THE UNIVERSITY OF CHICAGO

INFORMATIONAL MONOPOLY: AN EQUIVALENCE RESULT

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY

KAI HAO YANG

CHICAGO, ILLINOIS

JUNE 2020

TABLE OF CONTENTS

LIST OF FIGURES iv

LIST OF TABLES v

ACKNOWLEDGMENTS vi

ABSTRACT x

1 INTRODUCTION 1
1.1 Preface 1
1.2 Notation and Model 3
1.2.1 Notation 3
1.2.2 Model 4

2 FULLY INFORMED BUYER: SALE OF CONSUMER DATA 10
2.1 Introduction 10
2.1.1 Preface 10
2.1.2 Motivation 12
2.1.3 Related Literature 15
2.2 Model 17
2.3 Optimal Mechanisms 20
2.3.1 Optimal Mechanism under Regime $\mathcal{P}$ 20
2.3.2 Optimal Mechanism under Regime $\mathcal{I}$ 22
2.3.3 Further Remarks 31
2.4 Implications 33
2.4.1 Outcome-Equivalence 33
2.4.2 Welfare Implications and Consumer Data Ownership 35
2.5 Extensions 37
2.5.1 Sufficient Conditions and Relaxations of Assumption 2.1 37
2.5.2 Consumers' Private Information 39
2.5.3 Targeted Marketing 42
2.6 Discussions 45
2.7 Conclusion 47

3 UNINFORMED SELLER: SALE OF ADVERTISEMENT 49
3.1 Introduction 49
3.1.1 Preface 49
3.1.2 Motivation 51
3.1.3 Related Literature 53
3.2 Model 55
3.3 Optimal Mechanisms and Outcome-Equivalence 56
3.3.1 Optimal Mechanisms under Regime $\mathcal{P}$ 56

3.3.2 Optimal Mechanisms under Regime $\mathcal{I}$ 60
3.3.3 Outcome-Equivalence 65
3.4 Irrelevance of Targeting 65
3.5 Welfare Analysis and Comparative Statics 68
3.6 Discussions 70
3.7 Conclusion 72

4 FLEXIBLE INFORMATION: ONLINE PLATFORM DESIGNS 74
4.1 Introduction 74
4.2 Model 76
4.3 Optimal Mechanisms and Equivalence Results 78
4.4 Extension: Different Private Types 82
4.5 Discussions 86
4.6 Conclusion 88

APPENDIX FOR CHAPTER 2 90

APPENDIX FOR CHAPTER 3 168

REFERENCES 184

LIST OF FIGURES

2.1 Market segment $D_v^{\psi(c)}$ 30

3.1 Feasible Set $\mathcal{D}_0$. 50
3.2 Pointwise Upper Bound 58
3.3 Profit under $p = \mathbb{E}_{m^0}[v|v \geq \phi_G(c)]$ and $\boldsymbol{D}_{\mathrm{b}}(c)$ 59
3.4 Upper Censorship $\boldsymbol{D}_{\mathrm{u}}(c)$ with Cutoff $\phi_G(c)$. 63

LIST OF TABLES

1.1 Restrictions on Information Structures 9

ACKNOWLEDGMENTS

It has been more than a decade since I started dreaming about writing the acknowledgment pages of my Ph.D thesis, which would mark not only the reception of the long-dreamed Ph.D degree but also a conclusion of a decade-long chapter in my life of pursuing one. Although defending my dissertation via Zoom and writing the acknowledgement pages during quarantine was not exactly how I pictured my journey would end, the emotional complexity that comes along with retrospecting to the challenging and yet fruitful path does not fall short of my expectation. From the winter of 2010 to the spring of 2020, the journey that I've been through for the past ten years may be one of the most unforgettable decade in my life. From a political-science-major sophomore who did not know Calculus, to this Ph.D dissertation in economic theory, this past decade would have been unimaginably different if there weren't for the supports and guidance of numerous kind and heartwarming helping hands along the way.

First of all, I would like to give thanks to all the professors who have taught me, guided me and inspired me. I am extremely grateful for my thesis advisor, Professor Philip Reny, although "being extremely grateful" is probably not enough for me to describe how indebted I am to Phil. From his fantastic teaching during the first and second year courses, to countless meetings with him on the whiteboard to present ideas, to the RA meetings, and eventually to the (endless) rounds of paper revision and slides preparation meetings for the job market, Phil has always been patient, polite, and yet extremely sharp. I learned not only about research and economics from him, but also the attitude of being a professional economist. His dedication to rigor and precision, as well as to being gentle to the audiences and readers, will always be my Polaris on my journey as an economist in the future—just as the pictures on slides that shouldn't move. Thank you, Phil, for not only being my advisor, but also my mentor as an economist.

If having a great thesis advisor is a Ph.D student's good fortune, I cannot believe how lucky I am to have a committee full of fantastic faculty members. I sincerely appreciate everything my committee members, Professor Benjamin Brooks, Professor Emir Kamenica and Professor Doron Ravid, have done for me during these years. Their brilliance and generosity made every meetings extremely valuable. Their guidance led me from a second-year student who knew nothing about and was intimidated by economic theory to being able to go on the job market for theory, and eventually the completion of my dissertation. Looking back at my dissertation, I can still vividly remember their voices when making suggestions and comments during my pretensions and meetings with them. Thank you, Ben, Emir and Doron, for all your time and effort. I would not have been able to be here if it weren't for you.

Another unbelievable blessing I have had throughout the years is to be at the University of Chicago. The theory faculty in this community and the environment are fantastic. Besides my committee members, I have received numerous valuable advises and comments from the faculty members here: Professor Alexander Frankel, Professor Jacob Leshno, Professor Elliot Lipnowski, Professor Daniel Rappoport, Professor Roger Myerson and Professor Nancy Stokey. I am truly grateful for their constant supports and generosity.

I would also like to express my gratefulness to two professors who had a significant impact on my academic career. Professor Min-jen Lin and Professor Jih-wen Lin. Thank you Min-jen, for opening my eyes to the world of economics, for your constant persuasion/encouragement, and for eventually recommending me to Chicago. Thank you Jih-wen, for always being my role model as a scholar and as a Christian, for the lessons that you taught on the philosophy of social science, and for your kindness and warm encouragement when I lost faith. I wish I could show you my dissertation in person, but I believe you would see this, and hope that I made you proud.

This journey would have been lackluster if it wasn't accompanied by the incredible friends and colleagues. To my first-year study group buddies: Dan, David, Eunki, James, Kai-wei, Anna-Maria, Yuyao and Ziho, I can still vividly remember those sleepless nights when we tried to squeeze out just one more first-order condition for the price theory problem set, or those weekends at the grad lounge when we tried to solve Professor Shimer's system of difference equation with eight unknowns. Those are undoubtedly painful and yet unforgettable memories that I think I will carry with me for the rest of my life. To the theory group candidates: Andrew, John and Wenji, it is truly a blessing for me to have gone through this journey with you. From the small-group meetings at Ben's office, to the full-blown weekly theory group meetings with those random number generators, listening to everyone's recent research progress and thoughts on papers was one of the most exciting experiences throughout the years. Thank you for creating such a fantastic work environment and for being the best peers that one can ever imagine. To the Taiwanese econ group: Chien-yu, Kai-wei, Kuan-ming, Sheng-hao and Yu-ting, looking back at those late-night trash-talks about how to do economics experiments if we had an island; whether the educators in Taiwan care; or how to find real numbers, you brought me laughter, joy, and demonstrate how friendship among economists would be. To my dear friends in the incoming class of 2014 cohort, especially Fang, Wenji and Yuehao, it has been fortunate to have you as friends. Every lunch break and coffee/water break throughout the years, as well as the "support group session" during the job market season, did not only ease my anxiety, but also bound us together via the unforgettable memories over the years. I wish you all the best for the coming years in your career.

Finally, to my family, everything would not be possible if it weren't for you. I am grateful for all the supports from my mother and grandmother. For thirty years, you have been caring and supporting me without any hesitation. There is nothing in this world I can imagine that would empower your perseverance. Thank you, for always being there for me. To my beloved wife, Shiyu, there is absolutely no way for me to describe how grateful I am. You've always stood by me; you've always believed in me; you comfort my anxiety; you fought through

this entire battle alongside with me, and you never complained, never gave up, and never became weary of it. Having you as my life partner is the most amazing thing that have ever happened to me and I would like to share this moment of accomplishing the Ph.D program with you. Thank you Shiyu, for everything.

ABSTRACT

In a monopolistic pricing setting where the buyer has quasi-linear preference and unit demand and the seller has private production cost, I introduce an *informational monopolist* who has the ability to design information structures between the buyer and the seller about the buyer's value, and then sell these information structures to the seller.

I consider three environments that differ in the set of feasible information structures for the informational monopolist: The one where the buyer is fully informed and the informational monopolist can design and sell any information about the buyer's value to the seller; The one where both the buyer and the seller are initially uninformed and the informational monopolist can provide any information to the buyer and then charge the seller for this service; as well as the one where the informational monopolist is able to design any information structure between the buyer and the seller, subject to a constraint that the buyer must be better-informed. These three environments correspond to three leading examples: The sale of consumer data by data brokers, the sale of advertisements by agencies, and the design of online trading platforms. In each environment, I further consider two market regimes: One in which the informational monopolist can contract with the seller on product prices and one in which the informational monopolist cannot contract on prices.

The main results consist of two parts. First, I completely characterize the revenue-maximizing mechanisms for the informational monopolist under each environment and each market regime. Using this characterization, I further show that in each of these environments, the market outcomes are *equivalent* under any optimal mechanisms and any market regime, which implies that the informational monopolist's ability to contract on price is irrelevant when he can design and sell information flexibly.

Keywords: Informational monopolist, revenue-maximization, mechanism design, information design, virtual cost, outcome-equivalence

Jel classification:D42, D61, D82, D83, L12, M37

CHAPTER 1
INTRODUCTION

1.1 Preface

For the past decade, information technology has been vastly and drastically improved. Improvement of the internet and various devices facilitate personal data collection and information transmission; advancement in computation power makes data analysis significantly more efficient; and novel technology such as artificial intelligence and virtual reality creates many more possible ways to convey and transmit information. These rapid developments in information technology have brought several novel economic activities, or have injected new aspects to the existing ones. For instance, advancement of the ability to collect, store and analyze personal data has given rise to several *data brokers* (e.g., Axciom), who collect and sell consumer data to producers to facilitate their pricing strategies. Meanwhile, market consultants can also learn more about the potential customers and provide consulting services in a more flexible way. Alternatively, advertisers can now exploit the improved information channel to inform the consumers about products they might be interested in through various ways. Furthermore, online platforms (e.g., Amazon) now have many ways to design their websites and the information structure among the consumers and producers who trade on their platforms.

Given their novelty and sizes, especially during an era where the internet has become an indispensable part of the economy, it is thus imperative to understand the economics of the aforementioned activities. Specifically, how should entities exploit their information technologies and extract surplus from the market? What is the best way to sell information? What are the implications on the economy? In the following chapters, I explore the design and sale of information by an *informational monopolist* under various environments that are motivated by the aforementioned examples. Specifically, I consider an information monopolist's revenue-maximization problem. This information monopolist can design information about the consumers' values in a monopolistic pricing setting and then sell this *information structure* to the (monopolistic) producer. In Chapter 2, I consider an environment where the consumers are fully informed about their values and the informational monopolist sells information about consumers' values to the producer. This corresponds to the economic activities of data brokers and market consultants. In Chapter 3, I consider an environment where the producer is always uninformed about the consumers' values while the informational monopolist can provide different product-information to the consumers and then sell this to the producer as a service. This corresponds to the activities of advertisers. Finally, in Chapter 4, I consider an environment where the informational monopolist can design arbitrary information structures among the consumers and the producer—subject to a constraint that the producer always have less information about the consumers' values—and then sell this information structure to the producer. This corresponds to the design of online platforms and how these platforms contract with the retailers.

The main results of the following three chapters center around one subject: *Outcome-equivalence*. Specifically, I consider two major market regimes for the informational monopolist. Under one of the regimes, the informational monopolist can only design and sell information as a third party, and does not have control over the product market at all; while under the other market regime, the informational monopolist can contract with the producer on prices charged to consumers, in addition to simply selling information structures. The first set of main results are characterizations of revenue-maximizing mechanisms for the informational monopolist under the three aforementioned different environments and two regimes. The characterizations themselves shed lights on how the informational monopolist should design information structures to extract surplus. Another crucial implication of these characterizations is that the market outcomes—informational monopolist's revenue, producer's net profit, consumer surplus and the allocation of the product—are all the same under any revenue-maximizing mechanism and any market regime. This implies that as long has the informational monopolist is able to design and sell information structures in a sufficiently rich way, whether or not he is active in the product market (i.e., whether or not he can contract on prices) does not matter. From a bird's-eye perspective, this is because there are many ways to design information and hence the informational monopolist can exploit this flexibility to discipline the producer's pricing behavior *as if* he has control over them.

Among these chapters, Chapter 2 is extracted from Yang (2020). In this paper, I focus on the aspect of the sale of consumer information and how that affects the economy. Chapter 3 is modified from Yang (2019b), where I study the design and sale of information structures by an informational intermediary who can provide information to the buyer in a monopolistic pricing setting. Chapter 4 combines the results developed in these two papers, and establish a broader equivalence result, with a leading example being online platform design and its consequences. For the rest of this chapter, I will first introduce the notation that will be used throughout all chapters and will introduce the general environment in details.

1.2 Notation and Model

1.2.1 Notation

Below, I introduce the notation that will be used throughout the following chapters. For any Polish space X, $\Delta(X)$ denotes the set of probability measures on X where X is endowed with the Borel σ-algebra and $\Delta(X)$ is endowed with the with weak-* topology. When $X \subseteq \mathbb{R}$, for any probability measure $\mu \in \Delta(X)$, the expected value of a measurable function $f : X \to \mathbb{R}$ is denoted as

$$\mathbb{E}_\mu[f(x)] := \int_X f(x)\mu(\mathrm{d}x).$$

Similarly, for any non-empty, measurable set $A \subseteq X$, $\mathbb{E}_\mu[f(x)|A] := \mathbb{E}_\mu[f(x)|\sigma(A)]$, where $\sigma(A)$ is the σ-algebra generated by A. As notational conventions, when $X = [a, b]$ is an interval with some $-\infty \leq a < b \leq \infty$, write

$$\int_b^a f(x)\mu(\mathrm{d}x) := -\int_a^b f(x)\mu(\mathrm{d}x).$$

Also, when $X = [a, b]$, $b' > b$ and $a' < a$, define

$$\mathbb{E}_\mu[x|x > b'] := b, \quad \mathbb{E}_\mu[x|x < a'] := a.$$

Furthermore, given any (finite) collection of polish spaces $\{X_i\}_{i=0}^n$, any (finite) collection of transition kernels $\{\kappa_i\}_{i=1}^n$ with $\kappa_i : X_{i-1} \to \Delta(X_i)$, and any $\mu \in \Delta(X_0)$, the probability measure of the joint distribution induced by $\{\kappa_i\}_{i=1}^n$ and μ is denoted by $\mathbb{P}_{\kappa_n,\ldots,\kappa_1,\mu}$.[1] Meanwhile, the expectation induced by the measure $\mathbb{P}_{\kappa_n,\ldots,\kappa_1,\mu}$ is denoted by $\mathbb{E}_{\kappa_n,\ldots,\kappa_1,\mu}$.

1.2.2 Model

There is one buyer ($\mathfrak{B}$, they) and one seller ($\mathfrak{S}$, she). The buyer has value $v \in V := [\underline{v}, \overline{v}]$ for a single good which the seller is able to produce at cost $c \in C := [\underline{c}, \overline{c}]$, where $0 \leq \underline{v} < \overline{v} < \infty$ and $0 \leq \underline{c} < \overline{c} < \infty$. The buyer has quasi-linear preference and the seller has full bargaining power and hence is able to design any selling mechanism to sell this product to the buyer in order to maximize profit. A selling mechanism consists of an arbitrary set of messages and mappings that map the buyer's reported messages to probability of trade and the amount of payment.

The seller's production cost is privately known, and is drawn from a common prior G, where G is a CDF that admits density $g > 0$. A complete description of an *information structure* in this environment pertains to what the buyer and the seller know about v.

1. That is, for any measurable sets $\{A_i\}_{i=0}^n$ with $A_i \subseteq X_i$ for all $i \in \{0, \ldots, n\}$,

$$\mathbb{P}_{\kappa_n,\ldots,\kappa_1\mu}(A_0 \times \ldots \times A_n) := \int_{A_0}\int_{A_1}\cdots\int_{A_{n-1}}\int_{A_n} \kappa_n(\mathrm{d}x_n|x_{n-1})\kappa_{n-1}(\mathrm{d}x_{n-1}|x_{n-2})\cdots\kappa_1(\mathrm{d}x_1|x_0)\mu(\mathrm{d}x_0).$$

Specifically, the value v is drawn from a common prior $m^0 \in \Delta(V)$. An *information structure* is a transition kernel $\chi : V \to \Delta(S_{\mathfrak{B}} \times S_{\mathfrak{S}})$ that specifies, for each realized value v, a joint distribution of *signals* that will be privately received by the buyer and the seller respectively, where $S_{\mathfrak{B}}$ and $S_{\mathfrak{S}}$ are Polish spaces that are "rich enough".[2] In other words, for both the seller and the buyer, conditional on receiving a signal realization ($s_{\mathfrak{B}} \in S_{\mathfrak{B}}$ or $s_{\mathfrak{S}} \in S_{\mathfrak{S}}$), posteriors for the value v and the counter-part's signal can be formed, upon which the decisions are based. To keep the seller's problem tractable (i.e., to avoid informed principle problems), I assume that the buyer always knows more about v than the seller does. Let X be the collection of all such information structures.[3]

Given an information structure $\chi \in \mathrm{X}$, due to quasi-linearity, it is without loss for the seller to restrict attention to posted price mechanisms.[4] That is, the seller chooses a price $p \in \mathbb{R}_+$ and then the buyer, upon seeing price p, decides whether to buy the product at price p. With this simplification, the buyer's and the seller's (ex-post) payoffs can be written as follows:

$$u_{\mathfrak{B}}(a, p, v) = a \cdot (v - p)$$
$$u_{\mathfrak{S}}(a, p, v|c) = a \cdot (p - c),$$

2. In principle, the signal spaces $S_{\mathfrak{B}}$ and $S_{\mathfrak{S}}$ can be arbitrarily rich. However, to be able to formally define the set of all possible information structures (i.e., to avoid the Russell-Zermelo paradox), I hold fix the signal spaces to be some rich enough sets. In this setting, as long as the cardinalities of both $S_{\mathfrak{B}}$ and $S_{\mathfrak{S}}$ are at least the same as that of $\mathbb{R}$, this restriction would be without loss of generality.

3. Specifically, $\chi \in \mathrm{X}$ if and only if there exists transitional kernels $\chi_{\mathfrak{B}}$ and $\chi_{\mathfrak{S}}$ such that for any $v \in V$,

$$\chi(\mathrm{d}s_{\mathfrak{B}}, \mathrm{d}s_{\mathfrak{S}}|v) = \chi_{\mathfrak{S}}(\mathrm{d}s_{\mathfrak{S}}|s_{\mathfrak{B}})\chi_{\mathfrak{B}}(\mathrm{d}s_{\mathfrak{B}}|v).$$

4. It is without loss of generality to restrict attention to posted price mechanisms even though the seller has private information about c when designing selling mechanisms. This is because, given any signal realization $(s_{\mathfrak{B}}, s_{\mathfrak{S}})$, and hence the induced posteriors, the environment features independent private values and quasi-linear payoffs, and both the seller's and the buyer's payoffs are monotone in their types. By Proposition 8 of Mylovanov and Tröger (2014), it is as if c is commonly known when the producer designs selling mechanisms. Therefore, according to Myerson (1981) and Maskin and Zeckhauser (1983), it is without loss to restrict attention to posted price mechanisms.

where $a \in \{0, 1\}$ is the buyer's purchasing decision. The timing of the events is as follows:

1. Nature draws $v \in V$ from m^0 and $c \in C$ from G.
2. $\mathfrak{S}$ privately observes c.
3. Nature draws $(s_{\mathfrak{B}}, s_{\mathfrak{S}}) \in S_{\mathfrak{B}} \times S_{\mathfrak{S}}$ from $\chi(\cdot|v)$.
4. $\mathfrak{B}$ and $\mathfrak{S}$ privately observe $s_{\mathfrak{B}}$ and $s_{\mathfrak{S}}$, respectively.
5. $\mathfrak{S}$ chooses a price $p \in \mathbb{R}_+$.
6. $\mathfrak{B}$ observes p, and then chooses $a \in \{0, 1\}$.

For the buyer, as they always know more about v than the seller does, the only payoff-relevant statistic of an information structure is the interim expected value $\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}]$. As a result, given any information structure $\chi \in \mathrm{X}$ and any signal realization $s_{\mathfrak{B}} \in S_{\mathfrak{B}}$, the buyer's purchasing decision is simple. That is, $\mathfrak{B}$ chooses $a = 1$ if $\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}] \geq p$ and chooses $a = 0$ if $\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}] < p$.[5] Taking this as given, with any signal realization $s_{\mathfrak{S}} \in S_{\mathfrak{S}}$, the seller's goal is then to choose p to maximize

$$(p - c)\mathbb{P}_{(\chi,m^0)}(\{\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}] \geq p\}|s_{\mathfrak{S}}).$$

Therefore, the seller's profit generated by information structure χ is given by

$$U_{\mathfrak{S}}(\chi, c) = \mathbb{E}_{(\chi,m^0)}\left[\sup_{p\geq 0}(p - c)\mathbb{P}_{(\chi,m^0)}(\{\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}] \geq p\}|s_{\mathfrak{S}})\right]$$

An *informational monopolist* (he) sells information structures to the seller in order to maximize revenue. That is, the information monopolist designs a selling mechanism, which consists of an arbitrary set of messages and mappings that map from the seller's reported

5. The tie is broken in favor of the seller as this is the only tie-breaking rule that can occur in equilibrium.

message to an information structure and an amount of payment. If the seller does not participate in the mechanism, she has outside option $\pi_0(c)$ when her cost is c.[6] By the revelation principle, especially for mechanism design problems with general outside options (c.f. Myerson (1979); Jullien (2000)), it is without loss to restrict attention to incentive compatible and individually rational direct mechanisms. That is, a mechanism for the informational monopolist is a tuple of functions $(\alpha, \boldsymbol{\chi}, \tau)$ with domain C that asks the seller to report her cost, so that for each report $c \in C$, $\alpha(c) \in \{0, 1\}$ specifies whether the seller participates in the mechanism, $\boldsymbol{\chi}(c) \in \mathrm{X}$ stands for the information structure that will be provided (upon participation) and $\tau(c) \in \mathbb{R}$ stands for the amount of payment the seller has to pay to the informational monopolist (upon participation). A mechanism $(\alpha, \boldsymbol{\chi}, \tau)$ is said to be *incentive compatible* if it induces the seller to report truthfully. That is, for any $c, c' \in C$,

$$\alpha(c)[U_{\mathfrak{S}}(\boldsymbol{\chi}(c), c) - \tau(c)] + (1 - \alpha(c))\pi_0(c) \geq \alpha(c')[U_{\mathfrak{S}}(\boldsymbol{\chi}(c'), c) - \tau(c')] + (1 - \alpha(c'))\pi_0(c). \tag{IC-I}$$

Moreover, a mechanism $(\alpha, \boldsymbol{\chi}, \tau)$ is *individually rational* if for all $c \in C$ such that $\alpha(c) = 1$,

$$U_{\mathfrak{S}}(\boldsymbol{\chi}(c), c) - \tau(c) \geq \pi_0(c). \tag{IR-I}$$

Henceforth, a mechanism is said to be *incentive feasible* if it is incentive compatible and individually rational.

