sup_{q \in Q} & \int_C \left( \int_0^{q(c)} (D_0^{-1}(q) - \phi(c)) \, dq \right) G(dc) - \bar{\pi} \\
\text{s.t. } & \bar{\pi} + \int_c^{\bar{\tau}} q(z) \, dz \geq \bar{\pi} + \int_c^{\bar{\tau}} D_0(p_0(z)) \, dz,
\end{align*}
\]

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

The proof of Lemma 8 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 7), the objective function of (16) 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 (16) are equivalent to the individual rationality constraints.

With Lemma 8, the price-controlling data broker’s revenue maximization problem can be solved explicitly.
Proof of Proposition 2. Let $R^*$ be the value of (16) and consider the dual problem of (16). By weak duality, it suffices to find a Borel measure $\mu^*$ on $C$ and a feasible $q^* \in \mathcal{Q}$ such that $q^*$ is a solution of
\[
\sup_{q \in \mathcal{Q}} \left[ \int_C \left( \int_0^{q(c)} \left( D_0^{-1}(q) - \phi(c) \right) dq \right) G(dc) - \pi + \int_C \left( \int_0^z \left( q(z) - D_0(p_0(z)) \right) dz \right) \mu^*(dc) \right]
\] (17)
and that
\[
\int_C \left( \int_0^z \left( q^*(z) - D_0(p_0(z)) \right) dz \right) \mu^*(dc) = 0.
\] (18)

To this end, define $M^*: C \to [0, 1]$ as the following:
\[
M^*(c) := \lim_{z \uparrow c} q(z)(\phi(z) - p_0(z))^+, \forall c \in C.
\] (19)

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 $c^* := \inf \{ c \in C : \phi(c) > p_0(c) \}$.

For any $q \in \mathcal{Q}$, by interchanging the order of integrals and then rearranging, (17) can be written as
\[
\sup_{q \in \mathcal{Q}} \left[ \int_C \left( \int_0^{q(c)} \left( D_0^{-1}(q) - \bar{\phi}(c) \right) dq \right) G(dc) - \bar{\pi} - \int_C M^*(c)D_0(p_0(c)) dc \right].
\] (20)

To solve (20), notice that
\[
\int_C \left( \int_0^{q(c)} \left( D_0^{-1}(q) - \bar{\phi}(c) \right) dq \right) G(dc) \leq \int_C \left( \int_0^{D_0(\bar{\phi}(c))} \left( D_0^{-1}(q) - \bar{\phi}(c) \right) dq \right), \forall q \in \mathcal{Q}.
\] (21)

Thus, as $\bar{\phi}$ is nondecreasing, $D_0 \circ \bar{\phi}$ is indeed a solution of (20) and hence a solution of (17).

Moreover, since $\bar{\phi} \leq p_0$, for all $c \in C$, $\int_c^z D_0(\bar{\phi}(z)) dz \geq \int_c^z D_0(p_0(z)) dz$. Therefore, $D_0 \circ \bar{\phi} \in \mathcal{Q}$ is feasible in the primal problem (16). Meanwhile, since $M^*(c) = 0$ for all $c \in [c^*, c^*]$ and since $\bar{\phi}(c) = p_0(c)$ for all $c \in (c^*, \bar{c}]$, the complementary slackness condition (18) is also satisfied. Together, $D_0 \circ \bar{\phi}$ is indeed a solution of (16). Finally, by definition of $D_0^{-1}$, it then follows that
\[
R^* = \int_C \left( \int_0^{D_0(\bar{\phi}(c))} \left( D_0^{-1}(q) - \phi(c) \right) dq \right) G(dc) - \bar{\pi} = \int_C \left( \int_{\{v \geq \bar{\phi}(c)\}} (v - \phi(c))D_0(dv) \right) G(dc).
\]