To establish the main equivalence result, the market regime where the informational monopolist can contract on prices should also be considered. Henceforth, I refer the regime introduced above as regime $\mathcal{I}$, where the informational monopolist cannot contract on price; and refer the regime where the information monopolist can contract on price as regime $\mathcal{P}$. Under regime $\mathcal{P}$, in addition to selling information structures, the informational monopolist can also contract with the seller on what prices she should charge based on each of her signal

6. The function $\pi_0 : C \to \mathbb{R}$ will be further specified in later chapters.

realizations. More specifically, a mechanism under this regime becomes a tuple $(\alpha, \boldsymbol{\chi}, \tau, \boldsymbol{\gamma})$, where for each report $c \in C$, $\boldsymbol{\gamma}(c)$ is a transition kernel that maps from $S_{\mathfrak{S}}$ to $\Delta(\mathbb{R}_+)$ so that for any reported cost c and for any signal realization $s_{\mathfrak{S}} \in S_{\mathfrak{S}}$, $\boldsymbol{\gamma}(\cdot|s_{\mathfrak{S}}, c)$ is the distribution from which the seller must draw her price. As a result, a mechanism under regime $\mathcal{P}$ is incentive compatible if for any $c, c' \in C$,

$$\begin{aligned}&\alpha(c)[\mathbb{E}_{(\boldsymbol{\chi}(c),m^0,\boldsymbol{\gamma}(c))}\left[(p-c)\mathbb{P}_{(\boldsymbol{\chi}(c),m^0,\boldsymbol{\gamma}(c))}(\{\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}] \geq p\}|s_{\mathfrak{S}})\right] - \tau(c)] + (1-\alpha(c))\pi_0(c)\\ \geq &\alpha(c')[\mathbb{E}_{(\boldsymbol{\chi}(c'),m^0,\boldsymbol{\gamma}(c'))}\left[(p-c)\mathbb{P}_{(\boldsymbol{\chi}(c'),m^0,\boldsymbol{\gamma}(c'))}(\{\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}] \geq p\}|s_{\mathfrak{S}})\right] - \tau(c')] + (1-\alpha(c'))\pi_0(c),\end{aligned} \tag{IC-P}$$

and is individually rational if for all $c \in C$ such that $\alpha(c) = 1$,

$$\mathbb{E}_{(\boldsymbol{\chi}(c),m^0,\boldsymbol{\gamma}(c))}\left[(p-c)\mathbb{P}_{(\boldsymbol{\chi}(c),m^0,\boldsymbol{\gamma}(c))}(\{\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}] \geq p\}|s_{\mathfrak{S}})\right] - \tau(c) \geq \pi_0(c). \tag{IR-P}$$

The main goal for the following chapters is to solve the informational monopolist's revenue maximization problem (i.e., choosing an incentive feasible mechanism to maximize expected revenue) under both regime $\mathcal{I}$ and regime $\mathcal{P}$ and under various economically relevant restrictions on the set of feasible information structures X. In Chapter 2, I consider a subset $\mathrm{X}_{\mathrm{Ch2}}$ of X that contains all information structures under which the buyer is fully informed and characterize the informational monopolist's revenue-maximizing mechanisms when he is only allowed to provide information structures in $\mathrm{X}_{\mathrm{Ch2}}$. This corresponds to economic activities such as sales of consumer data by a data broker. In Chapter 3, I consider another subset $\mathrm{X}_{\mathrm{Ch3}}$ of X that contains all information structures where the seller is entirely uninformed. This describes activities such as sales of advertisement by advertising agents. Finally, in Chapter 4, I impose no restrictions on X and derive the revenue-maximizing mechanisms. This captures the activities of online platforms that both have information about the consumers and can manipulate the information about the products provided to the consumers (e.g., Amazon). These restricted sets are summarized by the table below. For the following

	Buyer	**Seller**
$\mathrm{X}_{\mathrm{Ch2}}$	Fully-informed	Flexible
$\mathrm{X}_{\mathrm{Ch3}}$	Flexible	Uninformed
X	Flexible, more-informed	Flexible, less-informed

Table 1.1: Restrictions on Information Structures

chapters, the cost distribution G is said to be *regular* if the induced virtual cost function $\phi_G(c) := c + G(c)/g(c)$ is increasing. Similarly, the value distribution m^0 is said to be *regular* if it is regular in the Myersonian sense (see Monteiro and Svaiter (2010) for formal definition for regularity under arbitrary value distributions).

It is noteworthy that, regardless of the restrictions of X, the revenue maximization problem under regime $\mathcal{P}$ is a relaxed problem of that under regime $\mathcal{I}$ and hence the optimal revenue for the informational monopolist must be weakly larger under regime $\mathcal{P}$. The main results in the following chapters, however, imply that the informational monopolist's revenue (in fact, the entire market outcome) would be exactly the same regardless of which regime he operates under.

More specifically, given any subset $\widehat{\mathrm{X}} \subseteq \mathrm{X}$, regime $\mathcal{I}$ and regime $\mathcal{P}$ are said to be *outcome-equivalent* under $\widehat{\mathrm{X}}$ if, when restricting the informational monopolist's feasible information structures to $\widehat{X}$, every optimal mechanism under either regime $\mathcal{I}$ or regime $\mathcal{P}$ induce the same revenue for the informational monopolist, the same profit for the producer, the same consumer surplus, and the same ex-post allocation of the product. In the following chapters, I will show that regimes $\mathcal{I}$ and $\mathcal{P}$ are outcome-equivalent under $\mathrm{X}_{\mathrm{Ch2}}$, $\mathrm{X}_{\mathrm{Ch3}}$ and X.

CHAPTER 2

FULLY INFORMED BUYER: SALE OF CONSUMER DATA

2.1 Introduction

2.1.1 Preface

This chapter contains the main constituents of Yang (2020). In this chapter, I restrict the set of feasible information structures to $\mathrm{X}_{\mathrm{Ch2}} \subseteq \mathrm{X}$, where the buyer is fully informed, while the seller can be arbitrarily informed about v.[1] Meanwhile, I specialize the seller's outside option to her optimal profit when there is no information. That is, $\pi_0(c) := \max_p (p - c) m^0([p, \overline{v}])$. Under these specifications, the informational monopolist can be thought of as selling information about the buyer's value to the seller. This corresponds to economic activities such as sales of consumer data by a data broker, or sales of pricing advises by a market consultant. For concreteness, I use the sale of consumer data as the leading example in this chapter. Throughout this chapter, the informational monopolist is referred as the *data broker*.

The set of feasible information structures, when restricted to $\mathrm{X}_{\mathrm{Ch2}}$, has a more tractable representation. Specifically, since the buyer is always fully informed, their purchasing decision is simple: Given any realized value v and any posted price p, $\mathfrak{B}$ chooses $a = 1$ if $v \geq p$ and $a = 0$ if $v < p$. This in turn simplifies the seller's problem given an information structure $\chi \in \mathrm{X}_{\mathrm{Ch2}}$: Given any signal realization $s_{\mathfrak{S}} \in S_{\mathfrak{S}}$ and any realized cost $c \in C$, $\mathfrak{S}$ chooses p to maximize

$$(p - c)\mathbb{P}_{(\chi, m^0)}(\{v \geq p\} | s_{\mathfrak{S}}).$$

1. Formally, since $S_{\mathfrak{B}}$ has cardinality larger than $\mathbb{R} \supseteq V$, there is an injective function f that maps from V to $S_{\mathfrak{B}}$. As such, $\mathrm{X}_{\mathrm{Ch2}}$ can be defined as:

$$\mathrm{X}_{\mathrm{Ch2}} := \{\chi \in \mathrm{X} : \chi(\{f(v)\} \times S_{\mathfrak{S}} | v) = 1, \forall v \in V\}.$$

As a result, using Blackwell's characterization (Blackwell, 1953), the set of feasible information structures $\mathrm{X}_{\mathrm{Ch2}}$ can be represented by the set of distribution over posteriors over v that are mean-preserving spreads of a dirac measure on m^0. For the ease of notation, let $D_0(p) := m^0([p, \overline{v}])$ for all p and let $\mathcal{D}$ be the collection of nonincreasing, upper-semicontinuous functions $D : \mathbb{R}_+ \to [0,1]$ such that $D(\underline{v}) = 1$ and $D(\overline{v}^+) = 0$.[2] $\mathrm{X}_{\mathrm{Ch2}}$ can be represented by the collection of distributions $s \in \Delta(\mathcal{D})$ such that

$$\int_{\mathcal{D}} D(p)s(\mathrm{d}D) = D_0(p), \forall p \geq 0. \tag{2.1}$$

This representation has a natural interpretation. To see this, consider an equivalent model where there is one producer who sells a product at a (private) marginal cost c and there is a unit mass of consumers with unit demand of the product and heterogeneous values. Each consumer knows their own value and the values across the consumers are distributed according to $m^0 \in \Delta(V)$, so that $D_0(p)$ would be the share of consumers who buy the product if the price is p (i.e., D_0 is the *market demand*). According to (2.1), an information structure in X_{CH2} is equivalent to a *market segmentation* that splits the market demand into several *market segments* $D \in \mathcal{D}$. This captures the economic feature that data brokers often use consumer data to segment the consumers and sell these segmentations to the producer to facilitate price discrimination. In the remanding parts of the chapter, I will use this interpretation instead of the languages developed in Chapter 1. However, it should be noted that these two interpretations are entirely equivalent once the set of feasible information structures is restricted to $\mathrm{X}_{\mathrm{Ch2}}$.

2. Notice that $\mathcal{D}$ is isomorphic to the collection of probability measures on V, which can be interpreted as *demand functions* of the buyer, as the buyer buys the product if and only if $v \geq p$. Henceforth, $\mathcal{D}$ is endowed with the weak-* topology and the induced Borel σ-algebra. In addition, for any $D \in \mathcal{D}$, the measure associated with D will be denoted by m^D.

2.1.2 Motivation

The rapid development of informational technology has enabled one to collect, process and analyze vast volumes of consumer data. By the use of consumer data, the scope of price discrimination has moved 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 the prediction of their values, and therefore to create numerous sorts of *market segmentation*—a way to split the market demand into several segments by partitioning the consumers' characteristics—to facilitate price discrimination. Moreover, because of their specialization in information technology, several "data brokers" trade vast amounts of consumer data with retailers, which effectively means these data brokers can create market segmentations and sell them as a product that facilitates price discrimination. For instance, online platforms such as Facebook collect and sell[3] a significant amount of consumers' personal information, including personal characteristics, traveling plans, lifestyles, and text messages via its own platform. Alternatively, data companies such as Acxiom and Datalogix gather and sell personal information such as government records, financial activities, online activities and medical records to retailers (Federal Trade Commission, 2014).

This chapter studies the design of optimal selling mechanisms of a data broker. In this chapter, I specialize the model introduced in Chapter 1 into a model where there is one producer with privately known constant marginal cost, who produces and sells a single product to a unit mass of consumers. The consumers have unit demand and the distribution of their values is described by commonly known market demand. The informational monopolist is thought of as a *data broker*, who does not know the producer's marginal cost of production

3. 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

but can sell *any* market segmentation to the producer via *any* selling mechanism. As the data broker is uncertain about the production cost, and only affects the product market indirectly by selling consumer data to the producer, it is not obvious how the data broker should sell market segmentations to the producer, what market segmentations will be created, and how the sale of consumer data affects economic welfare and allocative outcomes.

The main result of this chapter is a complete characterization of the revenue-maximizing mechanisms for the data broker (under both regime $\mathcal{I}$ and $\mathcal{P}$). The optimal mechanisms feature *quasi-perfect price discrimination*, which is an outcome where all the purchasing consumers pay exactly their values, although not every consumer with values above the marginal cost buys the product. With the ability to contract on price, the revenue-maximization problem is essentially a non-linear pricing problem (with type-dependent outside option, see Jullien (2000)) as in Mussa and Rosen (1978), and the characterization of optimal mechanisms is given by Theorem 2.1. The main contribution of this chapter, both technically and economically, is that even without the ability to contract on prices, the data broker can still achieve the same revenue and even the entire market outcome would be the same. Theorem 2.2 completely characterizes the optimal mechanisms for the data broker under regime $\mathcal{I}$, and shows that every optimal mechanism must create *quasi-perfect segmentations* described by a cost-dependent cutoff. Namely, all the consumers with values above the cutoff are separated from each other whereas the consumers with values below the cutoff are pooled with the separated high-value consumers. When pricing optimally under this segmentation, the producer only sells to high-value consumers and induces quasi-perfect price discrimination. As a result, the ability to contract on prices does not matter, as the market segmentations can be designed to incentivize the producer to price in the way as if it is contracted. Furthermore, the cutoff function under any optimal mechanism has a closed form, which is exactly the minimum between the (ironed) *virtual marginal cost* function and the optimal uniform price as a function of marginal cost.

The outcome-equivalence between regime $\mathcal{I}$ and regime $\mathcal{P}$ can be interpreted as the equivalence in different two business models that can be employed by a data broker: Regime $\mathcal{I}$ corresponds to a business model where the data broker only sells consumer data and does not participate in the product market at all, whereas regime $\mathcal{P}$ can be interpreted as a business model where the data broker becomes an (exclusive) retailer, who buys products from the producer and then sells these products to the consumers as a monopoly via perfect-price discrimination. As a result, the outcome-equivalence between regime $\mathcal{I}$ and regime $\mathcal{P}$ indicates that it does not matter which business model is employed by the data broker, both from the broker's perspective and from a policymaker's perspective.

In addition to the equivalence result, the characterization of the optimal mechanisms further leads to several welfare implications. As the defining feature of quasi-perfect price discrimination, under any optimal mechanism, all the consumers pay their values conditional on buying. This implies that the consumer surplus under any optimal mechanism is zero (Proposition 2.1). In other words, in terms of consumer surplus, it is *as if* all the information about the consumers' values is revealed to the producer (even though it is not fully revealed under regime $\mathcal{I}$). Furthermore, Theorem 2.2 also allows a comparison between data brokership and uniform pricing, where no consumer data can be shared. More specifically, I show that data brokership always increases total surplus (Proposition 2.3), 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 Proposition 2.4).

The rest of this chapter is organized as follows. Continuing in this section, several related literatures are discussed. Section 2.2 presents the (specialized) model in further details. In Section 2.3, I characterize the optimal mechanisms of the data broker. The equivalence result and other consequences of consumer-data brokership are discussed in Section 2.4. Finally, some further discussions are in Section 2.6.

2.1.3 Related Literature

This chapter is related to several streams of literature. In the recent literature of price discrimination, several theoretical works center around the discussion of the welfare effects of price discrimination (see, for instance, Varian (1985), Aguirre et al. (2010) and Cowan (2016)) and provide conditions under which third-degree price discrimination increases or decreases total surplus and output. In addition, Bergemann et al. (2015) show that any surplus division between the consumers and the monopolist can be achieved by some market segmentation. In these papers, market segmentation is treated as an exogenous object, whereas in this chapter, market segmentation is determined endogenously by a data broker's revenue-maximization behavior. The welfare implications in this chapter are clear: With a data broker, consumer surplus always decreases, while volume of trade and total surplus always increases. On the other hand, several empirical works have focused on quantifying the improvements on welfare and profit by the use of price discrimination (see, for instance Shiller and Waldfogel (2011), Shiller (2014) and Dubé and Misra (2017)), as well as detecting and identifying the use of price discrimination (Mikians et al. (2012) and Goldberg (1996)).

Additionally, this chapter is related to the literature on the interplay between monopolistic pricing and information structures. Specifically, Lewis and Sappington (1991) and Johnson and Myatt (2006) characterize the seller-optimal information that the consumers have within a parameterized family of information structures. They show that either full-information or no-information is optimal for the seller. On the other hand, Roesler and Szentes (2017) characterize the buy-optimal information structure and further provide a set-valued prediction of all possible surplus division that can arise under some information structure that the buyer has. Yang (2019a) further characterizes the buyer-optimal information structure when the seller has cost uncertainty, which can be further extended to a characterization of all possible surplus division in an environment that features second-degree price discrimination (Yang, 2019c). In addition, Ravid et al. (2019) study a strategic environment where the buyer can acquire information covertly with some cost of learning. While these papers focus on the implications of different information the consumers can have, this chapter can be interpreted as a model that focuses on different information the producer can have.

Furthermore, this chapter can also be categorized into the mechanism design and information design literature. In particular, this chapter is closely related to the mechanism design problems where the information structures are also part of the design objects. Within this branch, Bergemann and Pesendorfer (2007) study an optimal auction problem in which the seller can also disclose information about the value to the consumers independently. Dworczak (2020) studies a mechanism design problem where the designer can also disclose information to affect the aftermarket. Kolotilin et al. (2017) examine a persuasion problem in which the sender can use different information structures to screen the receiver who has (independent) private information. Bergemann and Bonatti (2015) explore a pricing problem of a data provider who can use information about consumers' match value to facilitate a firm's targeting. Bergemann et al. (2018) solve a mechanism design problem in which the designer sells experiments to decision makers who have private information about his belief.

Methodologically, the main feature of this chapter is to exploit the richness of the design object (segmentation/information) to discipline double deviations. This is related to the dynamic mechanism design problems in Esö and Szentes (2007), where they show that in a two-period optimal auction problem with independent signals, the second-period incentive constraints are irrelevant and the designer's optimal revenue is as if the second-period signal were public.[4]

Among the aforementioned studies, Bergemann et al. (2015), Bergemann et al. (2018) are the closest to this chapter. Specifically, Bergemann et al. (2015) explore the welfare

4. The original statement of Esö and Szentes (2007) does not have the independence condition. However, as shown in Li and Shi (2017), such independence condition is needed in order for the optimal disclosure result to be valid.

implications of different market segmentation, while this chapter introduces a data broker who designs market segmentation in order to maximize revenue. Meanwhile, Bergemann et al. (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, while this chapter studies a revenue maximization problem where the agent's private information is part of her intrinsic preference and has a rich action space. Also, this chapter allows a large class of priors, including those with infinite support.

2.2 Model

In what follows, I specialize the general model introduced in Chapter 1 by restricting the set of feasible information structures to $\mathrm{X}_{\mathrm{Ch2}}$ and specializing the seller's outside option as

$$\pi_0(c) := \max_{p \geq 0}(p - c)D_0(p).$$

As discussed in Section 2.1, this model is equivalent to a model where there is a unit mass of consumers with unit demand and heterogeneous values, as well as one producer who sells a product at a constant marginal cost that is her private information. An information structure is equivalent to a market segmentation $s \in \Delta(\mathcal{D})$ satisfying (2.1) that splits the market demand D_0 into several market segments so that the producer can price-discriminate the consumers accordingly. In other words, given any market segmentation s, the producer with marginal cost c solves

$$\max_{p \geq 0}(p - c)D(p)$$

for all $D \in \mathrm{supp}(s)$. For the rest of this chapter, the set of market segmentations is denoted by $\mathcal{S}$. In addition, for any $c \in C$ and any $D \in \mathcal{D}$, let $\boldsymbol{P}_D(c)$ denote the set of optimal prices for the producer with marginal cost c under market segment D. As a convention, regard $\boldsymbol{P}$

as a correspondence on $\mathcal{D} \times C$ and if $\boldsymbol{p}$ is a selection for $\boldsymbol{P}$, write $\boldsymbol{p} \in \boldsymbol{P}$.[5] Furthermore, for any $c \in C$ and any $D \in \mathcal{D}$, let

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

denote the maximized profit of the producer. Also, let

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

be the largest optimal price for the producer with marginal cost c under market segment D.[6] For conciseness, let $\overline{\boldsymbol{p}}_0 := \overline{\boldsymbol{p}}_{D_0}$.

Throughout this chapter, I impose the following technical assumption on the market demand D_0 and the distribution G.

Assumption 2.1. *The function* $c \mapsto \max\{g(c)(\phi_G(c) - \overline{\boldsymbol{p}}_0(c)), 0\}$ *is nondecreasing.*

Assumption 2.1 permits a wide class of (D_0, G) and includes many common examples.[7] Also, it does not require regularities of either $m^0 0$ or G (nor is it implied by regularities of m^0 and G). In Section 2.5, I will further discuss this assumption, including how the results rely on it, its relaxations, as well as several economically interpretable sufficient conditions.

With this specification, a mechanism for the data broker (i.e., the informational monopolist introduced in Chapter 1) becomes a pair (σ, τ) under regime $\mathcal{I}$, and a tuple $(\sigma, \tau, \boldsymbol{\gamma})$

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

6. $\overline{\boldsymbol{p}}$ is well-defined by Lemma 2.5 in Appendix B. Moreover, according to the notational convention stated in footnote 5, whenever $c > \max(\text{supp}(D))$, $\overline{\boldsymbol{p}}_D(c) = \max \overline{V}$.

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

under regime $\mathcal{P}$,[8] where $\sigma : C \to \mathcal{S}$ is a measurable function that assigns a market segmentation $\sigma(c) \in \mathcal{S}$ given reported cost c and will be referred as a *segmentation scheme* (or sometimes, a *scheme*); $\tau : C \to \mathbb{R}$ is a measurable function that maps report c to the amount of payment $\tau(c)$; and $\boldsymbol{\gamma}$ is a measurable function so that for any report c, $\boldsymbol{\gamma}(c)$ is a transition kernel that maps from $\mathcal{D}$ to $\mathbb{R}_+$ and stands for the (stochastic) prices that must be charged to segment of consumers when the report is c.

Under regime $\mathcal{I}$, the incentive compatibility constraint (IC-I) and the individual rationality constraint (IR-I) can be simplified to

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

for all $c, c' \in C$, and

$$\int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c) - \tau(c) \geq \pi_0(c),$$

for all $c \in C$, respectively. Similarly, under regime $\mathcal{P}$, (IC-P) reduces to

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

for all $c, c' \in C$ while (IR-P) reduces to

$$\int_{\mathcal{D}\times\mathbb{R}_+} (p-c)D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\sigma(\mathrm{d}D|c) - \tau(c) \geq \pi_0(c),$$

for all $c \in C$.

From this regard, 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

8. Notice that a mechanism in this chapter does not need to specify exclusion decisions (i.e., the mapping $\alpha : C \to \{0, 1\}$ is not needed). This is because the seller's outside option is her optimal uniform pricing profit, which can be achieved by providing no information without excluding the seller.

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. Second, the producer's outside option is type-dependent. This is because the producer has access to the consumers, and is only buying the additional information about the consumers' values.

2.3 Optimal Mechanisms

In what follows, I characterize the data broker's optimal mechanisms under both regime $\mathcal{I}$ and $\mathcal{P}$. To begin with, I first examine the revenue maximization problem under regime $\mathcal{P}$ and establish an upper bound for the broker's revenue under regime $\mathcal{I}$.

2.3.1 Optimal Mechanism under Regime $\mathcal{P}$

With the ability to contract on price, the data broker's problem becomes straightforward despite the rich allocation space and the type-dependent outside option. Indeed, with the ability to contract price, as formally proved in Appendix B, the data broker can be thought of as a retailer, who buys the product from the producer as a monopsony and then sell all the purchased units to the consumers via perfect price discrimination. As a result, the data broker's problem is equivalent to choosing a purchasing contract $(\boldsymbol{q}, t)$, with $\boldsymbol{q} : C \to [0, 1]$ and $t : C \to \mathbb{R}$ being measurable and satisfying

$$t(c) - c\boldsymbol{q}(c) \geq \max\{t(c') - c\boldsymbol{q}(c'), \pi_0(c)\}$$

to maximize

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

where D_0^{-1} is the inverse market demand.[9] The interpretation is that the data broker, as a monopsony, uses an incentive compatible and individually rational direct mechanism $(\boldsymbol{q}, t)$ to purchase the product from the producer, so that given report c, he purchases $\boldsymbol{q}(c)$ units and pays $t(c)$ to the producer. Then he sells these $\boldsymbol{q}(c)$ units to the consumers via perfect price discrimination. As a result, the data broker's problem is reduced to a standard nonlinear pricing problem with one-dimensional allocation space and type-dependent outside option. Using the techniques developed by Myerson (1981) and Jullien (2000), optimal mechanisms for the data broker can be fully characterized. To state this result, let φ_G be the ironed virtual value and let $\overline{\varphi}_G(c) := \min\{\varphi_G(c), \boldsymbol{p}_0(c)\}$ for all $c \in C$. The characterization of optimal mechanisms can be stated as Theorem 2.1 below.

Theorem 2.1. *Under regime $\mathcal{P}$, a revenue-maximizing mechanism exists. Furthermore, under any revenue-maximizing mechanism, for G-almost all $c \in C$, all consumers with $v \geq \overline{\varphi}_G$ buy the product by paying their values while all consumers with $v < \overline{\varphi}_G(c)$ do not buy. In particular, the optimal revenue is*

$$R^* = \int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \pi_0(\overline{c}).$$

The proof of Theorem 2.1 can be found in Appendix B. The intuition behind Theorem 2.1 is simple: Due to the asymmetric information, the data broker's marginal cost is effectively larger than the producer's actual marginal cost by the amount of $G(c)/g(c)$ as he has to pay information rents to the producer. As a result, given each report $c \in C$ the data broker is essentially choosing a quantity $\boldsymbol{q}(c) \in [0, 1]$ to maximize profit. As the data broker sells all $\boldsymbol{q}(c)$ units via perfect price discrimination, had there been no type-dependent outside options, the optimal quantity is to sell to all the consumers with value above the (ironed)

9. More precisely,

$$D_0^{-1}(q) := \sup\{p \in V : D(p) \geq q\}, \forall q \in [0, 1].$$

virtual cost. Since the allocation space is one-dimensional, monotonicity of the optimal quantity as a function of c would be sufficient for global incentive compatibility. However, since the producer's outside option is type-dependent, this cutoff has to be adjusted so that the individual rationality constraints could be satisfied. Under Assumption 2.1, the adjustment is exactly to replace $\varphi_G(c)$ by $\overline{\varphi}_G(c)$, as argued in the formal proof.

Under any optimal mechanism, according to Theorem 2.1, all consumers with value $v \geq \overline{\varphi}_G(c)$ end up buying the product by paying their values, whereas all consumers with $v < \overline{\varphi}_G(c)$ do not buy. This outcome will be referred as *quasi-perfect price discrimination*. It is "perfect" as all the purchasing consumers pay their values, while it is "quasi" because not all the consumers with $v \geq c$ buy. More generally, for any $\kappa \geq 0$, an outcome is said to be *κ-quasi-perfect price discrimination for c* if all consumers with $v \geq \kappa$ buy the products by paying their values while all consumers with $v < \kappa$ do not buy. This outcome will be crucial for deriving the results in the following subsection.

2.3.2 Optimal Mechanism under Regime $\mathcal{I}$

In this subsection, I characterize the optimal mechanisms for the data broker under regime $\mathcal{I}$. Without the ability to contract on price, the rich allocation space creates substantial difficulties and thus standard methods for finding optimal mechanisms cannot be applied. Instead, in what follows, I exploit the characterization given by Theorem 2.1 and the observation that the data broker's revenue must be weakly higher under regime $\mathcal{P}$ than under regime $\mathcal{I}$ to characterize the optimal mechanisms under regime $\mathcal{I}$. In brief, in this subsection, I will construct an incentive feasible mechanism under regime $\mathcal{I}$ that achieves the same outcome as the optimal mechanisms under regime $\mathcal{P}$.