To see that any solution of the price-controlling data broker’s problem must induce $\bar{\phi}(c)$-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 $\mathbb{E}[\tau(c)] = R^*$ and that the indirect utility of the producer with marginal cost $\bar{c}$ is $\bar{\pi}$. Thus, by Lemma 8, it must be that
\[
\int_C \left( \int_{\mathbb{R}_+} \left( p - \phi(c) \right) D(p) \gamma(dp|D, c) \right) \sigma(dD|c) G(dc) = \int_C \left( \int_{\{v \geq \bar{\phi}(c)\}} (v - \phi(c))D_0(dv) \right) G(dc),
\] (22)
which is equivalent to
\[
\int_C \left( \int_{\mathbb{R}_+} \left( p - \bar{\phi}(c) \right) D(p) \gamma(dp|D, c) \sigma(dD|c) \right) G(dc) + \int_C (\bar{\phi}(c) - \phi(c))q^*_c(c) G(dc) = \int_C \left( \int_{\{v \geq \bar{\phi}(c)\}} (v - \bar{\phi}(c))D_0(dv) \right) G(dc) + \int_C (\bar{\phi}(c) - \phi(c))D_0(\bar{\phi}(c)) G(dc),
\] (23)
where \( q_\gamma^\ast(c) := \int_{D \times \mathbb{R}_+} D(p)\gamma(dp|D,c)\sigma(dD|c) \) for all \( c \in C \). Moreover, since for any \( c \in C \),
\[
\int_{D \times \mathbb{R}_+} (p - \overline{\phi}(c))D(p)\gamma(dp|D,c)\sigma(dD|c) \leq \int_D \max\{[(p - \overline{\phi}(c))D(p)]\sigma(dD|c) \leq \int_V (v - \overline{\phi}(c))^+ D_0(dv), \tag{24}
\]
it must be that
\[
\int_C (\overline{\phi}(c) - \phi(c))q_\gamma^\ast(c)G(d\gamma) \geq \int_C (\overline{\phi}(c) - \phi(c))D_0(\overline{\phi}(c))G(d\gamma). \tag{25}
\]
Furthermore, since \( \overline{\phi}(c) = p_0(c) \leq \phi(c) \) for all \( c \in (e^\ast, \overline{c}) \) and \( \overline{\phi}(c) = \phi(c) \), for all \( c \in [\underline{c}, e^\ast] \), by the definition of \( M^\ast \) given by (19), together with integration by parts, (25) is equivalent to
\[
\int_C \left( \int_\sigma (q_\gamma^\ast(z) - D_0(p_0(z))) \, dz \right) M^\ast(d\gamma) \leq 0 \tag{26}
\]
Lastly, since \( (\sigma, \tau, \gamma) \) is individually rational, for any \( c \in C \),
\[
\int_C (q_\gamma^\ast(z) - D_0(p_0(z))) \, dz \geq 0.
\]
Thus, as \( M^\ast \) is the CDF of a Borel measure, (26) must hold with equality, which in turn implies that (25) must hold with equality. Together with (23), (24) must hold with equality for \( G \)-almost all \( c \in C \). Therefore, \( (\sigma, \tau, \gamma) \) must induce \( \overline{\phi}(c) \)-quasi-perfect price discrimination for \( G \)-almost all \( c \in C \), as desired.

\[\square\]

**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 (2020a), \( \pi_D : C \rightarrow \mathbb{R}_+ \) is convex and continuous on \( C \) for any \( D \in \mathcal{D} \) with \( \pi'_D(c) = -D(\hat{p}_D(c)) \) for all selection \( \hat{p} \) of \( P \) and for almost all \( c \in C \). Moreover, since for any \( D \in \mathcal{D} \) and for any selection \( \hat{p} \) of \( P \), \( |\pi'_D(c)| = |D(\hat{p}_D(c))| \leq 1 \), for almost all \( c \in C \), the order of integral and differential can be interchanged. That is, for any \( c, c' \in C \),
\[
\frac{d}{dc} \int_D \pi_D(c)\sigma(dD|c') = \int_D \pi_D(c)\sigma(dD|c') = -\int_D D(\hat{p}_D(c))\sigma(dD|c'). \tag{27}
\]
As such, for any \( c', \) the function \( c \mapsto \int_D \pi_D(c)\sigma(dD|c') \) is convex and, by (27), has an almost-everywhere derivative - \( \int_D D(\hat{p}_D(c))\sigma(dD|c') \), for any selection \( \hat{p} \) of \( P \). 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), letting \( U(c) := u(c, c) \), we then have
\[
U(c) = U(\overline{c}) - \int_\overline{c} \frac{\partial}{\partial c} u(z, z) \, dz = U(\overline{c}) + \int_\overline{c} \left( \int_D D(\hat{p}_D(z))\sigma(dD|z) \right) \, dz. \tag{28}
\]
Assertion 1 then follows after rearranging.
Furthermore, for any mechanism \((\sigma, \tau)\) satisfying assertion 1 (and hence (28)) with any selection \(\tilde{\nu}\) of \(P\), we have