To begin with, recall that by Theorem 2.1, any optimal mechanisms under regime $\mathcal{P}$ must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for (almost) all $c \in C$. As a result, even without the ability to contract on prices, if there is an incentive feasible mechanism under regime $\mathcal{I}$ that induces the same outcome, then this mechanism must be optimal and the data broker's optimal revenue must be R^* under regime $\mathcal{I}$ as well. However, without the ability to contract on prices, the data broker cannot simply use the value-revealing segmentations and induce perfect price discrimination to achieve this outcome, as there would be consumers with values $v \in (c, \overline{\varphi}_G(c))$ to whom the data broker would not like to sell but the producer would. Instead, more sophisticated market segmentations have to be used. This leads to the following definition.

Definition 2.1. For any $c \in C$ and any $\kappa \geq c$, a segmentation $s \in \mathcal{S}$ is a *κ-quasi-perfect segmentation* for c if for s-almost all $D \in \mathcal{D}$, either $D(c) = 0$, or the set $\{v \in \text{supp}(D) : v \geq \kappa\}$ is a singleton and is a subset of $\boldsymbol{P}_D(c)$.

A κ-quasi-perfect segmentation for c is a segmentation that separates all the consumers with $v \geq \kappa$ while pooling the rest of the consumers together with them so that when the producer's marginal cost is c, every market segment with positive trading volume[10] must contain one and only one consumer-value $v \geq \kappa$ and this v is an optimal price for the producer. Notice that a κ-quasi-perfect segmentation for c induces κ-quasi-perfect price discrimination for c when the producer's marginal cost is c and when she charges the largest optimal price in (almost) all segments. Namely, a consumer with value v buys the product if and only if $v \geq \kappa$ and all purchasing consumers pay exactly their values.

With Definition 2.1, I now define the following:

Definition 2.2. Given any function $\psi : C \to \mathbb{R}$ with $c \leq \psi(c)$ for all $c \in C$:

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

2. A mechanism (σ, τ) is a *ψ-quasi-perfect mechanism* if σ is a ψ-quasi-perfect scheme and if the producer with marginal cost $\overline{c}$, when reporting truthfully, has net profit $\pi_0(\overline{c})$.

10. Notice that when the producer's marginal cost is c, no trade occurs in market segment D if and only if $D(c) = 1$.

With the definitions above, the main characterization of this chapter can be stated as Theorem 2.2.

Theorem 2.2 (Optimal Mechanism)**.** *Under regime $\mathcal{I}$, the set of optimal mechanisms is nonempty and is exactly the set of incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanisms. Furthermore, every optimal mechanism induces $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G-almost all $c \in C$.*

The proof of Theorem 2.2 can be found in Appendix D. Below, I introduce some simplifying assumptions and sketch the proof, which consists of three main steps: The revenue equivalence formula, constructing an incentive feasible ψ-quasi perfect mechanism, and the proof of uniqueness.

The Revenue Equivalence Formula and an Upper Bound for Revenue

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

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

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

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

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

$$\pi'_D(c) = -D(\overline{\boldsymbol{p}}_D(c)). \tag{2.2}$$

Together,

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

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

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

Together, these lead to Lemma 2.1 below.

Lemma 2.1. *A mechanism (σ, τ) is incentive compatible if and only if*

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

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

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

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

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

The proof of Lemma 2.1 can be found in Appendix D. It formalizes the heuristic arguments above by using the envelope theorem of Milgrom and Segal (2002). In essence, condition 1 in Lemma 2.1 is a generalized revenue-equivalence formula stating that the transfer τ must be determined by σ up to a constant, whereas condition 2 in Lemma 2.1 is stronger than the usual monotonicity condition due to the high-dimensionality nature of the allocation space and the lack of the single-crossing property.

From Lemma 2.1, for any incentive compatible mechanism (σ, τ), the data broker's expected revenue can be written as

$$\mathbb{E}[\tau(c)] = \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c))\, D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - U(\overline{c}), \tag{2.3}$$

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

Combining (2.3) and Theorem 2.1, together with the definition of quasi-perfect mechanisms, it then follows that any incentive feasible $\overline{\varphi}_G$ quasi-perfect mechanism must be optimal. Therefore, to prove the first part of Theorem 2.2, it suffices to construct an incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism.

Attaining R^*

By the definition of quasi-perfect segmentations, for any nondecreasing function $\psi : C \to \mathbb{R}_+$ and for any ψ-quasi-perfect scheme σ, given any 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. That is, 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 $v \geq \psi(c)$.[11] Specifically, for all $c \in C$,

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

and

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

As a result, if there is an incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism (σ, τ), then by Lemma 2.1, the data broker can attain revenue

$$\begin{aligned} \mathbb{E}[\tau(c)] &= \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \\ &= \int_C \left(\int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \qquad (2.4) \\ &= \bar{R}. \end{aligned}$$

However, not every quasi-perfect scheme is implementable. To ensure incentive compatibility, the integral inequality given by condition 2 in Lemma 2.1 must be satisfied. While this condition involves a continuum of constraints and is difficult to check, the following lemma provides a simpler sufficient condition.

Lemma 2.2. *For any nondecreasing function $\psi : C \to \mathbb{R}_+$ with $\psi(c) \geq c$ for all $c \in C$, and for any ψ-quasi-perfect scheme σ, there exists a transfer scheme $\tau : C \to \mathbb{R}$ such that (σ, τ)*

11. The formal arguments can be found in the proof of Lemma 2.10 in Appendix C.

is incentive compatible if for any $c \in C$,

$$\psi(z) \leq \overline{\boldsymbol{p}}_D(z), \tag{2.5}$$

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

In essence, Lemma 2.2 reduces the integral inequalities given by condition 2 of Lemma 2.1 to pointwise conditions. Details about the proof can be found in Appendix D. The crucial step is to notice that for a ψ-quasi-perfect scheme, there are always no downward-deviation incentives (i.e., a producer with cost c would never have an incentive to misreport $c' < c$), as a higher-cost producer would find the gains from reducing the cutoff less beneficial than the increment in transfer. With this observation, as the pointwise condition (2.5) is sufficient to rule out upward-deviation incentives, Lemma 2.2 then follows.

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

Lemma 2.3. *For any nondecreasing function $\psi : C \to \mathbb{R}_+$ such that that $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$, there exists a ψ-quasi-perfect scheme σ^* that satisfies* (2.5).

A direct consequence of Lemma 2.2 and Lemma 2.3 is that there exists an incentive compatible $\overline{\varphi}_G$-quasi-perfect mechanism (σ^*, τ^*). Furthermore, for any $c \in C$, (2.1) also implies that

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

Together, by Lemma 2.1 and (2.2), after possibly adding a constant to τ^* so that the indirect utility of the producer with cost $\overline{c}$ equals to $\pi_0(\overline{c})$, (σ^*, τ^*) is an incentive feasible ϕ_G-quasi-perfect mechanism, which in turn implies that (σ^*, τ^*) is optimal. Together with (2.4), it then follows that any incentive feasible ϕ_G-quasi-perfect mechanism is optimal.

The proof of Lemma 2.3 is by construction and the details can be found in Appendix D. This construction is especially simple when m^0 is regular (equivalently, when the profit function $p \mapsto (p-c)D_0(p)$ is single-peaked on $\text{supp}(D_0)$ for all $c \in C$). Specifically, for any $c \in C$ and for any $v \in [\psi(c), \overline{v}]$, define market segment $D_v^{\psi(c)} \in \mathcal{D}$ as

$$D_v^{\psi(c)}(p) := \begin{cases} D_0(p), & \text{if } p \in [\underline{v}, \psi(c)] \\ D_0(\overline{\varphi}(c)), & \text{if } p \in (\psi(c), v] \\ 0, & \text{if } p \in (v, \overline{v}] \end{cases}, \tag{2.6}$$

for all $p \in V$, which is illustrated in Figure 2.1. Also, let $\sigma^* : C \to \Delta(\mathcal{D})$ be defined as (2.10) with $\overline{\varphi}_G$ being replaced by ψ. By construction, $\sigma^*(c) \in \mathcal{S}$ for all $c \in C$. Furthermore, σ^* is a ψ-quasi-perfect scheme satisfying (2.5). To see this, consider any $c \in C$. By regularity of D_0 and by the hypothesis that $\psi(c) \leq \overline{\boldsymbol{p}}_0(c)$, when the producer's marginal cost is c, she would prefer charging price $\psi(c)$ (or the lowest price in $\text{supp}(D_0)$ that is above $\psi(c)$, if $\psi(c) \notin \text{supp}(D_0)$) than charging any price $p < \psi(c)$ under D_0. Therefore, for any $v \geq \psi(c)$ and for any $p < \psi(c)$, since $D_v^{\psi(c)}(p) = D_0(p)$ and $D_v^{\psi(c)}(v) = D_0(\psi(c))$, charging price v in segment $D_v^{\psi(c)}$ must be optimal for the producer as $v \geq \psi(c)$. On the other hand, when the producer has marginal cost $z < c$, for any $v \geq \psi(c)$, since $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}$ is nonincreasing, it must be that either $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z) = v$ or $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z) < \psi(c)$. In the former case, since ψ is nondecreasing, it then follows that $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z) \geq \psi(c) \geq \psi(z)$. In the latter case, as $D_v^{\psi(c)}(p) = D_0(p)$ for all $p < \psi(c)$, $\overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z)$ must have been optimal for the producer under D_0 as well. Therefore, it must be that $\overline{\boldsymbol{p}}_0(z) \leq \overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z)$. Combining with the hypothesis that $\psi(z) \leq \overline{\boldsymbol{p}}_0(z)$, this then implies $\psi(z) \leq \overline{\boldsymbol{p}}_{D_v^{\psi(c)}}(z)$. As a result, σ^* is indeed a ψ-quasi-perfect scheme satisfying (2.5).

In general, for any arbitrary $D_0 \in \mathcal{D}$, the construction is more convoluted. In brief, the segmentation scheme σ^* is constructed by approximating D_0 with a sequence of step functions $\{D_n\} \subseteq \mathcal{D}$ that converges to D_0, followed by finding a desired segmentation scheme σ_n of each D_n. Together with several continuity lemmas in Appendix B, the limit of $\{\sigma_n\}$ converges to the desired segmentation scheme σ^*.

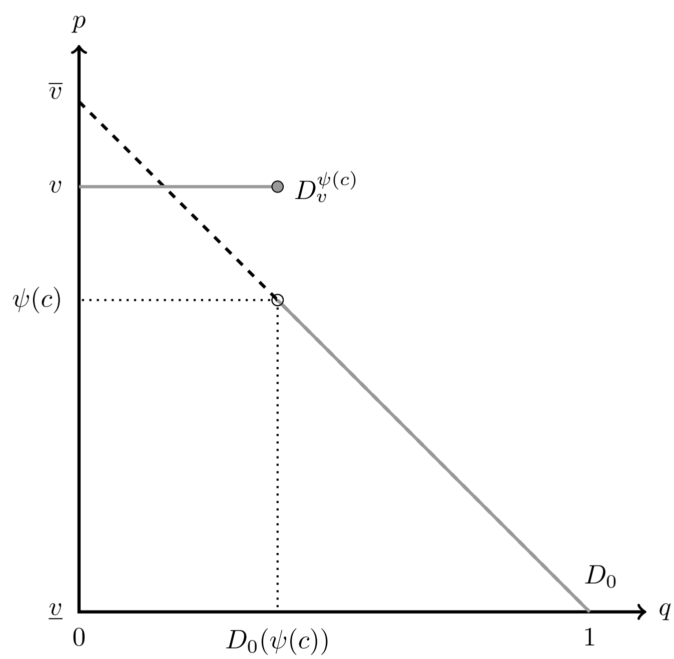

Figure 2.1: Market segment $D_v^{\psi(c)}$

Uniqueness

To see why any optimal mechanism of the data broker is a $\overline{\varphi}_G$-quasi-perfect mechanism, I sketch the intuition under an assumption that is stronger than Assumption 2.1, the formal proof for the more general case involves details of the duality argument used to prove Theorem 2.1 and thus is relegated to Appendix D. To this end, in this subsection, assume that

$$\phi_G(c) \leq \overline{\boldsymbol{p}}_0(c), \forall c \in C \tag{2.7}$$

and that G is regular. Notice that $\overline{\varphi}_G(c) = \phi_G(c)$ for all $c \in C$ when (2.7) holds and when G is regular.

Under these assumptions, suppose that a mechanism (σ, τ) is optimal. Then,

$$
\begin{aligned}
R^* &= \int_C \left(\int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \pi_0(\bar{c}) \\
&= \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \pi_0(\bar{c}),
\end{aligned} \tag{2.8}
$$

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

$$
\int_{\{v \geq \phi_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) = \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma(\mathrm{d}D|c), \tag{2.9}
$$

since the left-hand side is the efficient surplus in an economy where the producer's cost is $\phi_G(c)$ and hence must be an upper-bound of the right-hand side, and (2.8) implies that the right-hand side must attain this upper bound.

It then follows σ must be a $\overline{\varphi}_G$-quasi-perfect mechanism. Indeed, if σ is not a $\overline{\varphi}_G$-quasi-perfect scheme, it must be that there is a positive G-measure of $c \in C$ and a positive $\sigma(c)$-measure of $D \in \operatorname{supp}(\sigma(c))$ such that either $D(v) > 0$ for some $v > \overline{\boldsymbol{p}}_D(c)$, or $D(\phi_G(c)) \neq D(\overline{\boldsymbol{p}}_D(c))$. That is, either there are some consumers with $v \geq \phi_G(c)$ who do not buy the product or buy the product at a price below v, or there are some consumers with $v < \phi_G(c)$ who end up buying the product. This contradicts (2.9). As a result, (σ, τ) must be a ϕ_G-quasi-perfect mechanism. Moreover, (σ, τ) must also induce quasi-perfect price discrimination since $\overline{\boldsymbol{p}}$ can be replaced with any $\boldsymbol{p} \in \boldsymbol{P}$ according to Lemma 2.1.

2.3.3 Further Remarks

From the definition of quasi-perfect segmentations, there are some degrees of freedom regarding the ways to pool the low-value consumers with the high-values. Therefore, Theorem 2.2 implies that there might be multiple optimal mechanisms—as long as the low-value consumers are pooled with the high-values in a way such that the mechanism is incentive feasible and is $\overline{\varphi}_G$-quasi-perfect. Nevertheless, the outcome induced by any optimal mechanism is unique. That is, under any optimal mechanism, for (almost) all marginal cost $c \in C$, a consumer with value v buys the product if and only if $v \geq \overline{\varphi}_G(c)$ and all the purchasing consumers pay their values. In other words, the multiplicity only accounts for the off-path incentives. Furthermore, there is always an explicit construction of an optimal mechanism (see details in Appendix D). In fact, when the market demand m^0 is regular, the argument above for the sketched proof of Lemma 2.3 implies that there is an optimal mechanism that takes a particularly simple form particularly simple form: The low-value consumers are pooled with the high values in a way that preserves the likelihood ratios among values below the cutoff. Specifically, for any $c \in C$ and for any $p \in [\overline{\varphi}_G(c), \overline{v}]$, let

$$\sigma^{**}\left(\left\{D_v^{\overline{\varphi}_G(c)} : v \geq p\right\} \middle| c\right) := \frac{D_0(p)}{D_0(\overline{\varphi}_G(c))}. \tag{2.10}$$

In other words, for any $c \in C$, $\sigma^{**}(c)$ only assigns positive measure to market segments $\{D_v^{\overline{\varphi}_G(c)}\}_{v \in [\psi(c), \overline{c}]}$ and its distribution is exactly the distribution of consumers' values conditional on being above the cutoff $\overline{\varphi}_G(c)$ given by the market demand. Then, by Lemma 2.1, there exists a unique transfer scheme $\tau^{**} : C \to \mathbb{R}$ such that (σ^{**}, τ^{**}) is an incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism and hence is optimal. Henceforth, I refer the mechanism (σ^{**}, τ^{**}) as the *canonical* $\overline{\varphi}_G$-quasi-perfect mechanism. This then gives the following result.

Theorem 2.3. *Suppose that m^0 is regular. Then the canonical $\overline{\varphi}_G$-quasi-perfect mechanism (σ^{**}, τ^{**}) is optimal.*

As an example, consider the case where D_0 is linear and G is a uniform distribution with $V = C = [0, 1]$. It then follows $\overline{\varphi}_G(c) = 2c$ for all $c \in [0, 1/3]$ and $\overline{\varphi}_G(c) = (1 + c)/2$ for all $c \in (1/3, 1]$. In this case, the canonical $\overline{\varphi}_G$-quasi-perfect mechanism is described by a uniform distribution on the market segments $\{D_v^{\overline{\varphi}_G(c)}\}_{v \in [\overline{\varphi}_G(c), 1]}$, where each market segment $D_v^{\overline{\varphi}_G(c)}$ is defined by (2.6).

In addition, Theorem 2.2 underlines a noteworthy feature of the optimal mechanisms. According to Theorem 2.2, for any optimal mechanism (σ, τ), the segmentation scheme σ does not generate value-revealing segmentations in general. Specifically, for any report c such that $\overline{\varphi}_G(c) > \underline{v}$, there are market segments $D \in \text{supp}(\sigma(c))$ containing consumers with distinct values. The reason is that in order to incentivize the producer to set prices in desirable ways and to elicit information from the producer, some market segments must contain consumers with values below the desirable threshold $\overline{\varphi}_G(c)$. By pooling the high-value consumers with the low-value ones in the same market segment while separating them from other high-value consumers, the data broker is able to incentivize the producer to set prices at the highest value in each market segment and induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for all c, which in turn enables the data broker to elicit private information by discouraging trade and extract surplus from the purchasing consumers at the same time. This also means it is not optimal for the data broker to release all the information about consumers' values.

2.4 Implications

2.4.1 Outcome-Equivalence

The results in Section 2.3 immediately leads to the outcome-equivalence result. Indeed, by Theorem 2.1 and Theorem 2.2, every optimal mechanism under either regime $\mathcal{I}$ or regime $\mathcal{P}$ must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for (almost) all $c \in C$. This then implies that the data broker's revenue, the producer's profit, consumer surplus and the allocation of the product, must remain unchanged regardless of the market regime, as summarized below.

Theorem 2.4 (Outcome-Equivalence)**.** *Regime $\mathcal{I}$ and regime $\mathcal{P}$ are outcome-equivalent (under* $\mathrm{X}_{\mathrm{Ch2}}$*).*

Theorem 2.4 implies that, the ability to contract on prices is irrelevant. The ability to design market segmentations (equivalently, information structures) is powerful enough so that the data broker can achieve the same outcome even if prices are not contractable. The essence of this result relies on the richness of the allocation space. Indeed, as the space of market segmentations is large, the data broker can exploit the allocations to provide extra incentives for the producer to discipline her pricing decisions.

Furthermore, Theorem 2.4 means that even when the data broker only affects the product market indirectly by selling consumer data, the market outcomes he induces would be the same as those when he can control product prices. More specifically, from the data broker's perspective, having control over how the product is sold in addition to consumer data adds no extra values to his revenue. As for the producer, her profit in face of a data broker is the same as if she sells the product, as well as the exclusive right to sell the product, to this data broker. Preserving the 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.

This outcome-equivalence result has several further implications. First, it implies that there are no incentives for the data broker to become more active, as the data broker's revenue would remain the same even if he gains control over prices. Second, from a policymaker's perspective, it means that no further concerns should be raised even if a data broker eventually becomes more active. After all, the market outcomes and the amount of deadweight loss would remain the same.

2.4.2 Welfare Implications and Consumer Data Ownership

One of the most pertinent questions about consumer-data brokership is how it affects consumer surplus. Are the data broker's possession of consumer data and the ability to sell them to a producer detrimental for the consumers? If so, to what extent? Meanwhile, can the consumers benefit from the fact that the data broker does not have access to the consumers and only affects the product market indirectly by selling data to the producer? While currently being a focus of policy debates, the following result, as an implication of Theorem 2.2, answers a certain aspect of this question.

Proposition 2.1 (Surplus Extraction)**.** *Consumer surplus is zero under any optimal mechanism.*

Proposition 2.1 follows directly from the characterization given by Theorem 2.2. Indeed, according to Theorem 2.2, any optimal mechanism must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for (almost) all $c \in C$, which means that every purchasing consumer must be paying their values. Notably, Proposition 2.1 provides an unambiguous and substantial assertion about the consumer surplus under data brokership. According to Proposition 2.1, even though the data broker does not sell the product to the consumers directly and only affects the market by creating market segmentations for the producer, it is as if the consumers are perfectly price discriminated and all the surplus is extracted away (even though the optimal mechanisms do not induce perfect price discrimination in general). This means that as long as the data broker possesses consumer data and can sell them to a producer, from the consumers' perspective, it is the same as buying the product from a monopolist who can implement perfect price discrimination. More practically, this result means it is impossible to expect the consumers to benefit from the gap between the ownership of production technology and ownership of consumer data.

Proposition 2.2. *The data broker's optimal revenue is no less than the consumer surplus under uniform pricing.*

An immediate consequence of Proposition 2.2 is that total surplus under data brokership is greater compared with uniform pricing, as summarized below.

Proposition 2.3 (Total Surplus Improvement)**.** *Data brokership always increases total surplus compared with uniform pricing.*

Proposition 2.3 means that even though data brokership is extremely harmful to the consumers, in terms of total surplus it creates, however, it is always better than the environment where no information about the consumers' values can be disclosed.

Another implication of Proposition 2.2 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 setting, the data broker does not have any data in the first place and has to purchase them from the consumers.[12] Proposition 2.2 immediately implies that, if the data broker has to purchase data by compensating the consumers with monetary transfers *before* they learn their values,[13] then the optimal mechanism would be to purchase all the data by paying the consumers their ex-ante surplus under uniform pricing and then use any optimal mechanism characterized by Theorem 2.2 to sell these data to the producer. Furthermore, since the data broker's revenue is greater than the consumer surplus under uniform pricing according to Proposition 2.2, and since the producer always has an outside option of uniform pricing, this outcome is in fact Pareto improving compared with uniform

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

13. 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. Such ex-ante purchase of consumer data is plausibly suitable for online activities. After all, in online settings, consumers often do not consider their values about a particular product when they agree that their personal data such as browsing histories, IP address and cookies, can be collected by the data brokers. Nevertheless, other purchase timing would also be a relevant question, which can be explored in future research.

pricing in the ex-ante sense, as stated below.

Proposition 2.4 (Data Ownership)**.** *If the consumers own their data and if the data broker can purchase data from the consumers before they learn their values, then data brokership is Pareto improving compared with uniform pricing in the ex-ante sense.*

2.5 Extensions

2.5.1 Sufficient Conditions and Relaxations of Assumption 2.1

Despite being a technical condition, Assumption 2.1 has an economically interpretable sufficient condition (2.7). To better understand this condition, recall that by definition, $\phi_G(c) = c + G(c)/g(c)$, and therefore $\phi_G(c)$ is the actual marginal cost c plus the information rent $G(c)/g(c)$. On the other hand, instead of regarding $\overline{\boldsymbol{p}}_0(c)$ as the optimal uniform price for the producer when her marginal cost is c, $\overline{\boldsymbol{p}}_0(c)$ can be written as $\overline{\boldsymbol{p}}_0(c) = c + \xi_0(c)$, where $\xi_0(c) := \overline{\boldsymbol{p}}_0(c) - c$ is the *monopoly mark-up* that the producer charges under uniform pricing. From this perspective, (2.7) is equivalent to

$$\frac{G(c)}{g(c)} \leq \xi_0(c), \ \forall c \in C.$$

That is, the information rent that the producer retains due to asymmetric information about her marginal cost is less than her monopoly mark-up.

Furthermore, (2.7) can also be interpreted as the gains from trade being large enough. More specifically, for any demand $D_0 \in \mathcal{D}$, define a location family $\{D_0^k\}_{k \geq 0}$ by moving the support of D_0 without changing the shape of the distribution. That is, $D_0^k(p) := D_0(p-k)$ for all $p \in V$ and for all $k > 0$. Within this family, it is natural to rank the gains from trade by the location parameter k. In the Supplemental Material, I show that there exists $\overline{k} \geq 0$

such that (2.7) holds if and only if $k \geq \overline{k}$.[14]

Although the results introduced above rely on Assumption 2.1, the sole purpose of Assumption 2.1 is to ensure that as a revenue upper bound, the data broker's problem under regime $\mathcal{P}$ has a closed form solution. After all, as argued in Section 2.3, the data broker's problem under regime $\mathcal{P}$ is essentially a nonlinear screening problem with one-dimensional allocation space and type-dependent outside options. A common feature of such problems is that the characterization of the optimal mechanisms involves Lagrange multipliers in general (see, for instance, Lewis and Sappington (1989) and Jullien (2000)). Assumption 2.1, however, yields a closed form solution for the data broker's problem (Theorem 2.1), which in turn allows an explicit construction of an incentive feasible mechanism for the data broker that attains the revenue upper bound.

Consequently, many of the results can be extended to environments without Assumption 2.1. First, Proposition 2.1 actually does not rely on Assumption 2.1 at all. A strengthened version of Proposition 2.1 can be found in the Supplemental Material of Yang (2020), which ensures both the existence of an optimal mechanism for the data broker and the fact that any optimal mechanism must yield zero consumer surplus. A crucial step of the proof is to take any mechanism (σ, τ) under which the consumers retain positive surplus and apply Lemma 2.3 to every market segment $D \in \text{supp}(\sigma(c))$ for every report c, with the cutoff function ψ being $\overline{\boldsymbol{p}}_D$. This would induce another segmentation scheme. The fact that all the market segments $D \in \text{supp}(\sigma(c))$ are decomposed according to a $\overline{\boldsymbol{p}}_D(c)$-quasi-perfect segmentation and the hypothesis that consumers retain positive surplus under (σ, τ) yield a strict improvement on the data broker's revenue. Moreover, (2.5) ensures that such decomposition relaxes the incentive compatibility and individual rationality constraints.

In addition, the main characterization (Theorem 2.2) can be generalized as well. More

14. Clearly, if k is large enough so that $\underline{v} \geq \phi_G(\overline{c})$, then there is common knowledge of gains from trade even after incorporating the information rents and hence the value-revealing scheme would be optimal. In the Supplemental Material of Yang (2020) I show that there exists k such that (2.7) holds even then $\underline{v} < \phi_G(\overline{c})$.

specifically, in the Supplemental Material of Yang (2020), I show that as long as D_0 is continuous, there exists a nondecreasing function φ^* such that every optimal mechanism must be a φ^*-quasi-perfect mechanism. However, unlike $\overline{\varphi}_G$, the cutoff function φ^* does not have a closed form solution.

2.5.2 Consumers' Private Information

Given the amount of consumer data that can be collected, their predictive power is approaching perfect estimations of consumers' values. Nonetheless, it is still imperative to explore the economic implications of the possibility when the consumers have some private information. This section extends the baseline model in Section 2.2 and allows the consumers to retain some pieces of information.

To formally model this, let Θ be a Polish space that denotes a set of consumer characteristics which can be disclosed by the data broker. Suppose that among the consumers, their available characteristics θ are distributed according to $\beta_0 \in \Delta(\Theta)$. These characteristics are informative about the consumers' values but there may still be variation in values even among the consumers who share the same characteristics. Specifically, given any $\theta \in \Theta$, suppose that among the consumers who share characteristic θ, their values are distributed according to $m^\theta \in \Delta(V)$ and m^θ induces a demand $D_\theta \in \mathcal{D}$ (i.e., $D_\theta(p) := m^\theta([p, \overline{v}])$ for all $p \in P$) for each $\theta \in \Theta$. The data broker is only able to segment the market according to θ but not v. In this environment, a market segmentation is then defined by $s \in \Delta(\Delta(\Theta))$ such that

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

for any measurable $A \subseteq \Theta$. As a result, there is now a limit on how predictive the data can be and the environment is described by $(\{D_\theta\}_{\theta \in \Theta}, \beta_0)$.

To simplify analyses, I further specialize the environment. Suppose that there are finitely many possible characteristics. That is, $|\Theta| < \infty$. Moreover, suppose that $\{\text{supp}(D_\theta)\}_{\theta \in \Theta}$

forms a partition of V and $\text{supp}(D_\theta)$ is an interval for all $\theta \in \Theta$. This specialization will be referred as *partitional*. In other words, the data broker only has partial information about the consumers' values and can at most identify which interval a consumer's value belongs to. Even when θ is perfectly revealed, the producer would still be unable to identify each consumer's value. In this environment, the market demand D_0 is given by

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

for all $p \in V$. Moreover, a market segmentation s induces market segments $\{D_\beta\}_{\beta \in \text{supp}(s)}$ and

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

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

When the consumers' values can never be fully disclosed, it is clear that their surplus will increase. After all, it is no longer possible for the producer to charge the consumers their values as the additional variation in values given by D_θ always allows some consumers to buy the product at a price that is below their values. Nevertheless, as shown in Theorem 2.5, under any optimal mechanism, consumer surplus must be lower than the case when all the information about θ is revealed to the producer. That is, the main implication of Proposition 2.1— for the consumers, the presence of a data broker is no better than a scenario where their data is fully revealed to the producer—is still valid even when the consumers retain some private information through a partitional $(\{D_\theta\}_{\theta \in \Theta}, \beta_0)$.

Theorem 2.5. *For any partitional* $(\{D_\theta\}_{\theta \in \Theta}, \beta_0)$ *and any distribution of marginal cost G, an optimal mechanism exists. Furthermore, the consumer surplus under any optimal mechanism of the data broker is lower than the case when* θ *is fully disclosed.*

The intuition behind Theorem 2.5 is simple. Since there are only finitely many characteristics, identifying the consumers' characteristic θ effectively enables the producer to categorize the consumers into finitely many "blocks" so that every possible value belongs to one and only one block. As a result, when changing prices within each block of values, the trading volume is only affected by purchasing decisions of the consumers whose values are within that block. Such separability allows an analogous argument as in the proof of the generalized version of Proposition 2.1 (provided in the Supplemental Material) which shows that the data broker can always construct a mechanism that increases its revenue if the consumer surplus is higher than that when the characteristic θ is not full-revealed.

In addition to the surplus extraction result, the characterization of the optimal mechanisms can be generalized as well. That is, with proper regularity conditions, there is an optimal mechanism that is analogous to the canonical $\overline{\varphi}_G$-quasi-perfect mechanism introduced in Section 2.3. To state this result, given any partitional $(\{D_\theta\}_{\theta\in\Theta}, \beta_0)$, for each $\theta \in \Theta$, write $\mathrm{supp}(D_\theta)$ as $[l(\theta), u(\theta)]$. For any $p \in V$, let $\theta_p \in \Theta$ be the unique θ such that $p \in (l(\theta), u(\theta)]$. For any $c \in C$, let $\widehat{\boldsymbol{p}}_0(c)$ be the largest optimal price for the producer with marginal cost $c \in C$ under the demand whose support contains $\overline{\boldsymbol{p}}_0(c)$.[15] Also, let $\widehat{\varphi}_G(c) := \min\{\varphi_G(c), \widehat{\boldsymbol{p}}_0(c)\}$ for all $c \in C$. Furthermore, given any function $\psi : C \to \mathbb{R}_+$, say that a mechanism (σ, τ) is a canonical ψ-quasi-perfect segmentation if the producer with marginal cost $\overline{c}$, when reporting truthfully, receives $\pi_0(\overline{c})$, and if for any $c \in C$, and for any $\beta \in \mathrm{supp}(\sigma(c))$, either

$$\beta(\theta') = \beta^{\theta}_{\psi(c)}(\theta') := \begin{cases} \beta_0(\theta'), & \text{if } u(\theta') < \psi(c) \text{ and } u(\theta) \geq \psi(c) \\ \sum_{\{\hat{\theta}:u(\hat{\theta})\geq\psi(c)\}} \beta_0(\hat{\theta}), & \text{if } u(\theta') \geq \psi(c) \text{ and } \theta' = \theta \\ 0, & \text{otherwise} \end{cases}, \quad (2.11)$$

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

$$\mathrm{supp}(\beta) = \{\theta' : l(\theta') \leq \psi(c)\} \cup \{\theta\} \quad (2.12)$$

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

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

$$\beta(\theta') = \beta_0(\theta'). \tag{2.13}$$

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

With these definitions, Theorem 2.6 below prescribes an optimal mechanism for the data broker.

Theorem 2.6. *For any partitional* $(\{D_\theta\}_{\theta \in \Theta}, \beta_0)$ *and any distribution of marginal cost* G *such that the function* $c \mapsto \max\{(\phi_G(c) - \widehat{\boldsymbol{p}}_0(c)), 0\}$ *is nondecreasing and that* m^0 *is regular, there is a canonical* $\widehat{\varphi}_G$*-quasi-perfect mechanism that is optimal.*

2.5.3 Targeted Marketing

So far, the discussions have been abstracting away the possibility that the data broker can use consumer data to facilitate targeted marketing by assuming there is only one product. In fact, one of the most common arguments in favor of the usage and provision of consumer data is that it also benefits the consumers because more relevant products can be advertised to the consumers and therefore more surplus can be created. The following extension explores this aspect.

Formally, suppose that, instead of a single product, there are $J \in \mathbb{N}$ different producers who sell J different products. In addition, there are $I \in \mathbb{N}$ (equally populated) groups of consumers. Each group of consumers has different preferences about different products. More specifically, let $\mathcal{J} := \{1, \dots, J\}$ be the set of producers and let $\mathcal{I} := \{1, \dots, I\}$ be the set of all possible groups. For each $i \in \mathcal{I}$ and each $j \in \mathcal{J}$, the distribution of consumers' values in group i for product j is $D_0^{ij} \in \mathcal{D}$. For one of the results below, it is further assumed that for each product $j \in \mathcal{J}$, $\{D_0^{ij}\}_{i \in \mathcal{I}}$ can be ranked by pointwise ordering. The interpretation is that for each product, different groups value a product differently and some group prefers a product more than others.

For each producer $j \in \mathcal{J}$, her marginal production cost $c_j \in C_j = [\underline{c}_j, \overline{c}_j]$ is her private information that follows a distribution G_j. Assume that the marginal costs are independent across producers. Define $C := \prod_{j \in \mathcal{J}} C_j$ and use $c = (c_1, \ldots, c_J)$ to denote a typical element of C. Also, let $G := \prod_{j \in \mathcal{J}} G_j$ be the joint distribution of the producers' marginal costs. As in the baseline model, each producer can sell her product to the consumers but does not know the individual consumer's value a priori. Furthermore, the producer does not have the targeting technology and thus the consumers she faces in absence of the data broker are summarized by

$$D_0^j := \frac{1}{I} \sum_{i \in \mathcal{I}} D_0^{ij}.$$

That is, without targeting, the consumers who see producer j's product are uniformly drawn from each group.

The data broker can create market segmentations and sell them to the producers. In addition, he can help the producers target their products to different group of consumers. Formally, for any $i \in \mathcal{I}$ and any $j \in \mathcal{J}$, let $\mathcal{S}_{ij}$ denote the collection of $s \in \Delta(\mathcal{D})$ satisfying (2.1) with D_0 being replaced by D_0^{ij}. A mechanism is defined as a tuple $(\sigma, \tau, q) = (\sigma_{ij}, \tau_j, q_{ij})_{i \in \mathcal{I}, j \in \mathcal{J}}$, where for each $i \in \mathcal{I}$, $j \in \mathcal{J}$, $\sigma_{ij} : C \to \mathcal{S}_{ij}$ is the *segmentation scheme*; $q_{ij} : C \to [0, 1]$ such that $\sum_{i \in \mathcal{I}} q_{ij} \leq 1$ is the *targeting scheme* so that $q_{ij}(c)$ stands for the fraction of consumers of group i that can see product j;[16] and $\tau_j : C \to \mathbb{R}$ is the transfer scheme for producer j. A mechanism (σ, τ, q) is said to be incentive compatible if for any

16. Targeting can only re-direct the consumers who are able to see each product but cannot create new demand. As such, the total volume of consumers who can see product j must be less than $\sum_{i \in \mathcal{I}} 1/I = 1$.

$j \in \mathcal{J}$ and for any $c_j, c'_j \in C_j$,

$$\begin{aligned}
&\mathbb{E}_{c_{-j}}\left[\sum_{i\in\mathcal{I}}\int_{\mathcal{D}}\pi_D(c_j)\sigma_{ij}(\mathrm{d}D|c_j,c_{-j})q_{ij}(c_j,c_{-j})-\tau_j(c_j,c_{-j})\right]\\
\geq&\mathbb{E}_{c_{-j}}\left[\sum_{i\in\mathcal{I}}\int_{\mathcal{D}}\pi_D(c_j)\sigma_{ij}(\mathrm{d}D|c'_j,c_{-j})q_{ij}(c'_j,c_{-j})-\tau_j(c'_j,c_{-j})\right],
\end{aligned}$$

and is individually rational if for any $j \in \mathcal{J}$ and any $c_j \in C_j$,

$$\mathbb{E}_{c_{-j}}\left[\sum_{i\in\mathcal{I}}\int_{\mathcal{D}}\pi_D(c_j)\sigma_{ij}(\mathrm{d}D|c_j,c_{-j})q_{ij}(c_j,c_{-j})-\tau_j(c_j,c_{-j})\right]\geq\pi_{D_0^j}(c_j).$$

Proposition 2.1 can be generalized to the environment in which targeted marketing is possible, as summarized in Theorem 2.7.

Theorem 2.7. *For any demands $\{D_0^{ij}\}_{i\in\mathcal{I},j\in\mathcal{J}}$ and any marginal cost distributions $\{G_j\}_{j\in\mathcal{J}}$, there exists an incentive feasible mechanism that maximizes the data broker's revenue. Moreover, under any revenue-maximizing mechanism, consumers retain zero surplus.*

Theorem 2.7 implies that even with the additional targeting technology, the consumers still retain no surplus. The reason is that, even though the ability to target consumers increases total surplus, the data broker can always design segmentations and targeting schemes that extract all of the additional surplus created by targeting. The groups of consumers whose values are low will not be exposed to a product, whereas the surplus of the groups of consumers whose values are high enough are entirely extracted away due to price discrimination, even if they are targeted.

In addition to the implications for consumer surplus, since every group of consumer can buy from all of the J producers as long as they see the product, the data broker's problem is in fact similar to that in the baseline model. To maximize revenue, he will simply select the most profitable group of consumers for producer j and target producer j's product to that group. This observation leads to the following generalization of Theorem 2.4. That is, even in environments where targeted marketing is possible, under certain appropriate assumptions about the distributions of marginal costs and the market demands, regime $\mathcal{I}$ and $\mathcal{P}$ are still outcome-equivalent.

Theorem 2.8. *For any demands $\{D_0^{ij}\}_{i\in\mathcal{I},j\in\mathcal{J}} \subset \mathcal{D}$ such that $\{D_0^{ij}\}_{i\in\mathcal{I}}$ is ordered according to pointwise ordering for each $j \in \mathcal{J}$, and for any regular distributions of marginal costs $\{G_j\}_{j\in\mathcal{J}}$, suppose that for any $i \in \mathcal{I}$ and any $j \in \mathcal{J}$,$\phi_{G_j}(c) \leq \min\{\overline{\boldsymbol{p}}_{D_0^{ij}}(c), \overline{\boldsymbol{p}}_{D_0^{j}}(c)\}$ for all $c \in C$. Then regime $\mathcal{I}$ and $\mathcal{P}$ are outcome-equivalent.*

2.6 Discussions

The results above have several broader policy implications. First, in terms of welfare, although Proposition 2.1 implies that data brokership is undesirable for the consumers, Proposition 2.3 shows that the total surplus is always higher with 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 the question of whether a data broker is beneficial must depend on the objective of the policymaker and the kinds of redistributional policy tools available. If the policymaker's objective is simply maximizing total surplus, or if redistributional tools such as lump-sum transfers are available, then it is indeed beneficial to allow a data broker to sell consumer data. On the other hand, however, if the policymaker also concerns themselves with consumer surplus, and if no effective redistributional policies are available, then the presence of a data broker can be extremely unfavorable. Therefore, regarding the policy debates about whether a data broker should be allowed to collect, use and trade consumer data, it is imperative to first identify the available redistributional tools and the relative importance among consumer surplus, producer profit and total surplus.

In the case where the policymaker does wish to improve consumer surplus and no effective redistributional policies are available, Theorem 2.5 and Theorem 2.7 imply that there are limited possible policies that can be used to improve consumer surplus. Clearly, policies that help the consumers preserve some private information can improve consumer surplus. Nonetheless, according to Theorem 2.5, from the consumers' perspective, data brokership is still no better than all the characteristics being revealed. On the other hand, targeted marketing does not benefit the consumers either. As shown in Theorem 2.7, even though targeting technology can be used to further increase the total surplus, all the benefits will be extracted away from the consumers via price discrimination.

Furthermore, even if a policymaker attempts to improve consumer surplus by monitoring price discrimination, Theorem 2.2 implies that it is not enough to monitor only whether there is personalized pricing (i.e., first degree price discrimination). In fact, the optimal mechanisms given by Theorem 2.2 do not exhibit perfect price discrimination. Instead, it is a certain kind of third degree price discrimination (i.e., quasi-perfect price discrimination, induced by the quasi-perfect segmentations) that will arise in this environment. For a policymaker, prohibiting personalized pricing would not be effective in improving consumer surplus. Rather, identifying whether third degree price discrimination is operated in the form of quasi-perfect segmentation is indispensable for improving consumer surplus. Nevertheless, quasi-perfect segmentations may sometimes be difficult to identify. In general, it might be difficult to distinguish quasi-perfect segmentations from other basic forms of third degree price discrimination unless the policymaker has complete knowledge of the correlation between the disclosed consumer characteristics and the consumers' values. Whether it is possible to identify an quasi-perfect segmentation with less knowledge remains as a topic for future studies.

In contrast to the seemingly pessimistic implications discussed above, Proposition 2.4 prescribes a rather positive solution, both in terms of consumer surplus and in terms of total welfare. According to Proposition 2.4, if the data broker has to purchase the data from the consumers, and if the purchase takes 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 for their own data,[17] 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.

Finally, regardless of the policymaker's objective, as long as it depends only on the market outcomes, the equivalence result given by Theorem 2.4 implies that as long as it is the producer who bears the production cost, however active the data broker is in the product market does not affect market outcomes at all. This means that, on the one hand, the data broker has no incentive to become even more active in the product market rather than only 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, even if the data broker does become more active in the product market, it still raises no further concerns to the policymaker. Thus, any policy intervention that prohibits the data broker entering the product market by either gaining control over prices (e.g., by establishing an online platform and allows the producer to trade on this platform while controlling the prices) or obtaining the exclusive right to (re)-sell the product would be unnecessary.

2.7 Conclusion

In sum, this chapter restricts the set of feasible information structures for the informational monopolist to a set where the buyer is always fulling informed. This model is the used to

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

study a scenario where a data broker sells consumer data and creates market segmentations. In this chapter, I characterize the optimal mechanisms of the data broker and 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 extensions are also considered, including the case in which consumers possess some private information that cannot be disclosed and the environment where targeted marketing is available.

Several aspects can become future research topics. First, although one of the extensions of this chapter considers a scenario where targeted marketing is possible, it abstracts away from the possibility that a data broker can generate "match values" between the producers and consumers. By assuming that every group of consumers can buy every product as long as they see it, the matching aspect between consumers and producers is omitted. After all, there is effectively no competition among the producers when there is no "scarcity" of consumers. Furthermore, the consumers' characteristics that govern the match values can also be their private information. Second, although one of the extensions consider the case where the consumers can preserve some private information, it is restricted to certain environments. A natural direction of future research is to explore the data broker's optimal mechanisms and their implications in a setting where the feasible market segmentation is restricted by a Blackwell upper bound. Lastly, the producer is assumed to be a product monopoly in this chapter. It would be economically relevant to explore the consequences of consumer-data brokership under different industrial structures.

CHAPTER 3

UNINFORMED SELLER: SALE OF ADVERTISEMENT

3.1 Introduction

3.1.1 Preface

This chapter is constituted by the main results of Yang (2019b). In this chapter, I restrict the set of feasible information structures to $\mathrm{X}_{\mathrm{Ch3}} \subseteq \mathrm{X}$, where the seller is always completely uninformed about v.[1] Also, I specialize the seller's outside option π_0 to zero. Main examples of this specification include advertisement agencies or online platforms that help the seller provide information about the product to the buyer and charge the seller for such services. For concreteness, I use the sale of advertisement as the leading example throughout this chapter. As such, the informational monopolist is referred as the *intermediary* throughout this chapter.

As in Chapter 2, the set of feasible information structures X_{CH3} has a simple representation. To see this, recall that for the buyer, under any information structure $\chi \in \mathrm{X}_{\mathrm{Ch3}}$, given any signal realization $s_{\mathfrak{B}} \in S_{\mathfrak{B}}$ and any posted price p, he must buy the product if $\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}] \geq p$ and would not buy if $\mathbb{E}_{(\chi,m^0)}[v|s_{\mathfrak{B}}] < p$. In other words, the only payoff-relevant statistic of information structure $\chi \in \mathrm{X}_{\mathrm{Ch3}}$ is the interim expected value. As a result, the set of feasible information structures X_{CH3} can be represented by the collection of distributions on V that are mean-preserving contractions of m^0, as in Gentzkow and Kamenica (2016).[2]

1. Specifically, fix any $\underline{s} \in S_{\mathfrak{S}}$. The set of feasible information structures in this chapter is defined as

$$\mathrm{X}_{\mathrm{Ch3}} := \{\chi \in \mathrm{X} : \chi(S_{\mathfrak{B}} \times \{\underline{s}\})|v) = 1, \forall v \in V\}.$$

2. Similar characterizations appear in many recent developments in the literature of mechanism and information design, see for instance Neeman (2003), Bergemann and Pesendorfer (2007), Shi (2012), Roesler

Specifically, recall that $D_0(p) := m^0([p, \overline{v}])$, and that $\mathcal{D}$ is the collection of nondecreasing, left-continuous functions $D : \mathbb{R}_+ \to [0, 1]$ with $D(\underline{v}) = 1$, $D(\overline{v}^+) = 0$. For any $D \in \mathcal{D}$, let

$$I_D(p) := \int_p^\infty D(v)\,\mathrm{d}v,$$

for all $p \geq 0$. The set of feasible information structures $\mathrm{X}_{\mathrm{Ch3}}$ can be represented as follows:

$$\mathcal{D}_0 := \{D \in \mathcal{D} : I_D(p) \leq I_{D_0}(p),\ \forall p \in \mathbb{R}_+,\ \text{with equality at } p = 0\}$$

which can be illustrated by Figure 3.1, where $\underline{D}$ denotes the distribution that assigns probability 1 at $\mathbb{E}_{m^0}[v]$.

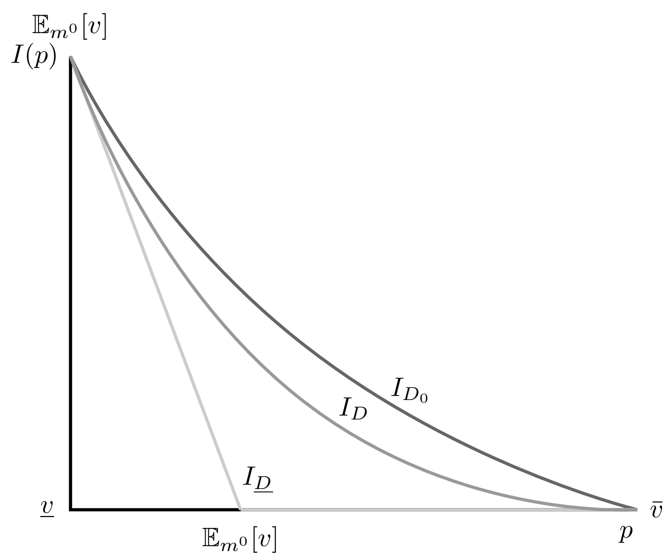

Figure 3.1: Feasible Set $\mathcal{D}_0$.

---

and Szentes (2017), Kolotilin et al. (2017), Bergemann and Morris (2017), Du (2018), Dworczak and Martini (2019), Brooks and Du (2019). Skreta and Perez-Richet (2018).

3.1.2 Motivation

The information about a product that is provided to consumers plays a central role in defining profits and consumer surplus. In many real-world contexts, the channel through which consumers receive information relies on a third party, an *intermediary*, who possesses the technology to provide information about the products to the consumers. Consumers can be informed about the presence of a particular product and can further learn about its features via the intermediary.

For instance, online selling platforms such eBay have a well-publicized website on which the commodities are presented to the consumers and different information about the commodities can be provided via photos, certificates and descriptions of the product; Advertisement agencies such as QVC also have well-known broadcasting channels through which the producers can present their products and information about the commodity can be provided; Car dealers know a great deal about the cars they are selling and can provide different information to the consumers. These intermediaries can exploit such technology and extract surplus from the product market. Among these intermediaries, the business models they adopt exhibit great variations. Some of them (e.g., QVC) only provide informational services to the seller by broadcasting advertisements and then charge the seller for this service (regime $\mathcal{I}$); whereas some of them (e.g., car dealers and online platforms) are more active in the product market—they would buy the product first from the producer and then sell to the consumers directly while providing information about this product (regime $\mathcal{P}$).

Given the importance of consumers' information and the rapid growth of third party intermediaries, it is thus of great importance to understand how such intermediaries could exploit their information technology and how various business models (market regimes) would affect market outcomes. In this chapter, I explore the above questions. Specifically, I consider a monopolistic pricing model with one buyer and one seller, where the seller can produce a product with a nonegative production cost. The buyer does not know their value for the product or the presence of the product *a priori*, but can interact with the seller and learn about the product through an intermediary. More specifically, there is an intermediary who can inform the buyer about the seller's product. The goal of the intermediary is to maximize revenue by exploiting this technology. The main result of this chapter is a complete characterization of the revenue-maximizing mechanisms for the intermediary under regime $\mathcal{I}$ and regime $\mathcal{P}$, respectively.

The main result leads to the outcome-equivalence result—consumer surplus, producer profit, and the intermediary's revenue are the same under any optimal mechanism of both regimes. As such, impacts of the intermediaries on the market are independent of market regimes. Furthermore, I also show that even under the most extreme form of targeting—the intermediary knows exactly the consumers' values and can target them accordingly—the market outcomes are still the same and do not depend on the market regimes. Finally, using the equivalence result, I further provide welfare analyses and comparative statics that are valid for both market regimes. I show that under both market regimes, total surplus and volume of trade when an intermediary has the informational technology are less than a benchmark case when the seller has control of the informational technology directly. Also, I show that under both regimes, total surplus and the seller's profit are larger when the buyer's value is higher and when information rent of the seller is smaller.

The rest of this chapter is organized as follows: The next section summarizes the related literature. Section 3.2 introduces the model. Section 3.3 characterizes the revenue maximizing mechanisms for the intermediary under both regime $\mathcal{P}$ and regime $\mathcal{I}$ and then establishes the outcome-equivalence result. In Section 3.4, targeting technology is introduced and the equivalence result with targeting is established. In Section 3.5, comparative statics and welfare analyses are provided. Section 3.6 contains discussions of the results and Section 3.7 concludes. The proofs of the results can be found in the appendix.