\[
U(c) - u(c, c') = (U(c) - U(c')) + \int_{\mathcal{D}} (\pi_D(c) - \pi_D(c')) \sigma(dD|c')
\]

\[
= \int_{c}^{c'} \left( \int_{\mathcal{D}} D(\tilde{\nu}_D(z)) \sigma(dD|z) - \int_{\mathcal{D}} D(\tilde{\nu}_D(z)) \sigma(dD|c') \right) dz
\]

\[
= \int_{c}^{c'} \left( \int_{\mathcal{D}} D(\tilde{\nu}_D(z))(\sigma(dD|z) - \sigma(dD|c')) \right) dz,
\]

where the second equality follows from the fundamental theorem of calculus and (27). Therefore, for any mechanism \((\sigma, \tau)\) satisfying assertion 1 with any selection \(\tilde{\nu}\) of \(P\), \(U(c) \geq u(c, c')\) for all \(c, c' \in C\) if and only if assertion 2 holds. This completes the proof. 

\section*{B.4 Proof of Lemma 2}

\emph{Proof of Lemma 2.} Given any nondecreasing function \(\psi : C \rightarrow \mathbb{R}_+\), and any \(\psi\)-quasi-perfect scheme \(\sigma : C \rightarrow \mathcal{S}\), suppose that for any \(c \in C\), \(\psi(z) \leq p_D(z)\), 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 \(c < c'\),

\[
\int_{c}^{c'} \left( \int_{\mathcal{D}} D(p_D(z)) \sigma(dD|z) - \sigma(dD|c') \right) dz = \int_{c}^{c'} \left( D_0(\psi(z)) - \int_{\mathcal{D}} D(p_D(z)) \sigma(dD|c') \right) dz
\]

\[
\geq \int_{c}^{c'} \left( D_0(\psi(z)) - \int_{\mathcal{D}} D(\psi(z)) \sigma(dD|c') \right) dz
\]

\[
= \int_{c}^{c'} (D_0(\psi(z)) - D_0(\psi(z))) dz
\]

\[
= 0,
\]

where the first equality follows from assertion 3 of Lemma 6, the inequality follows from the hypothesis, and the second equality follows from \(\sigma(z) \in \mathcal{S}\) for all \(z \in [c, c']\). Meanwhile, for any \(c, c' \in C\) with \(c > c'\),

\[
\int_{c'}^{c} \left( \int_{\mathcal{D}} D(p_D(z)) \sigma(dD|c) - \sigma(dD|z) \right) dz = \int_{c'}^{c} \left( \int_{\mathcal{D}} D(p_D(z)) \sigma(dD|c) - D_0(\psi(z)) \right) dz
\]

\[
= \int_{c'}^{c} (\min\{D_0(\psi(c), D_0(z)) - D_0(\psi(z))\}) dz
\]

\[
\geq 0,
\]

where the first equality again follows from assertion 3 of Lemma 6, and the second equality follows from the fact that \(c < c'\) and from the definition of quasi-perfect segmentations. Therefore, by Lemma 1, there exists a transfer \(\tau\) such that \((\sigma, \tau)\) is incentive compatible, as desired. 

\footnote{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 \(p_D(z) = p_D(c)\) and hence \(D(p_D(z)) = D_0(\psi(c)) = D_0(z)\); if \(D(c) > 0\) and \(\max(\text{supp}(D)) < z\), then \(D(p_D(z)) = 0\); while if \(D(c) = 0\) then \(D(z) = 0\).}
B.5 Proof of Theorem 1