3.1.3 Related Literature

This chapter is related to several branches of literature in interplay between monopolistic pricing and information structure, selling information, and Bayesian persuasion. In the monopolistic pricing literature, Lewis and Sappington (1991) also examines how the change in consumer's information affects a monopolist's profit. They show that for a given monopolist with a constant production cost, the buyer having either full information or no information is optimal for the monopolist. The model in this chapter differs from theirs in two major aspects. First, I examine how a third party would price information structures for a monopolist to purchase, instead of examining optimal information structure from a monopolist's perspective directly. Second, although the setting of Lewis and Sappington (1991) is close to a benchmark of the model here in which the monopolist can choose information structure that the buyer has directly, the model in this chapter maintains an assumption that the commodity is indivisible so that buyers have 0-1 demand while the consumer's demand in their model can be more general. On the other hand, Lewis and Sappington (1991) restrict the information structures to vary within a one-dimensional family by assuming a particular disclosure rule, whereas the model here allows full flexibility of the choice of information structure. Johnson and Myatt (2006) also studies how change in the distribution of buyers' value, which is equivalent to the information that the buyer has in the model here, affects a monopolist's profit. In particular, they show that when the distributions are ordered by the rotational ordering, the monopolist will prefer two extremes of the order. Using the language of buyers' information, this result is similar to Lewis and Sappington (1991) in that it implies that under a particular one dimensional (and hence, totally-ordered) family of information structures, a monopolist will either prefer the most informative one or the least informative one. Again, the model here differs from theirs in that it focuses on a third party's surplus extraction and that complete flexibility in providing information structures is allowed. Recent developments, on the other hand, have adopted flexible information structures. Bergemann et al. (2015) characterizes all the possible surplus division that can arise by giving the monopolist different information about the buyer's value. Roesler and Szentes (2017) examines the buyer-optimal information structure when facing a monopoly.

There are several works that also study a problem of pricing information. Bergemann and Bonatti (2015) studies a pricing problem of a data provider who can provide information about the match value for a seller whose profit from trade depends on the match value of each consumer and the amount of investments the seller makes in each consumer. Bergemann et al. (2018) solves an optimal menu for a data provider to sell experiments to a decision maker who has a private estimate about the state. The model in this chapter differs from the theirs in that the intermediary in this model sells information to affect the information of the *buyer* who is buying a product produced by the *seller*, which affects the seller's value function of different information indirectly, whereas the model in Bergemann et al. (2018) focuses on selling information structure to a decision maker whose value function depends on the information she purchases directly. Segura-Rodriguez (2019) considers a model where a data broker sells information for forecasting a multidimensional variable to firms who differ in the dimensions that are payoff-relevant. Yang (2020) studies the revenue-maximizing mechanisms for a data broker who sells data about the consumers value to a producer who has private information about her production cost and derive the economic implications. Although Yang (2020) also establishes an outcome equivalent result, the reasons and the techniques involved are distinct from this chapter, since the information being sold in this chapter is to inform the buyer instead of the seller.

Furthermore, the screening framework in the model for selling information structure is analogous to standard monopolistic screening problems as in Mussa and Rosen (1978), Myerson (1981) and Maskin and Riley (1984). The screening problem in the model here is more complicated in that the outcome space is infinite dimensional. Moreover, under regime $\mathcal{I}$, the revenue maximizing problem is a mixture of screening and moral hazard problem from the intermediary's perspective. Also, the assumption that intermediary is able to commit to a mechanism and the characterization of information structure follows from the Bayesian persuasion literature, as Kamenica and Gentzkow (2011), Gentzkow and Kamenica (2016).

3.2 Model

In this section, I specialize the model established in Chapter 1 by restricting the set of feasible information structures to $\mathrm{X}_{\mathrm{Ch3}}$ and specifying the seller's outside option as $\pi_0(c) = 0$ for all $c \in C$. As discussed in Section 3.1, the set of feasible information structures $\mathrm{X}_{\mathrm{Ch3}}$ can be represented by the set $\mathcal{D}_0$. As a result, a mechanism for the intermediary becomes a tuple $(\alpha, \boldsymbol{D}, \tau)$ under regime $\mathcal{I}$ and a tuple $(\boldsymbol{D}, \tau, \boldsymbol{\gamma})$ under regime $\mathcal{P}$,[3] where $\alpha : C \to \{0, 1\}$ a measurable function that maps each report c to a publicizing decision $\alpha(c) \in \{0, 1\}$ that denotes whether the seller's product will be seen by the buyer; $\boldsymbol{D} : C \to \mathcal{D}_0$ is a measurable function that maps each report c to an information structure $\boldsymbol{D}(c) \in \mathcal{D}_0$; $\tau : C \to \mathbb{R}$ is a measurable function that maps reported cost c to the amount of payment $\tau(c)$ made by the seller; and $\boldsymbol{\gamma} : C \to \Delta(\mathbb{R}_+)$ is a measurable function that specifies, for each report c, a stochastic price $\boldsymbol{\gamma}(c)$ that should be charged to the buyer. For simplicity, throughout this chapter, I assume that both m^0 and G are regular and that the sufficient condition for Assumption 2.1, (2.7), holds.[4]

Under these specifications, under regime $\mathcal{I}$, the incentive compatibility constraints (IC-I) and individual rationality constraints (IR-I) becomes

$$\alpha(c) \cdot \max_{p \geq 0} (p - c)\boldsymbol{D}(p|c) - \tau(c) \geq \alpha(c') \cdot \max_{p \geq 0} (p - c)\boldsymbol{D}(p|c') - \tau(c'),$$

3. Under regime $\mathcal{P}$, the intermediary can exclude the seller by contracting on price zero. Thus, the publicizing decision α is not needed when specifying a mechanism under regime $\mathcal{P}$.

4. These assumptions are not necessary, and the relaxations can be found in the appendix of Yang (2019b).

for all $c, c' \in C$ and

$$\alpha(c)\max_{p\geq 0}(p-c)\boldsymbol{D}(p|c)-\tau(c)\geq 0,$$

for all $c \in C$, respectively. Similarly, under regime $\mathcal{P}$, (IC-P) and (IR-P) becomes

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

for all $c, c' \in C$ and

$$\int_{\mathbb{R}_+}(p-c)\boldsymbol{D}(p|c)\boldsymbol{\gamma}(\mathrm{d}p|c)-\tau(c)\geq 0,$$

for all $c \in C$, respectively.

As in Chapter 2, compared with classical nonlinear pricing problems, the allocation space in the intermediary's revenue-maximization problem is infinite dimensional. After all, the allocation involves the information provided to the buyer, which belongs to a rich set $\mathcal{D}_0$. In the next chapter, I will characterize the revenue-maximizing mechanisms for the intermediary under both regimes, which will then lead to the outcome-equivalence result.

3.3 Optimal Mechanisms and Outcome-Equivalence

3.3.1 Optimal Mechanisms under Regime $\mathcal{P}$

As in Chapter 2, under regime $\mathcal{P}$, the intermediary's revenue maximization problem can in fact be reduced to a standard screening problem with one-dimensional allocation space. To see this, first notice that as in classical mechanism design problems, the transfer can pinned down by allocations up to a constant due to local incentive constraints, and a monotonicity condition would ensure global incentive compatibility. This is formally stated in Lemma 3.1.

Lemma 3.1. *Under regime $\mathcal{P}$, a mechanism $(\boldsymbol{D}, \tau, \boldsymbol{\gamma})$ is incentive compatible if and only if:*

1. *There exists* $\bar{\tau} \in \mathbb{R}$ *such that*

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

for all $c \in C$.

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

Using the characterization given by Lemma 3.1, the revenue maximization problem under regime $\mathcal{P}$ can be rewritten as (recall that $\phi_G(c) := c + G(c)/g(c)$ is the virtual cost):

$$\begin{aligned} &\sup_{\boldsymbol{\gamma},\boldsymbol{D}} \int_C \left(\int_{\mathbb{R}_+} (p - \phi_G(c))\boldsymbol{D}(p|c)\boldsymbol{\gamma}(\mathrm{d}p|c) \right) G(\mathrm{d}c) \qquad (3.1) \\ &\text{s.t. } c \mapsto \int_{\mathbb{R}_+} \boldsymbol{D}(p|c)\boldsymbol{\gamma}(\mathrm{d}p|c) \text{ is nonincreasing.} \end{aligned}$$

Thus, from Lemma 3.1, the intermediary's problem under regime $\mathcal{P}$ is essentially a standard nonlinear pricing problem with quasi-linear preference. Therefore, under the regularity assumptions, (3.1) can be solved by pointwise maximization. To construct a solution, consider any $c \in C$ and notice that maximizing the integrand of (3.1) evaluated at c is equivalent to maximizing the seller's profit by choosing a stochastic price $\boldsymbol{\gamma}(c)$ and an information structure $\boldsymbol{D}(c) \in \mathcal{D}_0$ when her cost is $\phi_G(c)$. Given any information structure $\boldsymbol{D}(c) \in \mathcal{D}_0$, according to Myerson (1981) and Maskin and Zeckhauser (1983), it is without loss to use deterministic prices. Furthermore, notice that the seller's profit must be bounded from above by the efficient surplus in the economy where the production cost is $\phi_G(c)$. That is,

$$(p - \phi_G(c))D(p) \le \int_{\phi_G(c)}^{\infty} (v - \phi_G(c))D_0(\mathrm{d}v) = \int_{\psi(c)}^{\infty} D_0(v)\,\mathrm{d}v = I_{D_0}(\phi_G(c)),$$

for all $p \ge 0$ and for all $D \in \mathcal{D}_0$, which can be also be seen geometrically from Figure 3.2.

Now, for any $c \in C$, consider the information structure $\boldsymbol{D}_{\mathrm{b}}(c) \in \mathcal{D}_0$ that informs the

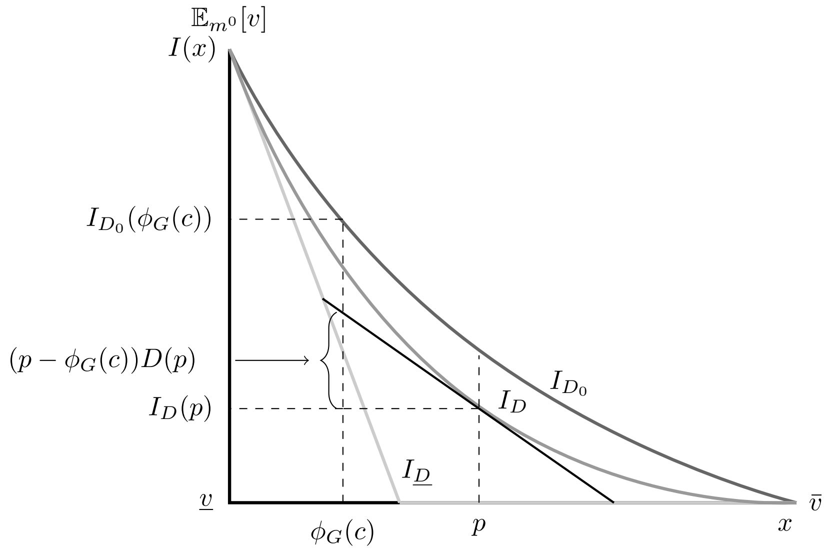

Figure 3.2: Pointwise Upper Bound

buyer whether their value is above or below $\phi_G(c)$.[5] Together with a (deterministic) price equal to the buyer's interim expected value conditional on knowing that their value is above $\phi_G(c)$, $v(c) := \mathbb{E}_{m^0}[v|v > \phi_G(c)]$, the seller's profit is

$$(v(c) - \phi_G(c))\boldsymbol{D}_{\mathrm{b}}(v(c)|c) = (\mathbb{E}_{m^0}[v|v > \phi_G(c)] - \phi_G(c))D_0(\phi_G(c)) = I_{D_0}(\phi_G(c)),$$

as illustrated by Figure 3.3.

As a result, for any $c \in C$, the seller's profit attains the upper bound $I_{D_0}(\phi_G(c))$ under price $v(c)$ and information structure $\boldsymbol{D}_{\mathrm{b}}(c)$. Furthermore, the induced trading probability is $\boldsymbol{D}_{\mathrm{b}}(v(c)|c) = D_0(\phi_G(c))$, which is nonincresing since ϕ_G is increasing. Thus, for any $c \in C$, let $\boldsymbol{\gamma}_{\mathrm{b}}(c)$ be the dirac measure on $v(c)$ and let $\tau_{\mathrm{b}}(c)$ be defined by $\boldsymbol{D}_{\mathrm{b}}$ and $\boldsymbol{\gamma}_{\mathrm{b}}$ using condition 1 of Lemma 3.1 with $\bar{\tau}$ being zero. Then, by Lemma 3.1, the revenue maximization problem

5. Formally, $\boldsymbol{D}_{\mathrm{b}}(c)$ is defined as

$$\boldsymbol{D}_{\mathrm{b}}(p|c) := \begin{cases} 0, & \text{if } p \in [0, \mathbb{E}_{m^0}[v|v \leq \phi_G(c)]] \\ D_0(\phi_G(c)), & \text{if } p \in (\mathbb{E}_{m^0}[v|v \leq \phi_G(c)], \mathbb{E}_{m^0}[v|v > \phi_G(c)]] \\ 1, & \text{if } p \in (\mathbb{E}_{m^0}[v|v > \phi_G(c)], \infty) \end{cases}.$$

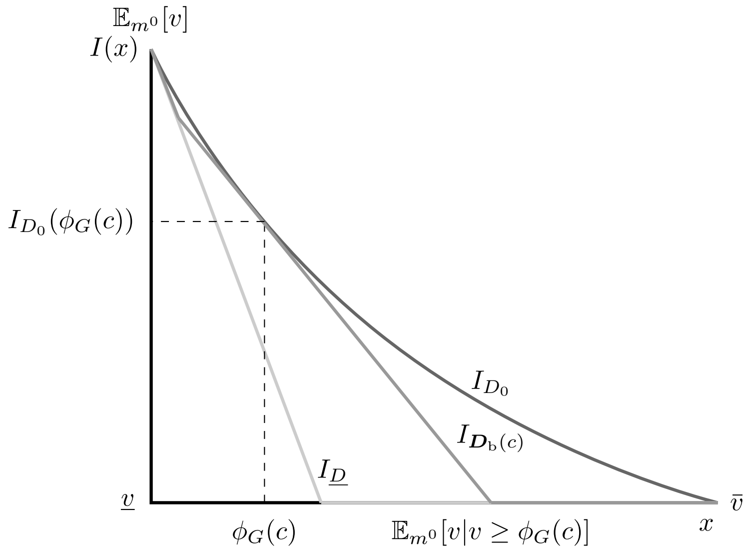

Figure 3.3: Profit under $p = \mathbb{E}_{m^0}[v|v \geq \phi_G(c)]$ and $\boldsymbol{D}_{\mathrm{b}}(c)$

under the regime $\mathcal{P}$ can be completely solved, which leads to Theorem 3.1 below.

Theorem 3.1. *Under regime $\mathcal{P}$, the mechanism $(\boldsymbol{D}_{\mathrm{b}}, \tau_{\mathrm{b}}, \boldsymbol{\gamma}_{\mathrm{b}})$ is optimal and the optimal revenue for the intermediary is:*

$$R^* := \int_C \left(\int_{\phi_G(c)}^{\infty} D_0(v)\,\mathrm{d}v \right) G(\mathrm{d}c).$$

As demonstrated above, when the intermediary has control over prices, a combination of tools in the mechanism design literature and the information design literature turns out to be extremely useful in solving this problem. Contractability of prices enables the intermediary to discipline the seller well so that the revenue maximization problem becomes separable. However, under regime $\mathcal{I}$, the intermediary does not have control over prices and hence the revenue maximization problem becomes more complicated. Nevertheless, the solution under regime $\mathcal{P}$ would be a useful benchmark for this more complex problem, which will be examined in the next section.

3.3.2 Optimal Mechanisms under Regime $\mathcal{I}$

The intermediary's problem under regime $\mathcal{I}$ is substantially different from that under regime $\mathcal{P}$. Specifically, under regime $\mathcal{I}$, the intermediary fully delegates the pricing decision to the seller and only provides the "informational service" by publicizing the product and providing information to the buyer. The intermediary has no control over the sale and the seller retains full control over prices. The inability to contract on price, together with the infinite-dimensional allocation space, leads to a mixture problem between moral hazard and screening. As such, the solution constructed above may not be feasible due to the possibility of double deviations and thus the insufficiency of local incentive constraints.

To show that regime $\mathcal{P}$ and regime $\mathcal{I}$ are outcome-equivalent, I first solve for the optimal mechanism in closed form under regime $\mathcal{I}$ (Theorem 3.2). This solution would then imply that the intermediary's revenue is also R^*, which in turn implies that the buyer's surplus and the seller's profit are also the same as those under regime $\mathcal{P}$.

To begin with, first notice that even though the allocation space is infinite-dimensional, there is still a revenue equivalence formula that characterizes incentive compatible mechanisms. Specifically, by the envelope theorem, transfer in any incentive compatible mechanism is pinned down by the information structure and the publicizing policy up to a constant. The difference is that because of the infinite-dimensional allocation space and the possibility of double deviations, monotonicity is no longer sufficient for incentive compatibility and the entire integral constraint should be kept tracked. To formally state the result, recall from Section 2.2 of Chapter 2 that for any $c \in C$ and for any $D \in \mathcal{D}$, $\pi_D(c) := \max_{p \geq 0}(p-c)D(p)$, $\boldsymbol{P}_D(c) := \mathrm{argmax}_{p\geq 0}(p-c)D(p)$ and $\overline{\boldsymbol{p}}_D(c)$ is the largest element of $\boldsymbol{P}_D(c)$.

Lemma 3.2. *Under regime $\mathcal{I}$, a mechanism $(\alpha, \boldsymbol{D}, \tau)$ is incentive compatible if and only if*

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

$$\tau(c) = \alpha(c) \cdot \pi_{\boldsymbol{D}(c)}(c) - \int_c^{\overline{c}} \alpha(z)\boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(z)}(z)|z)\,\mathrm{d}z + \bar{\tau}.$$

2. For any $c, c' \in C$,

$$\int_c^{c'} [\alpha(z)\boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(z)}(z)|z) - \alpha(c')\boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c')}(z)|c')]\,\mathrm{d}z \geq 0.$$

Furthermore, the "only if" part holds for any $\boldsymbol{p} \in \boldsymbol{P}$.

By Lemma 3.2 and individual rationality, after interchanging the order of integrals, it follows that optimal mechanisms can be found by solving the following maximization problem:

$$\sup_{\boldsymbol{D},\alpha} \int_C \alpha(c) \left((\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(c) - \phi_G(c))\boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(c)|c) \right) G(\mathrm{d}c) \tag{3.2}$$
$$\text{s.t. } \int_c^{c'} [\alpha(z)\boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(z)}(z)|z) - \alpha(c')\boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c')}(z)|c')]\,\mathrm{d}z \geq 0,\ \forall c, c' \in C.$$

In what follows, I will show that the value of (3.2) is exactly R^*, which is the same as that under regime $\mathcal{P}$. To this end, I construct explicitly an optimal mechanism for the intermediary. This optimal mechanism features *upper censorship*, in the sense that for each reported cost c, the intermediary will provide an information structure so that whenever the buyer's value is below a certain report-dependent cutoff, they learns exactly their value, whereas when the buyer's value is above the cutoff, they learns nothing else than the fact that their value is above the cutoff. More specifically, an upper censorship information structure is defined below.

Definition 3.1. An information structure $D \in \mathcal{D}_0$ is an *upper censorship* with cutoff $\kappa \geq 0$ if

$$D(p) = \begin{cases} D_0(p), & \text{if } p \in [0, \kappa] \\ D_0(\kappa), & \text{if } x \in (\kappa, \mathbb{E}_{m^0}[v|v > \kappa] \\ 1, & \text{if } x \in (\mathbb{E}_{m^0}[v|v > \kappa], \infty). \end{cases}, \forall p \in \mathbb{R}_+.$$

With the definition above, the solution for the intermediary's problem under regime $\mathcal{I}$

can be stated as Theorem 3.2 below.

Theorem 3.2. *Under regime $\mathcal{I}$, for each $c \in C$, let $\boldsymbol{D}_{\mathrm{u}}(c)$ be an upper censorship with cutoff $\phi_G(c)$, let $\alpha_{\mathrm{u}}(c) = 1$ and let*

$$\tau_{\mathrm{u}}(c) := (v(c) - c)D_0(\phi_G(c)) - \int_c^{\overline{c}} D_0(\phi_G(z))\,\mathrm{d}z.$$

Then $(\alpha_{\mathrm{u}}, \boldsymbol{D}_{\mathrm{u}}, \tau_{\mathrm{u}})$ is optimal for the intermediary. Moreover, the optimal revenue is R^*.

In other words, under regime $\mathcal{I}$, the intermediary can still achieve the same revenue as that under regime $\mathcal{P}$ even without the ability to contract on price by revealing all the information to the buyer when their value is is below $\phi_G(c)$, as illustrated by Figure 3.4 (where the dotted line depicts the optimal binary disclosure rule under regime $\mathcal{P}$). The formal proof of Theorem 3.2 is in the appendix. To see the intuition behind the proof, recall that the main distinction between regime $\mathcal{I}$ and regime $\mathcal{P}$ is that the intermediary cannot contract on price under regime $\mathcal{I}$, which leads to additional constraints that are associated with the seller's pricing incentive and double deviation incentives. The presence of these extra constraints complicates the revenue maximization problem. Indeed, as price is not contractable under regime $\mathcal{I}$, the mechanism consisting of binary disclosure information structures may not be optimal since it may not induce a truthfully-reporting seller with cost c to set price at $v(c)$. Furthermore, there might be incentives for double deviations—a seller with cost c may have incentive to misreport a cost c' and then set a price distinct from $v(c')$ under the binary disclosure information structure.

The property that the buyer learns exactly their value whenever it is below $\phi_G(c)$ under an upper censorship (with cutoff $\phi_G(c)$), however, deters the aforementioned deviations given that (2.7) holds. To see this, consider any $c \in C$. Regularity of m^0 and (2.7) imply that any price $p < \phi_G(c)$ yields a lower profit for the seller with cost c than the price $\phi_G(c)$, as $\phi_G(c)$ is below the optimal monopoly price under (2.7) and the profit function is singled-peaked

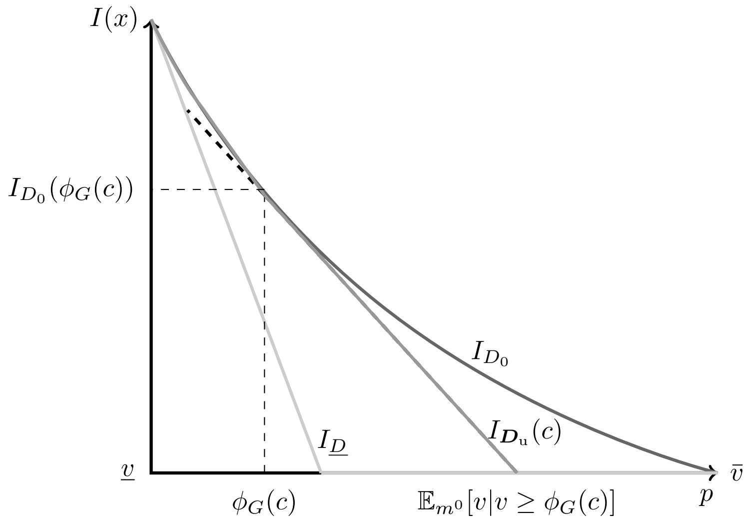

Figure 3.4: Upper Censorship $\boldsymbol{D}_{\mathrm{u}}(c)$ with Cutoff $\phi_G(c)$.

by regularity. This implies that charging price $v(c)$ would be better than charging price $\phi_G(c)$ under $\boldsymbol{D}_{\mathrm{u}}(c)$, since the probability of trade would be the same but $v(c) > \phi_G(c)$. As a result, a truthfully reporting seller will optimally set a price at $v(c)$ under upper censorship $\boldsymbol{D}_{\mathrm{u}}(c)$. Furthermore, under (2.7), the second condition in Lemma 3.2 is satisfied under this mechanism. Thus, the mechanism $(\alpha_{\mathrm{u}}, \boldsymbol{D}_{\mathrm{u}}, \tau_{\mathrm{u}})$ is incentive feasible and generates revenue R^*, as

$$(\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c) - \phi_G(c))\boldsymbol{D}_{\mathrm{u}}((\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c)|c) = (v(c) - \phi_G(c))D_0(\phi_G(c)) = I_{D_0}(\phi_G(c)),$$

for all $c \in C$. Since R^* is an upper bound for the intermediary's revenue under regime $\mathcal{I}$, this implies that $(\alpha_{\mathrm{u}}, \boldsymbol{D}_{\mathrm{u}}, \tau_{\mathrm{u}})$ is optimal and the optimal revenue is R^*, which establishes Theorem 3.2.

On a higher level, the proof of Theorem 3.2 exploits the richness of information structures to accommodate additional incentives for the seller when the selling decision is fully delegated, just as the proof of Theorem 2.2 in Chapter 2. Since the intermediary is allowed to use any information structure and the set of feasible information structures, $\mathcal{D}_0$, is rich, the intermediary can tailor the information structures in order to compensate the lack of ability to contract on prices. Any mechanism that attains R^* has to provide correct incentives for the seller to set the desirable price and eliminate double deviation concerns. Under (2.7), the mechanism consisting of upper censorships achieves this goal. When (2.7) does not hold, solving (3.2) would be more difficult and would entail more complex arguments than pointwise maximization. I now outline the solution for (3.2) without (2.7) below. In the appendix of Yang (2019b), I fully solve the intermediary's revenue maximization problem in this case.

As a benchmark, the optimal mechanism given by Theorem 3.2 has following features:

1. Given any report $c \in C$, the buyer's value is *fully* disclosed when it is below $\phi_G(c)$ and is completely not disclosed otherwise.

2. The product is publicized for all reported cost.

3. For any seller with cost $c \in C$, truth-telling and setting a price at $\mathbb{E}_{m^0}[v|v \geq \phi_G(c)]$ is optimal.

In contrast, the optimal mechanism without (2.7) constructed in the appendix of Yang (2019b) differs from the optimal mechanism given by Theorem 3.2 in following ways: There exists $c^* \geq 0$, which only depends on m^0 and G such that

1. Given any report $c \in C$, the buyer's value is *partially* disclosed when it is below $\phi_G^*(c) := \{\phi_G(c), \phi_G(c^*)\}$ and is completely not disclosed otherwise.

2. The product is publicized for reported cost below a threshold $\hat{c} \geq c^*$ and is not publicized otherwise.

3. For any seller with cost $c \in C$, truth-telling is optimal. For any seller with cost $c \leq \hat{c}$, setting a price at $\mathbb{E}_{m^0}[v|v \geq \phi_G^*(c)]$ is optimal.

3.3.3 Outcome-Equivalence

From the characterizations of the optimal mechanisms under both regime $\mathcal{I}$ and regime $\mathcal{P}$, the outcome-equivalence result then follows immediately. Indeed, according to Theorem 3.1 and Theorem 3.2, the intermediary's optimal revenue is R^* under both regime $\mathcal{I}$ and regime $\mathcal{P}$. Furthermore, since R^* is the total expected surplus in the economy where the seller's cost is replaced by the virtual cost, the only way to achieve R^* is for the buyer to buy whenever their value is above $\phi_G(c)$ by paying $v(c)$ and not to buy otherwise, for (almost) all realized cost $c \in C$. In other words, even though there might be multiple optimal mechanisms for the intermediary under both regimes, the outcome induced by the optimal mechanisms is unique. This leads to the outcome-equivalence result.

Theorem 3.3 (Outcome-Equivalence). *Regime $\mathcal{I}$ and regime $\mathcal{P}$ are outcome-equivalent (under $\mathrm{X}_{\mathrm{Ch3}}$).*

As in Chapter 2, this outcome-equivalence result means that the ability for the intermediary to contract on price is irrelevant, provided that he can design information provided to the buyer arbitrarily. Thus, from the intermediary's perspective, providing only informational services without involving in the product market is equally profitable compared with obtaining the ability to contract on price. Meanwhile, from the seller's perspective, preserving the right to set prices does not help improving her expected profit. Finally, for a policymaker, it does not matter which market regime the intermediary is operating under. Intermediaries that involve in product markets are the same as intermediaries that only provide informational services.

3.4 Irrelevance of Targeting

Previous analyses rely on a premise that the intermediary has to provide the same information structure and make the same publicizing decision regardless of the value of the buyer. However, it is reasonable to argue that in practice, the intermediary can do more. In some situations, the intermediary may have the technology to estimate the buyer's value—by collecting consumers' browsing history, personal information and purchasing behavior. Knowing the buyer's value, the intermediary can then design different publicizing and disclosure policy to different consumers in order to create more surplus. This section considers an extension where the intermediary is able to *target* buyers with different values.

For the ease of exposition, consider an alternative interpretation of the environment as adopted in Chapter 2, where there is a unit mass of consumers that have unit demand and heterogeneous values. The distribution of buyers' values is given by m^0 that has full support on V. Under this interpretation, availability of targeting means that the intermediary, when disclosing information, can select any (measurable) subset of buyers $A \subseteq V$, publicize the product to these consumers and provide arbitrary information to these buyers. The buyers who are not selected, $V \backslash A$, remain unaware of the product. As such, a *targeting policy* can be described by a measurable subset $A \subseteq V$. Thus, if the intermediary targets a group A, the prior distribution of values effectively becomes:

$$m^A = m^0|_A + \delta_{\{0\}} \cdot m^0(V \backslash A).$$

That is, the consumers who are not in the target set A are effectively the consumers with zero values. For any targeting set A, let $D_A := m^A([p, \overline{v}])$ for all $p \in \mathbb{R}_+$. By the same arguments as it were used when targeting is not available, together with an additional requirement that whoever is not targeted cannot receive any information, the collection of information structures can then be characterized by the set

$$\mathcal{D}_A := \{D \in \mathcal{D} : D(\underline{v}) = D_A(\underline{v}),\ I_D(p) \leq I_{D_A}(p),\ \forall p \geq 0,\ I_D(\underline{v}) = I_{D_A}(\underline{v})\}.$$

As such, under regime $\mathcal{P}$, a mechanism is given by $(\boldsymbol{A}, \boldsymbol{D}, \tau, \boldsymbol{\gamma})$; Under regime $\mathcal{I}$, a mechanism is given by $(\boldsymbol{A}, \boldsymbol{D}, \tau)$, where $\boldsymbol{A}$ maps from each reported cost c to a measurable set $\boldsymbol{A}(c) \subseteq V$

As in the previous section, the revenue-equivalence formulae are still valid and Lemma 3.1 and Lemma 3.2 can be revised by requiring that $\boldsymbol{D}(c) \in \mathcal{D}_{\boldsymbol{A}(c)}$ for all $c \in C$. Therefore, the revenue maximization problem under regime $\mathcal{P}$ can still be written as:

$$\begin{aligned} \sup_{\boldsymbol{A},\boldsymbol{D},\boldsymbol{\gamma}} & \int_C \left(\int_{\mathbb{R}_+} (p - \phi_G(c)) \boldsymbol{D}(p|c) \boldsymbol{\gamma}(\mathrm{d}p|c) \right) G(\mathrm{d}c) \\ \text{s.t. } & c \mapsto \int_{\mathbb{R}_+} \boldsymbol{D}(p|c) \boldsymbol{\gamma}(\mathrm{d}p|c) \text{ is nonincreasing,} \end{aligned} \tag{3.3}$$

and the revenue maximization problem under regime $\mathcal{I}$ can be written as:

$$\begin{aligned} \sup_{\boldsymbol{A},\boldsymbol{D}} & \int_C \left((\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(c) - \phi_G(c)) \boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(c)|c) \right) G(\mathrm{d}c) \\ \text{s.t. } & \int_c^{c'} [\boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(z)}(z)|z) - \boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c')}(z)|c')] \, \mathrm{d}z \geq 0, \ \forall c, c' \in C. \end{aligned} \tag{3.4}$$

With the additional leverage of targeting, it is clear that the intermediary can (weakly) improve its revenue under both regimes.[6] Below, I show that even with targeting, the intermediary cannot do better than generating revenue R^*. As such, the outcome equivalent result can be strengthened so that regardless of the market regimes *and* regardless of the availability of targeting technology, the induced outcomes are equivalent.

To see this, first notice that for any measurable set $A \subseteq V$, m^0 first-order stochastic dominates m^A and therefore

$$\int_p^\infty D_A(v) \, \mathrm{d}v \leq \int_p^\infty D_0(v) \, \mathrm{d}v,$$

for any $p \in \mathbb{R}_+$ and thus any prior after targeting must induce a convex function I_{D_A} that

6. To see this, under regime $\mathcal{P}$, for each $c \in C$, let $\boldsymbol{A}(c) = V$. Then any contract that is available when there is no targeting is feasible. On the other hand, under regime $\mathcal{I}$, for each $c \in C$, mechanisms with $\boldsymbol{A}(c) = V$ for all c are equivalent to mechanisms with $\alpha(c) = 1$, while mechanisms with $\boldsymbol{A}(c) = \emptyset$ are equivalent to mechanisms with $\alpha(c) = 0$.

is always below I_{D_0}. By using the same arguments as in the proof of Theorem 3.1 and Theorem 3.2, R^* is still an upper bound of the objective of (3.3) and (3.4). Therefore, since the optimal mechanisms without targeting generate revenue R^*, the intermediary's optimal revenue with targeting under both regimes must also be R^*. Moreover, since R^* can only be attained by selling to the buyer if and only if $v \geq \phi_G(c)$ at price $v(c)$, the rest of the market outcomes must also be the same.

Theorem 3.4. *Suppose that targeting is available. Then regime $\mathcal{I}$ and regime $\mathcal{P}$ are outcome-equivalent.*

3.5 Welfare Analysis and Comparative Statics

With the outcome-equivalence result, several comparative statics and welfare analyses independent of market regimes can be made. This section records such results.

Proposition 3.1 (Welfare Comparison)**.** *The expected total surplus generated by trade and the probability of efficient trade are larger when the seller has control of the information technology than when the intermediary has control of the information technology under any business model.*

In brief, Proposition 3.1 shows that, regardless of market regimes, when the seller does not have the technology to provide information to the buyer directly but has to do so by interacting with an intermediary who has this technology, as the seller has private information about production cost, the ownership of such information technology matters. Indeed, when the seller has to buy such information technology from the intermediary, due to the presence of incomplete information, the seller would demand information rent from the intermediary and total surplus will be reduced since the intermediary has to provide information structures so that the seller would be willing to internalize her information rent when making pricing decisions.

In addition to the welfare implication above, I now examine how the shifts of the distribution of value and the distribution of production cost affects the intermediary's revenue and the total surplus generated by trade. Specifically, take any two pairs of regular distributions (m_1^0, G_1) and (m_2^0, G_2) with $\text{supp}(m_1^0) = \text{supp}(m_2^0) = V$ and $\text{supp}(G_1) = \text{supp}(G_2) = C$ such that (2.7) holds. the previous results then gives the following comparative statics analysis:

Proposition 3.2 (Comparative Statics)**.**

1. *Suppose that m_1^0 first order stochastic dominates m_2^0. Then under both regime $\mathcal{I}$ and regime $\mathcal{P}$, the total surplus, the intermediary's revenue and seller's expected net profit under (m_1^0, G_i) are larger than those under (m_2^0, G_i), $i \in \{1, 2\}$.*

2. *Suppose that m_1^0 is a mean preserving spread of m_2^0. Then under both regime $\mathcal{I}$ and regime $\mathcal{P}$, the intermediary revenue under (m_1^0, G_i) is larger than that under (m_2^0, G_i), $i \in \{1, 2\}$.*

3. *Suppose that G_2 dominates G_1 in the hazard rate order. That is $g_1/G_1 \geq g_2/G_2$. Then under both regime $\mathcal{I}$ and regime $\mathcal{P}$, the total surplus, the intermediary's revenue and the seller's expected revenue are larger under (m_i^0, G_1) than those under (m_i^0, G_2), $i \in \{1, 2\}$.*

To summarize, regardless of the market regime, when the buyer's value becomes higher (in the sense of first order stochastic dominance), it becomes easier for the intermediary to generate trade surplus by providing proper information to the buyer and therefore total surplus, the intermediary's revenue and the seller's net profit all increase. When the distribution of value becomes more spread-out (in the sense of mean preserving spread), the informational tools for the intermediary becomes more flexible and therefore revenue increases. Finally, when the seller's cost shifts in the hazard rate order, causing a reduction of information rent and the costs (in the sense of first order stochastic dominance), the seller retains less information rent and the distortion on information structure for her to internalize pricing decision reduces. These two factors jointly increase total surplus and the intermediary's revenue as well. Furthermore, although the reduction of information rent and reduction of production cost has opposite effects on the seller's net profit, Proposition 3.2 shows that the gain in total surplus offsets the loss of information rent of the seller and hence also increases seller's expected net profit.

3.6 Discussions

The outcome-equivalence between regime $\mathcal{I}$ and regime $\mathcal{P}$ is surprising, since this implies that the ability to contract on prices does not matter. As remarked above, this is essentially because of the richness of information structures. Even though the intermediary under regime $\mathcal{P}$ plays a far less active role in the product market, by exploiting the details of the information provided, he can in fact incentivize the producers to behave as if he has control over price. As such, one of the insights of the equivalence result is that no matter what business model the intermediary uses—be it by selling itself, by contracting on the product price, or by fully delegating the sale to the producer—and regardless of whether the technology of targeting is possible, as long as it is the intermediary who provides the information about the product to the consumers, the induced market outcomes must remain the same.

As a result, even though there are many ways that the informational intermediary can operate with and it has become more and more likely that these intermediaries have the ability to precisely estimate the consumers' value and design policies accordingly, in environments where the intermediary can provide arbitrary information to the buyer, the market outcomes under all these possible business models and targeting technologies are in fact the same as if the intermediary plays a very passive role and only provides information.

Another insight of the equivalence result is that, as demonstrated in Proposition 3.1 and Proposition 3.2, since the market outcomes are equivalent between regime $\mathcal{I}$ and regime $\mathcal{P}$

and do not depend on whether targeting is available, the comparative statics and welfare analyses can be done in a robust way that does not require further specifications of the actual market regime and targeting technology in this environment. In other words, analyzing the solutions of either one of the specifications is sufficient to provide predictions and implications for all other possible specifications.

Finally, from the intermediary's perspective, the equivalence result also sheds light on choosing the optimal business model to operate. As the equivalence result suggests, as long as (2.7) holds, there are no differences between which business model (i.e., market regimes) he uses. This means that when deciding which business model to use, the intermediary only needs to consider the residual perspectives that are orthogonal to this model, such as costs for each business models (e.g. regime $\mathcal{P}$ might require more inventory costs or monitoring cost than regime $\mathcal{I}$) and how these models perform in terms of attracting customers.

Two final remarks are also noteworthy. First, the fact that the buyer always has zero surplus under any solution of any business models is a result of the assumption that buyer does not have any information about the value a priori and thus does not retain any rent when facing monopolies. It is natural to extend the model to generate positive buyer surplus by introducing some prior information to the buyer that is exclusive for himself. In the online appendix of Yang (2019b), I show that the main results of the paper still hold qualitatively. In particular, regime $\mathcal{I}$ and regime $\mathcal{P}$ are still outcome-equivalent under certain regularity conditions. Even though the characterizations of revenue-maximizing solutions become less explicit. Second, although this chapter focuses on the environment of monopolistic pricing in the spirit of Maskin and Riley (1984) and Myerson (1981), the same modeling techniques can also be applied in a setting where there is a regulator who wishes to regulate a monopolist's information disclosure of its product in spirit of Baron and Myerson (1982). The regulator can either control information provision and price at the same time (regime $\mathcal{P}$) or regulate only the information provision (regime $\mathcal{I}$). The payoffs between the monopoly and the consumers are transferable and thus taxes and subsidies are available. The regulator's goal is to maximize a weighed average surplus of the economy. Using similar arguments, with a modified definition for virtual cost function, the equivalence result suggests that both kinds of regulations yield the same outcome and provides optimal policies under each environment. This observation is also summarized in the online appendix of Yang (2019b).

3.7 Conclusion

In this chapter, I explored the revenue-maximizing mechanisms for an intermediary to sell services of providing information about the product to the buyer under two market regimes (i.e., regime $\mathcal{I}$ and regime $\mathcal{P}$). I characterize the revenue-maximizing mechanisms under both market regimes and show that the induced market outcomes are equivalent, regardless of the possibility of targeting. Welfare analysis suggests that total surplus and probability of trade is smaller in this environment comparing to a benchmark in which the seller has the technology to disclose information, independent of the market regime. Comparative statics show that, regardless of the market regime, total surplus and the seller's profit are higher when the buyer's value is higher and when the information rent of the seller is lower, as long as the gains from trade are large enough.

There are several aspects of the model that can be extended. First, the assumption that the seller is a monopoly can be replaced and competition can be introduced. Two natural extensions arise from this. One is to introduce different market structures to the seller. A conjecture is that as long as each seller has some of monopoly power(e.g. a monopolistic competition or a Cournot oligopoly model), the results will be similar, since the key aspect of the seller's behavior in the model above is the mark-up decisions rather than how much profit she can retain. This also allows introducing richer screening device to the intermediary. When there are multiple sellers and the buyer receives information through the intermediary, it is arguable that the order in which the information is presented matters.

As a result, the intermediary can include the order of the products through which the buyer will see as a part of his mechanisms. This is a topic for future research. Second, it is also plausible to consider an environment in which there are more than one intermediaries, these intermediaries compete in providing mechanisms and information outlet to attract buyers and sellers. This is also a direction for future research.

CHAPTER 4

FLEXIBLE INFORMATION: ONLINE PLATFORM DESIGNS

4.1 Introduction

In this chapter, I combine the analyses in the previous two chapters and examine the information monopolist's revenue-maximization problem when *any* information structure that leaves the buyer better-informed is feasible. The leading example of this environment is the design of online platforms. For the past decade, online platforms such as Amazon have developed rapidly. With extensive usage of "big data" and better computation power, these platforms are able to, on one hand, gather more information about the consumers (e.g., purchase histories and demographics) and estimate their values for products and, on the other hand, provide product information to the consumers in various ways (e.g., trial period, free return policy or pictures on the website). These advancements further enable platforms to manipulate information structures between buyers and sellers—They can inform the sellers about buyers' values for their products and facilitate the seller's pricing strategy while they inform the buyers about their values by using various disclosure policies. Meanwhile, these platforms have also developed various business models to operate with. For instance, Amazon has been using two major business models for the past decade—the "Amazon Vendor Central" and the "Amazon Seller Central". Under the Amazon Vendor Central, Amazon purchases products from producers and then sell directly on its website; Under the Amazon Seller Central, Amazon serves as a third party and allows the producers to sell their products directly on its website. In an era where online business becomes essential, it is crucial to understand how these platforms operate and what are the economic consequences. Namely, given that a platform's goal is to maximize revenue, what is the best way to exploit such information technology and which of these business models is the most ideal? What are the effects on market outcomes under various business models? Should policymakers encourage or discourage any ways to exploit this information technology or any particular business models?

In this chapter, I use the model developed in Chapter 1 to study these questions. Specifically, the online platforms discussed above can be thought of as an informational monopolist in Chapter 1, who can flexibly design and sell information structure $\chi \in \mathrm{X}$ to the seller in a monopolistic pricing environment. By examining the information monopolist's revenue-maximizing problems under both regime $\mathcal{I}$ and regime $\mathcal{P}$, I derive two equivalence results (Theorem 4.1). The first is that regime $\mathcal{I}$ and regime $\mathcal{P}$ are outcome-equivalent, even when there are no restrictions on the set of feasible information structures. This aligns with the two equivalence results in Chapter 2 and Chapter 3, and indicates that which business model is used by an online platform is irrelevant to the market outcomes, and that an online platform such as Amazon does not need to be active in the product market at all as long as it can design information structures flexibly. The second equivalence result, which connects the characterizations in Chapter 2 and Chapter 3, shows that even when the informational monopolist can design information structures flexibly, it would be without loss to either restrict attention to those in which the buyer is fully informed ($\mathrm{X}_{\mathrm{Ch2}}$), or those in which the seller is uninformed ($\mathrm{X}_{\mathrm{Ch3}}$). This means that, as an online platform, either fully informing the buyer and then selling information to the seller and facilitate price discrimination or merely manipulating the information provided to the buyer without disclosing anything to the seller are equally profitable, and can both be revenue-maximizing.

The remaining parts of this chapter is organized as follows: Section 4.2 recites the model introduced in Chapter 1 and establishes a convenient representation. Section 4.3 derives the equivalence result. Section 4.4 discusses several extensions that can be applied to the models throughout Chapter 2, Chapter 3 and Chapter 4. Section 4.5 discusses the results and their implications and Section 4.6 concludes.

4.2 Model

As introduced in Chapter 1, the model consists of one buyer, one seller and one informational monopolist. The seller is a monopolist with private cost $c \in C$ who sells a product to the buyer by posting a price. The buyer has quasi-linear preference and value $v \in V$. They decides whether to buy the product at the posted price given the information. The buyer's value v is drawn from a common prior m^0 and the seller's cost is drawn from a CDF G. The informational monopolist designs and sells an information structure $\chi \in \mathrm{X}$ to the seller. The seller with cost c has outside option $\pi_0(c)$. Throughout this chapter, I assume that m^0 and G are regular, that (2.7) holds and that $\pi_0(c) = 0$ for all $c \in C$.

Combining the discussion in Chapter 2 and Chapter 3, the set of feasible information structures X has a convenient representation. Specifically, since the buyer always knows more about v than the seller does and hence the purchasing decision only depends on their interim expected value, it is without loss to represent the buyer's information about v as a collection of distributions that are mean-preserving contractions of m^0, $\mathcal{D}_0$. Furthermore, since the seller's pricing problem only depends on v through the buyer's purchasing decision, given any information $\widehat{D} \in \mathcal{D}_0$ that the buyer has about v, the seller's information about v can be represented by the collection of probability measures $s \in \Delta(\mathcal{D})$ such that (2.1) holds, with D_0 being replaced by $\widehat{D}$. That is,

$$\int_{\mathcal{D}} D(p)s(\mathrm{d}D) = \widehat{D}(p), \; \forall p \in \mathbb{R}_+. \tag{4.1}$$

Given any $\widehat{D} \in \mathcal{D}_0$, let $\mathcal{S}_{\widehat{D}}$ be the collection of $s \in \Delta(\mathcal{D})$ satisfying (4.1). Then, the set of feasible information structures can be represented by a pair $(\widehat{D}, s)$, such that $\widehat{D} \in \mathcal{D}_0$ and $s \in \mathcal{S}_{\widehat{D}}$.

Using this representation, a mechanism for the informational monopolist can be written as $(\alpha, \boldsymbol{D}, \sigma, \tau)$ under regime $\mathcal{I}$, where $\alpha : C \to \{0, 1\}$ is the publicizing decision that specifies whether the seller can participate on the platform given her report c; $\boldsymbol{D} : C \to \mathcal{D}_0$ is a measurable function that maps each seller's report $c \in C$ to $\boldsymbol{D}(c) \in \mathcal{D}_0$ that represents the information provided to the buyer; $\sigma : C \to \Delta(\mathcal{D})$ such that $\sigma(c) \in \mathcal{S}_{\boldsymbol{D}(c)}$ for all $c \in C$ is a measurable function that maps each report c to the information provided to the seller; and $\tau : C \to \mathbb{R}$ is a measurable function that maps each report to a payment $\tau(c)$. Likewise, under regime $\mathcal{P}$, a mechanism for the informational monopolist becomes a tuple $(\boldsymbol{D}, \sigma, \boldsymbol{\gamma}, \tau)$, where $\boldsymbol{\gamma}$ is a measurable function that maps from each reported cost $c \in C$ to a transitional kernel from $\mathcal{D}$ to $\mathbb{R}_+$, which stands for the stochastic price that must be charged given every posterior of the seller. It then follows that the incentive compatibility and individual rationality constraints can be re-written as follows: Under regime $\mathcal{I}$, a mechanism $(\alpha, \boldsymbol{D}, \sigma, \tau)$ is incentive compatible if for any $c, c' \in C$,

$$\alpha(c) \left(\int_{\mathcal{D}} \pi_D(c) \sigma(\mathrm{d}D|c) - \tau(c) \right) \geq \alpha(c') \left(\int_{\mathcal{D}} \pi_D(c) \sigma(\mathrm{d}D|c') - \tau(c') \right),$$

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

$$\alpha(c) \left(\int_{\mathcal{D}} \pi_D(c) \sigma(\mathrm{d}D|c) - \tau(c) \right) \geq 0.$$

Meanwhile, under regime $\mathcal{P}$, a mechanism $(\boldsymbol{D}, \sigma, \boldsymbol{\gamma}, \tau)$ is incentive compatible if for all $c, c' \in C$,

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

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

$$\int_{\mathcal{D} \times \mathbb{R}_+} (p-c) D(p) \boldsymbol{\gamma}(\mathrm{d}p|D, c) \sigma(\mathrm{d}D|c) - \tau(c) \geq 0.$$

Notice that by restricting $\boldsymbol{D}$ so that $\boldsymbol{D}(c) = D_0$ for all c, the mechanisms reduce to those in Chapter 2, whereas by restricting σ so that $\sigma(c)$ must be a dirac measure on $\boldsymbol{D}(c)$, the mechanisms are reduced to those in Chapter 3.