Proof of Theorem 1. I first show that the data broker’s optimal revenue must be the same as the price-controlling data broker’s optimal revenue $R^*$. Since $R^*$ is an upper bound of the data broker’s revenue under any incentive feasible mechanism, it suffices to find an incentive feasible mechanism for the data broker that gives revenue $R^*$. To this end, notice that since $c \leq \bar{\phi}(c) \leq p_0(c)$ for all $c \in C$ and $\bar{\phi} : C \to \mathbb{R}_+$ is nondecreasing, by Lemma 3, there exists a $\bar{\phi}$-quasi-perfect scheme $\bar{\sigma} : C \to S$ that satisfies (12). Together with Lemma 2, there exists a transfer $\bar{\tau}$ such that $(\bar{\sigma}, \bar{\tau})$ is individually rational. Meanwhile, by possibly adding a constant to the transfer $\bar{\tau}$ so that the indirect utility of the producer with cost $\tau$, $U(\tau)$, equals to $\bar{\pi}$ under the mechanism $(\bar{\sigma}, \bar{\tau})$, it must be that, for any $c \in C$,

$$
\int_D \pi_D(c)\bar{\sigma}(dD|c) - \bar{\tau}(c) = U(\bar{\tau}) + \int_c^\pi \left( \int_D D(p_D(z))\bar{\sigma}(dD|z) \right) dz
$$

$$\begin{align*}
&= \int_c^\pi \bar{\tau}(c) + \int_c^\pi D_0(\bar{\phi}(z)) dz \\
&\geq \int_c^\pi \bar{\tau}(c) - \int_c^\pi D_0(p_0(z)) dz \\
&= \pi_D_0(c),
\end{align*}
$$

where the first equality follows from Lemma 1, the second equality follows from assertion 3 of Lemma 6, the inequality follows from $\bar{\phi} \leq p_0$ and the last equality follows from (7). As a result, $(\bar{\sigma}, \bar{\tau})$ is individually rational.

Furthermore, since $\bar{\sigma} : C \to S$ is a $\bar{\phi}$-quasi-perfect scheme, by assertion 3 and assertion 4 of Lemma 6, for any $c \in C$,

$$
\int_D (p_D(c) - \phi(c))D(p_D(c))\bar{\sigma}(dD|c) = \int_{\{v \geq \bar{\phi}(c)\}} (v - \phi(c))D_0(dv).
$$

and therefore, together with Lemma 1,

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

$$\begin{align*}
&= \int_C \left( \int_{\{v \geq \bar{\phi}(c)\}} (v - \phi(c))D_0(dv) \right) G(dc) - \bar{\pi} \\
&= R^*,
\end{align*}
$$

as desired.

Since the data broker’s optimal revenue is $R^*$ and since (29) holds for any $\bar{\phi}$-quasi-perfect scheme $\sigma$, by Lemma 1, any incentive feasible $\bar{\phi}$-quasi-perfect mechanism must give revenue $R^*$ and hence is optimal.

Conversely, to see why any optimal mechanism must be a $\bar{\phi}$-quasi-perfect mechanism, consider any optimal mechanism $(\sigma, \tau)$. As it is optimal and incentive compatible, by Lemma 1,

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

for any selection $\hat{p}$ of $P$. Meanwhile, since $(\sigma, \tau)$ is individually rational, by Lemma 1, it must be that

$$
\int_c^\pi \left( \int_D D(\hat{p}_D(z))\sigma(dD|z) \right) dz \geq \int_c^\pi D_0(p_0(z)) dz, \forall c \in C,
$$

(31)
for any selection \( \hat{p} \) of \( P \).

Now suppose that \((\sigma, \tau)\) is not a \( \bar{\phi} \)-quasi-perfect mechanism or it does not induce \( \bar{\phi}(c) \)-quasi-perfect price discrimination for a positive \( G \)-measure of \( c \), then there exists a selection \( \hat{p} \) of \( P \), a positive \( G \)-measure of \( c \) and a positive \( \sigma(c) \)-measure of \( D \in D \) such that either \( \hat{p}_D(c) < p_D(c) \), or \( D(c) > 0 \) and either \( \#\{v \in \text{supp}(D) : v \geq \bar{\phi}(c)\} \neq 1 \) 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

\[
\int_{\{v \geq \bar{\phi}(c)\}} (v - \bar{\phi}(c)) D(dv) \geq \int_{\{v \geq \hat{p}_D(c)\}} (v - \bar{\phi}(c)) D(dv) = (\hat{p}_D(c) - \bar{\phi}(c)) D(\hat{p}_D(c)) + \int_{\{v \geq \hat{p}_D(c)\}} (v - \hat{p}_D(c)) D(dv) \geq (\hat{p}_D(c) - \bar{\phi}(c)) D(\hat{p}_D(c)),
\]

with at least one inequality being strict. Therefore,

\[
\int_C \left( \int_D (\hat{p}_D(c) - \bar{\phi}(c)) D(\hat{p}_D(c)) \sigma(dD|c) \right) G(dc) < \int_C \left( \int_V (v - \bar{\phi}(c))^+ D_0(dv) \right) G(dc).
\]  
(32)