4.3 Optimal Mechanisms and Equivalence Results

In this section, I use the results from Chapter 2 and Chapter 3 to derive optimal mechanisms for the informational monopolist. To begin with, I first consider the informational monopolist's problem under regime $\mathcal{P}$. As in the previous two sections, the ability to contract on price effectively reduces the allocation-space to one-dimensional and hence a revenue-equivalence formula induced by local incentives and a monotonicity condition would be necessary and sufficient for global incentive compatibility, as formally stated by Lemma 4.1

Lemma 4.1. *Under regime $\mathcal{P}$, a mechanism $(\boldsymbol{D}, \sigma, \boldsymbol{\gamma}, \tau)$ is incentive compatible if and only if:*

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

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

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

Using Lemma 4.1, the informational monopolist's expected revenue under any incentive compatible mechanism $(\boldsymbol{D}, \sigma, \boldsymbol{\gamma}, \tau)$ can be written as

$$\mathbb{E}[\tau(c)] = \int_C \left(\int_{\mathcal{D}\times\mathbb{R}_+} (p - \phi_G(c))D(p)\boldsymbol{\gamma}(\mathrm{d}p|c,D)\sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \bar{\tau}. \tag{4.2}$$

Meanwhile, notice that the integrand of the right-hand side of (4.2) can be bounded from above as follows:

$$\begin{aligned}\int_{\mathcal{D}\times\mathbb{R}_+}(p-c)D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\sigma(\mathrm{d}D|c) &\leq \int_{\mathcal{D}}\max_{p\geq 0}(p-c)D(p)\sigma(\mathrm{d}D|c)\\ &\leq \int_{\{v\geq\phi_G(c)\}}(v-\phi_G(c))\boldsymbol{D}(\mathrm{d}v|c)\\ &\leq \int_{\{v\geq\phi_G(c)\}}(v-\phi_G(c))D_0(\mathrm{d}v),\end{aligned}$$

where the second inequality holds since $\sigma(c)\in\mathcal{S}_{\boldsymbol{D}(c)}$ and since the seller's profit when her cost is $\phi_G(c)$ is always below the total gains from trade in the economy with demand function $\boldsymbol{D}(c)$; whereas the last inequality follows from the fact that $\boldsymbol{D}(c)\in\mathcal{D}_0$ and convexity of the function $v\mapsto\max\{(v-\phi_G(c)),0\}$. Together, it then follows that for any incentive compatible mechanism $(\boldsymbol{D},\sigma,\boldsymbol{\gamma},\tau)$,

$$\mathbb{E}[\tau(c)]\leq\int_C\left(\int_{\{v\geq\phi_G(c)\}}(v-\phi_G(c))D_0(\mathrm{d}v)\right)G(\mathrm{d}c)=\int_C\left(\int_{\phi_G(c)}^{\infty}D_0(v)\,\mathrm{d}v\right)=:R^*.$$

Furthermore, by the analyses in Chapter 2 and Chapter 3, any optimal mechanisms under regime $\mathcal{P}$ even when restricting X to $\mathrm{X}_{\mathrm{Ch2}}$ or $\mathrm{X}_{\mathrm{Ch3}}$ attains R^* (possibly up to a constant, depending on the outside option of the seller when her cost is $\bar{c}$, provided that (2.7) holds). Therefore, any optimal mechanisms derived in Chapter 2 or Chapter 3 under regime $\mathcal{P}$, is optimal and the informational monopolist's optimal revenue under regime $\mathcal{P}$ is exactly R^*.

On the other hand, under regime $\mathcal{I}$, when prices are not contractible, the informational monopolist's revenue-maximization problem cannot be reduced to a standard one-dimensional nonlinear pricing problem. Thus, as in Chapter 2 and Chapter 3, I solve this revenue-maximization problem by constructing an incentive feasible mechanism that attains the revenue upper bound R^*. To this end, the mechanism constructed has to exploit the richness of the allocation space to discipline the pricing behavior of the seller so that she would optimally charge the same prices as that would have been contracted under regime

$\mathcal{P}$. As in the previous chapters, I first provide a characterization of incentive compatible mechanisms, where the transfers are pinned down by local incentive constraints and global incentives constraints are described by a family of integral inequalities.

Lemma 4.2. *Under regime $\mathcal{I}$, a mechanism $(\alpha, \boldsymbol{D}, \sigma, \tau)$ is incentive compatible if and only if:*

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

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

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

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

Furthermore, the "only if" part holds for any $\boldsymbol{p} \in \boldsymbol{P}$.

Thus, by Lemma 4.2, the expected revenue under any incentive compatible mechanism $(\alpha, \boldsymbol{D}, \sigma, \tau)$ can be written as

$$\mathbb{E}[\tau(c)] = \int_C \alpha(c) \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c))D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) \right) G(\mathrm{d}C) - \bar{\tau}.$$

Now notice that, when restricting $(\alpha, \boldsymbol{D})$ so that $\alpha(c) = 1$ and $\boldsymbol{D}(c) = D_0$ for all $c \in C$, the revenue maximization problem for the informational monopolist becomes the same as that in Chapter 2, except that the outside options are not type-dependent. As a result, since m^0 is regular and since (2.7) holdes, and hence Assumption 2.1 holds, by Theorem 2.6 in Chapter 2, the canonical ϕ_G-quasi-perfect mechanism is incentive feasible and hence gives revenue R^*.[1] Contrarily, when restricting σ so that $\sigma(c) = \delta_{\{\boldsymbol{D}(c)\}}$ for all $c \in C$, the

1. Notice that by (2.7) and by regularity of G, $\overline{\varphi}_G(c) = \phi_G(c)$ for all $c \in C$. Moreover, since $\pi_0(\overline{c}) = 0$, $\pi_0(\overline{c})$ is replaced by 0.

revenue-maximization problem becomes exactly the same as that in Chapter 3. Thus, by Theorem 3.2, the ϕ_G-upper-censorship mechanism is incentive feasible and gives revenue R^* as well. Together, this implies that all of the optimal mechanisms in Chapter 2 (i.e., when X is restricted to $\mathrm{X}_{\mathrm{Ch2}}$) and the optimal mechanisms in Chapter (3) (i.e., when X is restricted to $\mathrm{X}_{\mathrm{Ch3}}$) are optimal and the optimal revenue is also R^*. Furthermore, since R^* is the level of efficient surplus in an economy where the seller's cost is replaced by the virtual cost, and since the informational monopolist's revenue is bounded from above by the seller's optimal profit when her cost is replaced by the virtual cost, it must be that under any optimal mechanisms, the consumer surplus is zero and that the product is sold if and only if $v \geq \phi_G(c)$ for (almost) all realized cost c.

The observations above have two implications. First, the outcome-equivalence results derived in Chapter 2 and Chapter 3 are still valid when there are no further restrictions on the set of feasible information structures X. Second, regardless of the market regimes, the induced market outcomes are all the same across all the different restrictions on feasible information structures $\mathrm{X}_{\mathrm{Ch2}}$, $\mathrm{X}_{\mathrm{Ch3}}$, X, and all others "in between". These two observations are summarized below by Theorem 4.1.

Theorem 4.1 (Outcome-Equivalence)**.** *The induced market outcomes are the same for every optimal mechanism under both market regimes $\mathcal{I}$ and $\mathcal{P}$ and every set of feasible information structure $\widehat{\mathrm{X}}$ such that $\mathrm{X}_{\mathrm{Ch2}} \subseteq \widehat{\mathrm{X}}$ or $\mathrm{X}_{\mathrm{Ch3}} \subseteq \widehat{\mathrm{X}}$.*

Theorem 4.1 has two major implications. First, it means that both feasible information structures for the informational monopolist *and* market regimes are irrelevant for market outcomes. That is, both the informational monopolist's ability to contract on prices and the additional ability to manipulate information structures that is beyond those discussed in Chapter 2 and Chapter 3 does not affect market outcomes. Second, it means that the set of feasible information structures considered in Chapter 2 and Chapter 3 are "rich enough" for the informational monopolist to exploit and discipline the seller's pricing behavior. Given that the informational monopolist is able any create information structures in either $\mathrm{X}_{\mathrm{Ch2}}$ or $\mathrm{X}_{\mathrm{Ch3}}$, additional ways to manipulate information structures does not matter. As a result, the equivalence result derived in Chapter 2 is robust to allowing the data broker to garble the buyer's information about their value, whereas the equivalence result in Chapter 3 is robust to allowing the intermediary to further provide information about the buyer's signal realizations to the seller.

4.4 Extension: Different Private Types

Throughout Chapters 2, Chapter 3 and Chapter 4, the analyses have been focusing on the case where the seller has private cost. Although this is the case in many settings (e.g., data brokers sell consumer data to producers in various industries who know better about the industrial details), there are certainly other situations where the seller could have private information about something else. For instance, in the setting of Chapter 2, it is arguable that the producer would know better about how consumer characteristics are map into consumer values than the data broker does; In the setting of Chapter 3, it is sometimes the case that a seller would know better about consumers' "taste" than the intermediary does; In the setting of Chapter 4, it can be the case that the online platform only has the technology to manipulate information about consumers' values partially, while consumers and producers know about other determinants of values better than the platform does.

In this section, I consider an alternative model where the seller has different private type. That is, I consider the environment where the seller has private information about part of the buyer's value. For tractability, I focus on an one-dimensional environment and consider *additive* and *multiplicative* cases only. Specifically, consider the model introduced in Chapter 1, but suppose that the seller's cost is common knowledge and hence normalized to zero. In contrast, suppose that the buyer's value for the product is $f(v, \xi)$ for some measurable function f, where $v \in V$ is drawn from common prior m^0 and the informational monopolist can design information structure $\chi \in \mathrm{X}$ for v, whereas $\xi \in \Xi \subseteq \mathbb{R}$ is drawn from a commonly known CDF G and only the buyer and the seller can see the realization of ξ. Under this setting, the *additive case* corresponds to assuming that $f(v, \xi) = v - \xi$, $\underline{v} > 0$ and $\Xi = [0, \underline{v}]$. Contrarily, the *multiplicative* case corresponds to assuming that $f(v, \xi) = \xi \cdot v$ and $\xi > 0$.

For the additive case, it turns out that it is essentially equivalent to the original model where the seller has private information about her production cost. To see this, consider any realized ξ and any $D \in \mathcal{D}$. Since the buyer knows ξ, their purchasing decision given posted price p is to buy if their interim expected value is above $p + \xi$, and not to buy otherwise. Thus, the seller's profit when charging price $p' \geq 0$ is

$$p' D(p' + \xi),$$

which, by writing $p = p' + \xi$, becomes

$$(p - \xi) D(p),$$

which is exactly the seller's profit when she has cost ξ and charges price $p = p' + \xi$. As a result, all the analyses in Chapter 2, Chapter 3 and Chapter 4 can be interpreted in an alternative way, where the seller has private information about a parameter that affects the buyer's value in an additive way.

As for the multiplicative case, it turns out that the ability to contract on price does not matter at all, and therefore the informational monopolist's revenue maximization can be solved by the standard Myersonian approach. To see this, consider any $\xi > 0$ and any $D \in \mathcal{D}$, again the buyer's purchasing decision given posted price p is that they will buy if their interim expected value is above p/ξ and will not buy otherwise. Thus, the seller's profit from charging price p' is

$$p' D\left(\frac{p'}{\xi}\right),$$

which, by writting $p = p'/\xi$, becomes

$$\xi \cdot pD(p).$$

In particular, for the informational monopolist's perspective, the seller's private type only scales her payoff and does not direct affects her pricing decision. Thus, there are no moral hazard concerns as presented in the additive case and the informational monopolist's problem (under both regimes) can be solved by the standard Myersonian approach. For instance, consider the setting in this chapter and consider any mechanism $(\alpha, \boldsymbol{D}, \sigma, \tau)$ for the informational monopolist under regime $\mathcal{I}$, the seller with type ξ has payoff

$$\xi \cdot \alpha(\xi') \int_{\mathcal{D}} \pi_D(0)\sigma(\mathrm{d}D|\xi') - \tau(\xi')$$

when reporting ξ'. Thus, by letting

$$\Pi(\xi') := \alpha(\xi') \int_{\mathcal{D}} \pi_D(0)\sigma(\mathrm{d}D|\xi'),$$

the seller's payoff becomes

$$\xi\Pi(\xi') - \tau(\xi').$$

Using standard Myersonian arguments as in Myerson (1981), the incentive compatible mechanisms can be characterized by a revenue-equivalence formula and monotonicity of Π and the expected revenue under any incentive compatible mechanism can be written as

$$\mathbb{E}_G[\tau(\xi)] = \int_{\Xi} \Pi(\xi) \left(\xi - \frac{1 - G(\xi)}{g(\xi)} \right) G(\mathrm{d}\xi) - \bar{\tau}.$$

for some constant $\bar{\tau} \in \mathbb{R}$.

As a result, under regime $\mathcal{I}$, the revenue is maximized by finding $(\alpha, \boldsymbol{D}, \sigma)$ to maximize

$\Pi(\xi)$ whenever $\xi \geq \xi^*$ and minimize $\Pi(\xi)$ whenever $\xi < \xi^*$ subject to monotonicity of Π, where ξ^* is the largest solution of

$$\max_{\xi \in \Xi} \xi(1 - G(\xi)).$$

Thus, for any $\xi \geq \xi^*$, let $\alpha(\xi) = 1$ and let $\boldsymbol{D}(\xi)$ and $\sigma(\xi)$ be chosen so that the seller receives total surplus (by letting the seller and the buyer know equally about v); while for any $\xi < \xi^*$, let $\alpha(\xi) = 0$. Then Π is monotone and the revenue is maximized.

Meanwhile, under regime $\mathcal{P}$, given any mechanism $(\boldsymbol{D}, \sigma, \boldsymbol{\gamma}, \tau)$, the seller's payoff is

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

when her type is ξ and reports ξ'. By letting

$$\overline{\Pi}(\xi) := \int_{\mathcal{D} \times \mathbb{R}_+} pD(p)\boldsymbol{\gamma}(\mathrm{d}p|D)\sigma(\mathrm{d}D|\xi)$$

and by using the same arguments as above, the informational monopolist's expected revenue under any incentive compatible mechanism must be

$$\mathbb{E}_G[\tau(\xi)] = \int_{\Xi} \overline{\Pi}(\xi) \left(\xi - \frac{1 - G(\xi)}{g(\xi)} \right) G(\mathrm{d}\xi) - \bar{\tau}.$$

for some constant $\bar{\tau} \in \mathbb{R}$. As a result, for any $\xi \geq \xi^*$, $\overline{\Pi}(\xi)$ is maximized by giving the seller the same information as the buyer and pricing optimally. On the other hand, when $\xi < \xi^*$, $\overline{\Pi}$ is minimized by letting $\boldsymbol{\gamma}(\xi)$ a dirac measure on zero regardless of D.

It is noteworthy that in the multiplicative case, the equivalence between regime $\mathcal{I}$ and regime $\mathcal{P}$ follows mechanically. This is only because under this particular parametric assumption, the seller's profit maximization problem is effectively independent of her private type, which means that there is no moral hazard component at all. Generically, other functional forms of f do not produce this feature.

4.5 Discussions

The results derived above have several economic implications. First, they provide an outcome-equivalence result for online platforms that are similar to those stated in Chapter 2 (for data brokers and market consultants) and Chapter 3 (for intermediaries and advertisement agencies). According to Theorem 4.1, for online platforms such as Amazon who is able to design information structures to both inform the seller and the buyer (subject to a constraint that the buyer always knows more about their value than the seller does), it does not matter whether they have abilities to contract on price. Thus, as regime $\mathcal{P}$ is equivalent to a model where a platform becomes a retailer, Theorem 4.1 implies that for online platforms, being a retailer who buys the product first and then sell the product to the consumers is the same as being a third party that only provides informational services. There are no additional benefits for platforms to be more active in the product market. This observation provides some insights for how Amazon should evaluate its two major business models: Other things equal, the Vendor central and the Seller central are outcome-equivalent. In particular, the two business models generate the same level of revenue and thus comparisons between these two models can be simplified significantly. Meanwhile, from a policymaker's perspective, the outcome-equivalence result implies that it is not necessary to discourage one business model over another. In terms of market outcomes, whether a platform is active in the product market does not affect market outcomes at all.

Another implication that can be drawn from Theorem 4.1 pertains to how online platforms should exploit their informational technology. Although the model in this chapter allows the informational monopolist to design *any* information structure $\chi \in \mathrm{X}$, Theorem 4.1 implies that either $\mathrm{X}_{\mathrm{Ch2}}$ or $\mathrm{X}_{\mathrm{Ch3}}$ is enough. This means that for a platform that can both provide information about products to consumers and then provide further information about the consumers' interim expected values to the producer, it would be sufficient to simply provide full information to the consumers and then design information provided to the seller, or to only design information provided to the consumers without revealing anything to the producers. In the context of previous two chapters, this is equivalent to saying that it is sufficient for an online platform to either always provide full information to the consumers and act simply as a data broker who provides producers information to facilitate price discrimination; or to be just an advertiser who helps the producers provide information about their products to the consumers. This observation in turn provides insights for a policymaker. While currently many of the debates on welfare effects of online platforms are focusing on how platforms facilitate price discrimination using personal data, the equivalence result implies that it is equally important to understand the welfare effects even when platforms are not facilitating price discrimination at all but only exploiting their informational services by providing various kinds of information to the consumers.

Lastly, all the equivalence results in Chapter 2, Chapter 3 and Chapter 4 can be thought of as a way to gauge how powerful the ability to design information structures in a monopolistic pricing setting is. Indeed, the equivalence results in various settings (Theorem 2.4, Theorem 3.3 and Theorem 4.1) mean that the ability to design and sell information structures under the relevant restrictions is very powerful that even if the informational monopolist is given extra ability in the product market (i.e., the ability to contract on price), the revenue—and the induced market outcomes—would still remain the same. In other words, these results mean that the ability to sell information structures from a "rich" set implicitly empowers the informational monopolist to control prices in the product market, even when he only serves as a third party who does not participate in the market at all.

4.6 Conclusion

In this chapter, I examine the informational monopolist's revenue maximization problem when there are no restrictions on the set of information structures that can be provided. It turns out that regime $\mathcal{I}$ and regime $\mathcal{P}$ are still outcome-equivalent, which further implies the same equivalence under any other set of information structures that includes either $\mathrm{X}_{\mathrm{Ch2}}$ and $\mathrm{X}_{\mathrm{Ch3}}$. Overall, the result in this chapter provides a wide range of sufficient conditions on the set of feasible information structures for the outcome-equivalence result to hold. Furthermore, the optimal mechanisms derived in Chapter 2 and Chapter 3 are able to attain the optimal revenue when there are no restrictions on the set of information structures. This implies that for an online platform, either of the following approach is sufficient for revenue-maximization: 1) Fully revealing information to the consumers and then serve as a data broker who provide information about consumers' values to the producers and facilitate price discrimination. 2) Only serves as an advertiser who help the producers provide product information to the consumers while revealing nothing about the consumers' values to the producers.

There are several other interesting aspects that can be left for topics of future research. First, throughout Chapter 2, Chapter 3 and Chapter 4, the analyses are restricted to a simple monopolistic pricing setting where the seller has private information about her production cost. It would be an interesting question to see whether a similar outcome-equivalence result can hold in more general settings. This can be thought of as part of a broader research agenda in mechanism design, which aims for finding economically relevant environments and sufficient conditions for a mechanism designer to exploit richness of allocation spaces and discipline extra complicating incentives such as moral hazard components. Second, as part of another research agenda, a natural following step would be to explore the economic implications of other aspects of online platforms, such as the ability to rank sellers in different orders, as well as the ability to reveal different information about product prices charged by sellers and their competitors.

Appendix for Chapter 2

A Notational Conventions

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

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

This would corresponds to the weak-* topology on $\Delta(V)$ and hence this topology on $\mathcal{D}$ is also called the weak-* topology. As a result, $\mathcal{D}$ is a Polish space. Furthermore, notice that under this topology, $\{D_n\} \to D$ if and only if $\{D_n(p)\} \to D(p)$ for all $p \in V$ at which D is continuous.

Now I introduce some more notational conventions that are implicitly used in the main text and will be used throughout the proof. For any measurable sets X and Y, the collection of measurable functions $f : X \to Y$ is denoted by X^Y. Moreover, for any $f \in \mathbb{R}^X$, define f^+ by $f^+(x) := \max\{f(x), 0\}$ for all $x \in X$. For any $f, g \in \mathbb{R}^X$ and for any $\alpha, \beta \in \mathbb{R}$, define $\alpha f + \beta g \in \mathbb{R}^X$ by $[\alpha f + \beta g](x) := \alpha f(x) + \beta g(x)$ for all $x \in X$. If $X \subset \mathbb{R}^n$ and the partial derivative of f with respect to x_i exists for some $i \in \{1, \ldots, n\}$, use

$$f_i(x_1, \ldots, x_n) := \frac{\partial}{\partial x_i} f(x_1, \ldots, x_n)$$

to denote this partial derivative. When $X \subseteq \mathbb{R}$, for any $x \in \text{int}(X)$, let

$$f(x^+) := \lim_{x' \downarrow x} f(x') \text{ and } f(x^-) := \lim_{x' \uparrow x} f(x')$$

be the right and the left limits of f at x provided they exist, respectively.

For any measurable space X, let $\Delta^f(X) \subseteq \Delta(X)$ be the collection of probability measures on X that have finite support. The collection of probability measures $\Delta(X)$ is endowed with the algebraic structure so that for any $\mu_1, \mu_2 \in \Delta(X)$,

$$[\lambda\mu_1 + (1-\lambda)\mu_2](A) := \lambda\mu_1(A) + (1-\lambda)\mu_2(A),$$

for any measurable $A \subseteq X$. Furthermore, for any $x \in X$, $\delta_{\{x\}}$ denotes the Dirac measure that assigns probability 1 to the element x (whenever $\{x\}$ is measurable). For any probability measure $\mu \in \Delta^f(X)$ with finite support and for any measurable set $A \subseteq X$, the simplifying notation

$$\sum_{x \in A} \mu(x) := \sum_{x \in A \cap \text{supp}(\mu)} \mu(\{x\})$$

will be used.

Finally, for any $D \in \mathcal{D}$, let $\mathcal{S}_D$ denote the collection of $s \in \Delta(\mathcal{D})$ such that (2.1) holds with D_0 being replaced by D (so that $\mathcal{S}_{D_0} = \mathcal{S}$). Also, let D^{-1} denote the inverse demand of D, where D^{-1} is defined as

$$D^{-1}(q) := \sup\{p \in V : D(p) \geq q\}, \forall q \in [0,1]. \tag{2.14}$$

B Technical Lemmas and Proofs

This section contains technical lemmas that establish continuities of several critical objects. These continuity results are crucial for proving the main results.

Lemma 2.4. *For any $D \in \mathcal{D}$, $\pi_D \in \mathbb{R}_+^C$ is continuous and convex. Furthermore, for any $\boldsymbol{p} \in \boldsymbol{P}$, and for any $c \in C$, $-D(\boldsymbol{p}_D(c))$ is a subgradient of π_D at c. In particular, for any $\underline{c} \leq c < c' \leq \bar{c}$,*

$$\pi_D(c) - \pi_D(c') = \int_c^{c'} D(\boldsymbol{p}_D(z))\,\mathrm{d}z \tag{2.15}$$

for any $\boldsymbol{p} \in \boldsymbol{P}$

Proof. By definition of π_D, for any $c \in C$,

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

As such, π_D is convex for all $D \in \mathcal{D}$ since it is the pointwise supremum of a family of affine functions. Moreover, for any $\boldsymbol{p} \in \boldsymbol{P}$ and for any $c, c' \in C$,

$$\begin{aligned}
0 \leq & \pi_D(c') - (\boldsymbol{p}_D(c) - c')D(\boldsymbol{p}_D(c)) \\
= & \pi_D(c') + c'D(\boldsymbol{p}_D(c)) - \boldsymbol{p}_D(c)D(\boldsymbol{p}_D(c)) \\
= & \pi_D(c') - cD(\boldsymbol{p}_D(c)) + c'D(\boldsymbol{p}_D(c)) - (\boldsymbol{p}_D(c) - c)D(\boldsymbol{p}_D(c)) \\
= & \pi_D(c') - [-D(\boldsymbol{p}_D(c))(c' - c) + \pi_D(c)].
\end{aligned}$$

Thus, $-D(\boldsymbol{p}_D(c))$ is a subgradient of π_D at c. Together with convexity of π_D, π_D is differentiable almost everywhere and

$$\pi'_D(c) = -D(\boldsymbol{p}_D(c)),$$

for almost all $c \in C$. Thus, since π_D is convex, for any $\underline{c} \leq c < c' \leq \bar{c}$,

$$\pi_D(c) - \pi_D(c') = \int_c^{c'} D(\boldsymbol{p}_D(z))\,\mathrm{d}z,$$

for any $\boldsymbol{p} \in \boldsymbol{P}$.

For continuity, notice that for any $c \in C$,

$$\pi_D(c) = \max_{(p,q)\in\Xi} (p-c)q,$$

where $\Xi := \text{cl}(\{(p, D(p)) : p \in V\})$ is a compact set in $\mathbb{R}^2$. Therefore, by Berge's theorem of maximum, π_D is continuous on C. ■

Lemma 2.5. *The correspondence* $\boldsymbol{P}$ *is compact-valued and thus*

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

and

$$\underline{\boldsymbol{p}}_D(c) := \min \boldsymbol{P}_D(c)$$

are well-defined for all $c \in C$*,* $D \in \mathcal{D}$*. Furthermore, for any* $D \in \mathcal{D}$*, the correspondence* $\boldsymbol{P}_D : C \rightrightarrows \mathbb{R}_+$ *is upper-hemicontinuous. In particular,* $\overline{\boldsymbol{p}}_D \in \mathbb{R}_+^C$ *is right-continuous and* $\underline{\boldsymbol{p}}_D \in \mathbb{R}_+^C$ *is left-continuous.*

Proof. Consider any $c \in C$ and $D \in \mathcal{D}$. Suppose that $\{p_n\} \subseteq \boldsymbol{P}_D(c)$ and $\{p_n\} \to p$ for some $p \in \mathbb{R}_+$. Since the function $p \mapsto (p-c)D(p)$ is upper-semicontinuous,

$$\pi_D(c) = \limsup_{n\to\infty}(p_n - c)D(p_n) \leq (p-c)D(p) \leq \pi_D(c).$$

Thus, $p \in \boldsymbol{P}_D(c)$. As a result, since $\boldsymbol{P}_D(c) \subseteq \overline{V}$ (see footnote 5), for all $c \in C$ and $D \in \mathcal{D}$, $\boldsymbol{P}_D(c)$ is a closed subset of a compact set, which implies that $\overline{\boldsymbol{p}}_D(c)$ and $\underline{\boldsymbol{p}}_D(c)$ are well-defined.