Meanwhile, since by (30),

\[
\int_C \left( \int_D (\hat{p}_D(c) - \bar{\phi}(c)) D(\hat{p}_D(c)) \sigma(dD|c) \right) G(dc) + \int_C (\bar{\phi}(c) - \phi(c)) \left( \int_D D(\hat{p}_D(c)) \sigma(dD|c) \right) G(dc)
\]

\[
= \int_D \left( \int_C (\hat{p}_D(c) - \phi(c)) D_0(dv) \right) G(dc)
\]

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

\[
= \int_C \left( \int_{\{v \geq \hat{\phi}(c)\}} (v - \hat{\phi}(c))^+ D_0(dv) \right) G(dc) + \int_C (\bar{\phi}(c) - \phi(c)) D_0(\bar{\phi}(c)) G(dc),
\]

(32) implies that

\[
\int_C (\bar{\phi}(c) - \phi(c)) \left( \int_D D(\hat{p}_D(c)) \sigma(dD|c) \right) G(dc) > \int_C (\bar{\phi}(c) - \phi(c)) D_0(\bar{\phi}(c)) G(dc).
\]

Furthermore, since \( \bar{\phi}(c) = \phi(c) \) for all \( c \in [c, c^*] \) and \( \bar{\phi}(c) = p_0(c) \) for all \( c \in (c^*, \bar{c}] \), it then follows that

\[
\int_C (\phi(c) - p_0(c)) \left( \int_D D(\hat{p}_D(c)) \sigma(dD|c) \right) G(dc) < \int_C (\phi(c) - p_0(c)) D_0(\hat{\phi}(c)) G(dc),
\]

Using integration by parts, this is equivalent to

\[
\int_{c^*}^\bar{c} \left( \int_c^{c^*} \left( \int_D D(\hat{p}_D(z)) \sigma(dD|z) \right) dz \right) M^*(dc) < \int_{c^*}^\bar{c} \left( \int_c^{c^*} D_0(p_0(z)) dz \right) M^*(dc),
\]

where \( M^* \) is defined in (19). However, by (31) and by the fact that \( M^* \) is a CDF of a Borel measure, which is due to Assumption 1,

\[
\int_{c^*}^\bar{c} \left( \int_c^{c^*} \left( \int_D D(\hat{p}_D(z)) \sigma(dD|z) \right) dz \right) M^*(dc) \geq \int_{c^*}^\bar{c} \left( \int_c^{c^*} D_0(p_0(z)) dz \right) M^*(dc),
\]
a contradiction. Therefore, \( \sigma \) must be a \( \bar{\phi} \)-quasi-perfect scheme and must induce \( \bar{\phi}(c) \)-quasi-perfect price discrimination for \( G \)-almost all \( c \in C \). Together with Lemma 1, and the fact that \( U(\bar{\sigma}) = \bar{\tau} \) under any optimal mechanism, \( (\sigma, \tau) \) must be a \( \bar{\phi} \)-quasi-perfect mechanism. This completes the proof. \( \blacksquare \)
B.6 Proof of Theorem 2

Proof of Theorem 2. By the proof of Lemma 3 in the main text. Since \( c \leq \bar{\phi}(c) \leq p_0(c) \) for all \( c \in C \), the canonical \( \bar{\phi} \)-quasi-perfect scheme \( \sigma^* \) defined in (5) is implementable. Therefore, there exists \( \tau^* \) such that \((\sigma^*, \tau^*)\) is an incentive feasible \( \bar{\phi} \)-quasi-perfect mechanism. By Theorem 1, \((\sigma^*, \tau^*)\) is optimal. \[ \blacksquare \]

C Omitted Proofs for Other Results

C.1 Proof of Corollary 1

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

\[
\int_C \left( \int_D \left( \int_{\{v \geq \hat{p}_D(c)\}} (v - \hat{p}_D(c)) D(dv) \right) \sigma^*(dD|c) \right) G(dc) = \int_C \left( \int_D \left( \int_{\{v \geq p_0(c)\}} (v - p_0(c)) D(dv) \right) \sigma^*(dD|c) \right) G(dc) = 0,
\]
as desired. \[ \blacksquare \]

C.2 Proof of Lemma 4

Proof of Lemma 4. Since \( P_{D_0}(c) \) is a singleton for (Lebesgue)-almost all \( c \in C \) and since \( G \) is absolutely continuous, consumer surplus under uniform pricing does not depend which selection of \( P \) is used. Therefore, by Theorem 1, the difference between the data broker’s optimal revenue and the consumer surplus under uniform pricing is

\[
\int_C \left( \int_D \left( \int_{\{v \geq \bar{\phi}(c)\}} (v - \bar{\phi}(c)) D_0(dv) \right) G(dc) - \bar{\pi} \right) - \int_C \left( \int_D \left( \int_{\{v \geq p_0(c)\}} (v - p_0(c)) D_0(dv) \right) G(dc) \right) = \int_C \left( \int_D \left( \int_{\{v \geq \bar{\phi}(c)\}} (v - \bar{\phi}(c)) D_0(dv) \right) G(dc) \right) - \bar{\pi} - \int_C \left( \int_D \left( \int_{\{v \geq p_0(c)\}} (v - p_0(c)) D_0(dv) \right) G(dc) \right)
\]

\[
= \int_C \left( \int_D (p_0(c) - \bar{\phi}(c)) D_0(p_0(c)) \right) G(dc) + \int_D \left( \int_{\{v \geq \bar{\phi}(c), p_0(c)\}} (v - \bar{\phi}(c)) D_0(dv) \right) G(dc) = \int_C \left( \int_D (p_0(c) - \bar{\phi}(c)) D_0(p_0(c)) \right) G(dc) + \int_D \left( \int_{\{v \geq \bar{\phi}(c), p_0(c)\}} (v - \bar{\phi}(c)) D_0(dv) \right) G(dc) 
\]

\[
\geq 0
\]

where the first equality follows from the fact that \( \bar{\phi}(c) < p_0(c) \) if and only if \( \phi(c) < p_0(c) \), and the third equality follows from changing the order of integrals. This completes the proof. \[ \blacksquare \]

C.3 Proof of Theorem 3

Proof of Theorem 3. By Lemma 5, whose proof can be found in the Online Appendix, it suffices to prove the outcome-equivalence between data brokership and price-controlling data brokership. By Proposition 2
and Theorem 1, both the data broker and the price-controlling data broker have optimal revenue $R^*$. Furthermore, for any optimal mechanism $(\sigma, \tau)$ 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 $\bar{\phi}(c)$-quasi-perfect price discrimination for $G$-almost all $c \in C$. In particular, for $G$-almost all $c \in C$, all the consumers with $v \geq \bar{\phi}(c)$ buy the product by paying their values and all the consumers with $v < \bar{\phi}(c)$ 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 $\bar{\phi}$-quasi-perfect scheme and hence by assertions 3 and 4 of Lemma 6, and by Lemma 1, for Lebesgue almost all $c \in C$,

$$\int_D \pi_D(c)\sigma(dD|c) - \tau(c) = \bar{\pi} + \int_c^\bar{\pi} \left( \int_D D(p_D(z))\sigma(dD|z) \right) dz = \bar{\pi} + \int_c^\pi D_0(\bar{\phi}(z)) dz. \quad (33)$$

Meanwhile, for the price-controlling data broker’s optimal mechanism $(\hat{\sigma}, \hat{\tau}, \hat{\gamma})$, since, by Lemma 4, it induces $\bar{\phi}(c)$-quasi-perfect price discrimination for almost all $c \in C$, it must be that $\hat{q}_\gamma^\hat{\sigma}(c) = D_0(\bar{\phi}(c))$. Together with Lemma 7, for any $c \in C$,

$$\int_D \left( \int_{\mathbb{R}_+} (p - c)\hat{D}(p)\hat{\gamma}(dp|D,c) \right) \hat{\sigma}(dD|c) - \hat{\tau}(c) = \bar{\pi} + \int_c^\pi \hat{q}_\gamma^\hat{\sigma}(z) dz = \bar{\pi} + \int_c^\pi D_0(\bar{\phi}(z)) dz. \quad (34)$$

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

**References**