Now consider any $D \in \mathcal{D}$. To show upper-hemicontinuity of $\boldsymbol{P}_D$, it suffices to show that for any sequences $\{c_n\} \subseteq C$ and $\{p_n\} \subseteq \mathbb{R}_+$ such that $\{p_n\} \to p \in \mathbb{R}_+$ and $\{c_n\} \to c \in C$ and that $p_n \in \boldsymbol{P}_D(c_n)$ for all $n \in \mathbb{N}$, $p \in \boldsymbol{P}_D(c)$. Indeed, for any $n \in \mathbb{N}$, since $p_n \in \boldsymbol{P}_D(c_n)$,

$\pi_D(c_n) = (p_n - c_n)D(p_n)$ for all $n \in \mathbb{N}$. Moreover, since $\pi_D \in \mathbb{R}_+^C$ according to Lemma 2.4,

$$\lim_{n\to\infty} \pi_D(c_n) = \pi_D(c).$$

Therefore, since D is upper-semicontinuous,

$$\pi_D(c) = \lim_{n\to\infty} \pi_D(c_n) = \limsup_{n\to\infty}(p_n - c_n)D(p_n) \leq (p-c)D(p) \leq \pi_D(c).$$

Thus, $p \in \boldsymbol{P}_D(c)$ as desired. Finally, since for any $\boldsymbol{p} \in \boldsymbol{P}$, $\boldsymbol{p}_D \in \mathbb{R}_+^C$ is nondecreasing, upper-hemicontinuity of $\boldsymbol{p}_D$ then implies right-continuity of $\overline{\boldsymbol{p}}_D$ and left-continuity of $\underline{\boldsymbol{p}}_D$. This completes the proof. ■

Lemma 2.6. *For any $D \in \mathcal{D}$, the function $c \mapsto D(\overline{\boldsymbol{p}}_D(c))$ is right-continuous.*

Proof. Consider any $D \in \mathcal{D}$ and any $c \in C$. By Lemma 2.5,

$$\lim_{c'\downarrow c} \overline{\boldsymbol{p}}_D(c') = \overline{\boldsymbol{p}}_D(c).$$

Together with continuity of π_D, which is due to Lemma 2.4,

$$\begin{aligned}(\overline{\boldsymbol{p}}_D(c) - c)D(\overline{\boldsymbol{p}}_D(c)) =& \pi_D(c) \\ =& \lim_{c'\downarrow c} \pi_D(c') \\ =& \lim_{c'\downarrow c}(\overline{\boldsymbol{p}}_D(c') - c')D(\overline{\boldsymbol{p}}_D(c')) \\ =& (\overline{\boldsymbol{p}}_D(c) - c) \cdot \lim_{c'\downarrow c} D(\overline{\boldsymbol{p}}_D(c')),\end{aligned}$$

and hence

$$\lim_{c'\downarrow c} D(\overline{\boldsymbol{p}}_D(c')) = D(\overline{\boldsymbol{p}}_D(c)),$$

as desired. ■

Lemma 2.7. *For any $c \in C$, the function $D \mapsto \pi_D(c)$ is continuous on $\mathcal{D}$.*

Proof. Since $V \subseteq \mathbb{R}_+$ is bounded, this lemma is a special case of Theorem 12 of Hart and Reny (2019) when $k = 1$. ∎

Lemma 2.8. *For any $c \in (\underline{c}, \overline{c})$, the function $D \mapsto D(\overline{\boldsymbol{p}}_D(c))$ is lower-semicontinuous on $\mathcal{D}$.*

Proof. For any $c \in (\underline{c}, \overline{c})$ and for any $D \in \mathcal{D}$, define $\pi'_D(c^+)$ as

$$\pi'_D(c^+) := \lim_{c' \downarrow c} \frac{\pi_D(c') - \pi_D(c)}{c' - c}.$$

Since π_D is convex, $\pi'_D(c^+)$ is well-defined. Furthermore, by Lemma 2.4, $-D(\overline{\boldsymbol{p}}_D(c))$ is a subgradient of π_D at c and therefore, for any $c' > c$,

$$\frac{\pi_D(c') - \pi_D(c)}{c' - c} \geq -D(\overline{\boldsymbol{p}}_D(c)),$$

which implies that

$$\pi'_D(c^+) \geq D(\overline{\boldsymbol{p}}_D(c)). \tag{2.16}$$

Meanwhile, by (2.15), for any $c' > c$,

$$\frac{\pi_D(c') - \pi_D(c)}{c' - c} = \frac{1}{c' - c} \int_c^{c'} -D(\overline{\boldsymbol{p}}_D(z))\,\mathrm{d}z \leq -D(\overline{\boldsymbol{p}}_D(c')).$$

Thus, by Lemma 2.6,

$$\pi'_D(c^+) \leq \lim_{c' \downarrow c} -D(\overline{\boldsymbol{p}}_D(c')) = -D(\overline{\boldsymbol{p}}_D(c)). \tag{2.17}$$

Combining (2.16) and (2.17),

$$\pi'_D(c^+) = D(\overline{\boldsymbol{p}}_D(c)).$$

Now consider any $D \in \mathcal{D}$ and any $\{D_n\} \subseteq \mathcal{D}$ such that $\{D_n\} \to D$, Lemma 2.7 implies that $\{\pi_{D_n}\} \to \pi_D$ pointwise. Thus, for any $c \in (\underline{c}, \overline{c})$, by Theorem 24.5 of Rockafellar (1970),

$$-\liminf_{n\to\infty} D_n(\overline{\boldsymbol{p}}_{D_n}(c)) = \limsup_{n\to\infty} \pi'_{D_n}(c^+) \leq \pi'_D(c^+) = -D(\overline{\boldsymbol{p}}_D(c)).$$

Therefore, for any $c \in (\underline{c}, \overline{c})$,

$$\liminf_{n\to\infty} D_n(\overline{\boldsymbol{p}}_{D_n}(c)) \geq D(\overline{\boldsymbol{p}}_D(c)),$$

as desired. ∎

Lemma 2.9. *For any $c \in C$, the function $D \mapsto \overline{\boldsymbol{p}}_D(c)$ is upper-semicontinuous on $\mathcal{D}$.*

Proof. Consider any $c \in C$ and any sequence $\{D_n\} \subseteq \mathcal{D}$ such that $\{D_n\} \to D$ for some $D \in \mathcal{D}$. Let

$$\bar{p} := \limsup_{n\to\infty} \overline{\boldsymbol{p}}_{D_n}(c).$$

Take a subsequence $\{D_{n_k}\} \subseteq \{D_n\}$ such that

$$\lim_{k\to\infty} \overline{\boldsymbol{p}}_{D_{n_k}}(c) = \bar{p}.$$

First notice that since $D \in \mathcal{D}$ is upper-semicontinuous, for any sequence $\{\delta_k\} \subset \mathbb{R}_+$ such that $\{\delta_k\} \to 0$,

$$\limsup_{k\to\infty} (\overline{\boldsymbol{p}}_{D_{n_k}}(c) - c) D_{n_k}(\overline{\boldsymbol{p}}_{D_{n_k}}(c) - \delta_k) \leq (\bar{p} - c) D(\bar{p}). \tag{2.18}$$

Moreover, by the definition of the Lévy Prokhorov metric, for any $k \in \mathbb{N}$,

$$D_{n_k}(\overline{\boldsymbol{p}}_{D_{n_k}}(c)) \leq D\left(\overline{\boldsymbol{p}}_{D_{n_k}}(c) - \left(\rho(D_{n_k}, D) + \frac{1}{k}\right)\right) - \left(\rho(D_{n_k}, D) + \frac{1}{k}\right), \tag{2.19}$$

where $\rho : \mathcal{D} \times \mathcal{D} \to \mathbb{R}_+$ is the Lévy Prokhorov metric. Together, since $\{\rho(D_{n_k}, D)\} \to 0$ as

$k \to \infty$, which is because $\{D_{n_k}\} \to D$ as $k \to \infty$, we have

$$\begin{aligned}
\pi_D(c) =& \lim_{k\to\infty} \pi_{D_{n_k}}(c) \\
=& \lim_{k\to\infty} (\overline{\boldsymbol{p}}_{D_{n_k}}(c) - c) D_{n_k}(\overline{\boldsymbol{p}}_{D_{n_k}}(c)) \\
\leq& \limsup_{k\to\infty} (\overline{\boldsymbol{p}}_{D_{n_k}}(c) - c) \left[D\left(\overline{\boldsymbol{p}}_{D_{n_k}}(c) - \left(\rho(D_{n_k}, D) + \frac{1}{k} \right) \right) - \left(\rho(D_{n_k}, D) + \frac{1}{k} \right) \right] \\
\leq& (\bar{p} - c) D(\bar{p}) \\
\leq& \pi_D(c),
\end{aligned}$$

where the first equality follows from Lemma 2.7, the first inequality follows from (2.19), and the second inequality follows from $\{\rho(D_{n_k}, D)\} \to 0$ as $k \to \infty$, (2.18), as well as the fact that $\overline{\boldsymbol{p}}_{D_{n_k}}(c) \leq \max \bar{V}$ (see footnote 5). As a result, it then follows that $\bar{p} \in \boldsymbol{P}_D(c)$ and therefore $\bar{p} \leq \overline{\boldsymbol{p}}_D(c)$. Thus,

$$\limsup_{n\to\infty} \overline{\boldsymbol{p}}_{D_n}(c) = \bar{p} \leq \boldsymbol{p}_D(c),$$

as desired. ∎

C Crucial Properties of Quasi-Perfect Schemes

This section summarizes some crucial properties of a ψ-quasi-perfect scheme.

Lemma 2.10. *Consider any nondecreasing function $\psi \in \mathbb{R}_+^C$ with $c \leq \psi(c)$ for all $c \in C$. Suppose that for any $c \in C$, $\sigma(c) \in \mathcal{S}$ is a $\psi(c)$-quasi-perfect segmentation for c. Then,*

1. $\int_{\mathcal{D}} D(p)\sigma(\mathrm{d}D|c) = D_0(p)$ *for all $p \in V$ and for all $c \in C$.*

2. $\sigma : C \to \Delta(\mathcal{D})$ *is measurable.*

3. $\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) = D_0(\psi(c))$ *for all $c \in C$.*

4. $\int_{\mathcal{D}} \overline{\boldsymbol{p}}_D(c) D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) = \int_{\{v \geq \psi(c)\}} v D_0(\mathrm{d}v)$ *for all* $c \in C$*.*

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

$$\int_{\mathcal{D}} D(p)\sigma(\mathrm{d}D|c) = D_0(p), \tag{2.20}$$

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

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

and thus

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

where the last equality follows from (2.1). This proves assertion 3. Finally, to prove assertion 4, consider any $c \in C$. First notice that if $D_0(c) = 0$, then assertion 4 clearly holds as both sides would be zero. Now suppose that $D_0(c) > 0$. The fact that $\sigma(c) \in \mathcal{S}$ is a $\psi(c)$-quasi-perfect segmentation for c ensures that $D_0(\psi(c)) > 0$. Then, for any $v \in [\psi(c), \overline{v}]$, let

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

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

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

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

$$K(A|B) = \int_{v \in A} m^v(B) H(\mathrm{d}v). \tag{2.21}$$

In particular, by definition of K and by (2.1), for any measurable set $B \subseteq [\psi(c), \overline{v}]$,

$$\int_{[\psi(c),\overline{v}]} m^v(B) H(\mathrm{d}v) = K([\psi(c), \overline{v}]|B) = \int_{\mathcal{D}} m^D(B) \sigma(\mathrm{d}D|c) = m^0(B). \tag{2.22}$$

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

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

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

$$\int_A m^v(B) H(\mathrm{d}v) = \int_B m^v(A) H(\mathrm{d}v). \tag{2.23}$$

As a result,

$$\begin{aligned}
\int_{\mathcal{D}} \overline{\boldsymbol{p}}_D(c) D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) &= \int_{\mathcal{D}} \overline{\boldsymbol{p}}_D(c) D(\psi(c))\sigma(\mathrm{d}D|c) \\
&= \int_{\mathcal{D}} \max(\mathrm{supp}(D)) m^D([\psi(c), \overline{v}])\sigma(\mathrm{d}D|c) \\
&= \int_{[\psi(c),\overline{v}]} vK(\mathrm{d}v|[\psi(c), \overline{v}]) \\
&= \int_{[\psi(c),\overline{v}]} vm^v([\psi(c), \overline{v}])H(\mathrm{d}v) \\
&= \int_{v\in[\psi(c),\overline{v}]} \int_{v'\in[\psi(c),\overline{v}]} vm^v(\mathrm{d}v')H(\mathrm{d}v) \\
&= \int_{v\in[\psi(c),\overline{v}]} v \left(\int_{v'\in[\psi(c),\overline{v}]} m^{v'}(\mathrm{d}v)H(\mathrm{d}v') \right) \\
&= \int_{[\psi(c),\overline{v}]} vD_0(\mathrm{d}v),
\end{aligned}$$

where the second equality follows from the fact that $\sigma(c)$ is a $\psi(c)$-quasi-perfect segmentation for c, the third equality follows from the definition of K, the fourth equality follows from (2.21), the sixth equality follows from (2.23), and the last equality follows from (2.22). This completes the proof. ■

Lemma 2.11. *For any nondecreasing function $\psi \in \mathbb{R}_+^C$ with $c \leq \psi(c)$ for all $c \in C$, suppose that $\sigma \in \mathcal{S}^C$ is a ψ-quasi-perfect scheme. Then for any $c, c' \in C$ with $c < c'$,*

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

Proof. Consider any $c, c' \in C$ such that $c < c'$. Notice that since $\sigma \in \mathcal{S}^C$ is a ψ-quasi-perfect scheme, by Lemma 2.10,

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

In addition, since $\sigma(c)$ is a $\psi(c)$-quasi-perfect segmentation for c, for any $z > c$ and for any

$D \in \operatorname{supp}(\sigma(c))$, if $D(c) > 0$ and $\max(\operatorname{supp}(D)) \geq z$, then $\overline{\boldsymbol{p}}_D(z) = \overline{\boldsymbol{p}}_D(c)$. On the other hand, if $D(c) > 0$ and $\max(\operatorname{supp}(D)) < z$, then $D(\overline{\boldsymbol{p}}_D(z)) = 0$. Also, notice that $D(c) = 0$ implies $D(z) = 0$. Together, if $z \leq \psi(c)$,

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

and if $z > \psi(c)$,

$$\begin{aligned}\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|c) =& \int_{\{D\in\mathcal{D}:\max(\operatorname{supp}(D))\geq z\}} D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c) \\ =& \int_{\mathcal{D}} D(z)\sigma(\mathrm{d}D|c) \\ =& D_0(z).\end{aligned}$$

As a result,

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

where the inequality follows from the fact that $z \leq \psi(z)$ and $\psi(c) \leq \psi(z)$ for all $z \in [c, c']$, which in turn relies on the hypothesis that $c \leq \psi(c)$ for all $c \in C$ and that ψ is nondecreasing. This completes the proof. ■

Lemma 2.12. *Consider any function $\psi \in \mathbb{R}_+^C$ with $c \leq \psi(c)$ for all $c \in C$. Given any $\{D_n\} \subset \mathcal{D}$ and $\{\sigma_n\} \subset \mathcal{S}_{D_n}^C$. Suppose that $\{\sigma_n\} \to \sigma$ pointwise and $\{D_n\} \to D_0$ for some $\sigma \in \Delta(\mathcal{D})^C$ and $D_0 \in \mathcal{D}$. Then $\sigma \in \mathcal{S}_{D_0}^C$. Moreover, suppose further that σ_n is a*

ψ-quasi-perfect scheme for all $n \in \mathbb{N}$. Then σ is a ψ-quasi-perfect scheme.

Proof. First notice that since $\sigma_n \in \mathcal{S}_{D_n}^C$ is a ψ-quasi-perfect scheme, Lemma 2.10 ensures that σ_n is measurable. Then, since $\{\sigma_n\} \to \sigma$ pointwise, σ is also measurable. Moreover, since $\{D_n\} \to D_0$ and $\{\sigma_n\} \to \sigma$, for any bounded continuous function $f : V \to \mathbb{R}$ and for any $c \in C$

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

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

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

This proves that $\sigma \in \mathcal{S}^C$.

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

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

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

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

However, by Lemma 2.7 and Lemma 2.8, for Lebesgue almost all $c \in C$,

$$\begin{aligned}&\int_{\mathcal{D}}(\overline{\boldsymbol{p}}_D(c)-\psi(c))D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c)\\ =&\int_{\mathcal{D}}\pi_D(c)\sigma(\mathrm{d}D|c)-(\psi(c)-c)\int_{\mathcal{D}}D(\overline{\boldsymbol{p}}_D(c))\sigma(\mathrm{d}D|c)\\ \geq&\lim_{n\to\infty}\int_{\mathcal{D}}\pi_D(c)\sigma_n(\mathrm{d}D|c)-\liminf_{n\to\infty}(\psi(c)-c)\int_{\mathcal{D}}D(\overline{\boldsymbol{p}}_D(c))\sigma_n(\mathrm{d}D|c)\\ =&\limsup_{n\to\infty}\left[\int_{\mathcal{D}}\pi_D(c)\sigma_n(\mathrm{d}D|c)-(\psi(c)-c)\int_{\mathcal{D}}D(\overline{\boldsymbol{p}}_D(c))\sigma_n(\mathrm{d}D|c)\right]\\ =&\limsup_{n\to\infty}\int_V(v-\psi(c))^+D_n(\mathrm{d}v)\\ =&\lim_{n\to\infty}\int_V(v-\psi(c))^+D_n(\mathrm{d}v)\\ =&\int_V(v-\psi(c))^+D_0(\mathrm{d}v),\end{aligned} \tag{2.24}$$

a contradiction. Here, the first inequality follows from the fact that $\{\sigma_n(c)\} \to \sigma(c)$, Lemma 2.7 and Lemma 2.8; the second equality follows from the properties of the $\liminf$ and $\limsup$ operators;[2] the third equality follows from the fact that $\sigma_n(c) \in \mathcal{S}_{D_n}$ and is a

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

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

$\psi(c)$-quasi-perfect segmentation for c; and the last two equalities follow from the fact that the function $(v - \psi(c))^+$ is bounded and continuous in v and that $\{D_n\} \to D_0$. Therefore, σ must be a ψ-quasi-perfect scheme. ■

D Proofs for Optimal Mechanisms

Proof of Theorem 2.1

In this section, I first prove Theorem 2.1 and obtain an upper bound for the data broker's revenue. That is, I first solve the relaxed problem where the prices are also contractable. To this end, I first introduce the revenue-equivalence formula under regime $\mathcal{P}$

Lemma 2.13. *Under regime $\mathcal{P}$, a mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ is incentive compatible if and only if*

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

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

2. *The function*

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

is nonincreasing.

Proof. For necessity, consider any incentive compatible mechanism $(\sigma, \tau, \boldsymbol{\gamma})$. Let

$$u(c,c') := \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p-c)D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c') \right) \sigma(\mathrm{d}D|c') - \tau(c'),\ \forall c, c' \in C$$

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

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

denote the producer's net profit when her marginal cost is c and reports c'. By incentive compatibility, for any $c \in C$

$$U(c) := u(c,c) \geq u(c,c'),$$

Since $u_1(c,c') = -\int_{\mathcal{D}}\int_{\mathbb{R}_+} D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\sigma(\mathrm{d}D|c')$ is bounded for all $c' \in C$, by the envelope theorem (Milgrom and Segal, 2002), for any $c \in C$

$$U(c) = U(\bar{c}) - \int_c^{\bar{c}} u_1(z,z)\,\mathrm{d}z = U(\bar{c}) + \int_c^{\bar{c}} \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c) \right) \sigma(\mathrm{d}D|z) \right) \mathrm{d}z.$$

Also, by definition,

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

Rearranging, and letting $\bar{\tau} := U(\bar{c})$, for any $c \in C$,

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

which proves assertion 1. Furthermore, since $u(c,c')$ is affine in c for all c', U is convex as it is a pointwise maximum of a family of affine functions. Therefore, it's (almost everywhere) derivative

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

is nondecreasing. This proves assertion 2.

Conversely, given a mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ that satisfies assertions 1 and 2, for any $c, c' \in C$,

$$\begin{aligned}
& u(c,c) - u(c,c') \\
=& \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p-c) D(p) \boldsymbol{\gamma}(\mathrm{d}p|D,c) \right) \sigma(\mathrm{d}D|c) - \tau(c) \right) \\
& - \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p-c) D(p) \boldsymbol{\gamma}(\mathrm{d}p|D,c') \right) \sigma(\mathrm{d}D|c') - \tau(c') \right) \\
=& \int_c^{c'} \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p|D,z) \right) \sigma(\mathrm{d}D|z) \, \mathrm{d}z - \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p|D,c') \right) \sigma(\mathrm{d}D|c')(c'-c) \\
=& \int_c^{c'} \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p|D,z) \right) \sigma(\mathrm{d}D|z) - \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p|D,c') \right) \sigma(\mathrm{d}D|c') \right) \mathrm{d}z \\
\geq & 0,
\end{aligned}$$

where the inequality follows from assertion 2. As such, the mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ is indeed incentive compatible. ■

With Lemma 2.13, the producer's expected profit under an incentive compatible mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ can be written as

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

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

$$U(\overline{c}) + \int_c^{\overline{c}} \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p|D,z) \right) \sigma(\mathrm{d}D|z) \, \mathrm{d}z \geq \pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z)) \, \mathrm{d}z.$$

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

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

Therefore, the data broker's revenue maximization problem can be written under regime $\mathcal{P}$ as

$$\begin{aligned}
&\sup_{\sigma,\boldsymbol{\gamma}} \int_C \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \phi_G(c)) D(p) \boldsymbol{\gamma}(\mathrm{d}p|D,c) \right) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \\
&\text{s.t. } c \mapsto \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c) \right) \sigma(\mathrm{d}D|c) \text{ is nonincreasing,} \\
&\quad \int_c^{\overline{c}} \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,z) \right) \sigma(\mathrm{d}D|z)\,\mathrm{d}z \geq \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in C,
\end{aligned}$$

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

In what follows, let Γ be the collection of transition kernels that maps from $\mathcal{D}$ to $\Delta(\mathbb{R}_+)$. Let $\bar{s} \in \mathcal{S}$ denote the value-revealing segmentation and let $\bar{\sigma} \in \mathcal{S}^C$ be the segmentation scheme such that $\bar{\sigma}(c) = \bar{s}$ for all $c \in C$. Furthermore, for any $q \in [0,1]$, let $\rho_q := D_0^{-1}(q)$, where D_0^{-1} is defined by (2.14). Notice that by definition of D_0^{-1},

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

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

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

On the other hand, if $D_0(\rho_q) > D_0(\rho_q^+)$, then define $\tilde{\gamma}^q \in \Delta(\mathbb{R}_+)^V$ as

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

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

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

for any measurable $A \subseteq V$ and for any $D \in \mathcal{D}$. In other words, combining the segmentation $\bar{s}$ and the randomized price γ^q, this means that when the producer uses the randomized price γ^q under segmentation $\bar{s}$, then all the consumers with values above the $(1-q)$th-percentile buy the product by paying exactly their values while the other consumers do not buy, so that the traded quantity is exactly q. That is,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \gamma^q(\mathrm{d}p|D) \right) \bar{s}(\mathrm{d}D) = q. \tag{2.25}$$

With this notation, I now introduce the second auxiliary lemma.

Lemma 2.14. *For any $q \in [0,1]$, let $\bar{R}(q)$ be the value of the maximization problem*

$$\begin{aligned} \sup_{s \in \mathcal{S}, \gamma \in \Gamma} & \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} p D(p) \boldsymbol{\gamma}(\mathrm{d}p) \right) s(\mathrm{d}D) \\ \text{s.t.} & \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p) \right) s(\mathrm{d}D) \leq q. \end{aligned} \tag{2.26}$$

Then

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

where D_0^{-1} is defined by (2.14). *Moreover,* $(\bar{s}, \gamma^q)$ *is a solution of* (2.26).

Proof. Consider the dual problem of (2.26). That is, for any $\nu \geq 0$, let

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

Clearly, $d(\nu) \geq \bar{R}(q)$ for any $\nu \geq 0$. Thus, by weak duality, to solve (2.26), it suffices to find ν^* and (s^*, γ^*) such that (s^*, γ^*) is feasible in the primal problem (2.26), (s^*, γ^*) solves the dual problem

$$\sup_{s \in \mathcal{S}, \gamma \in \Gamma} \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \nu^*) D(p) \gamma(\mathrm{d}p|D) \right) s(\mathrm{d}D) \tag{2.27}$$

and that the complementary slackness condition

$$\nu^* \left[q - \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} p D(p) \gamma^*(\mathrm{d}p|D) \right) s^*(\mathrm{d}D) \right] = 0 \tag{2.28}$$

holds. Since this would imply that

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

and hence (s^*, γ^*) must be a solution to (2.26).

To this end, let

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

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

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

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

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

Therefore, $(\bar{s}, \gamma^q)$ solves the dual problem (2.27). On the other hand, by (2.25),

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

Together with $\bar{s} \in \mathcal{S}$ and $\gamma^q \in \Gamma$, it follows that $(\bar{s}, \gamma^q)$ is feasible in the primal problem (2.26) and, furthermore, the complementary slackness condition (2.28) also holds. Thus, $(\bar{s}, \gamma^q)$ is a solution to the primal problem (2.26).

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

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

This completes the proof. ■

With the two auxiliary lemmas above, the broker's problem under regime $\mathcal{P}$ can be effectively reduced to a one-dimensional screening problem with type-dependent individual rationality constraints. As stated by Lemma 2.15 below.

Lemma 2.15. *Under regime $\mathcal{P}$, there exists an incentive feasible mechanism that maximizes the data broker's revenue. Furthermore, the data broker's optimal revenue is*

$$\begin{aligned} R^* = \max_{\boldsymbol{q} \in \mathcal{Q}} & \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \phi_G(c)) \,\mathrm{d}q \right) G(\mathrm{d}c) - \pi_0(\bar{c}) \\ \text{s.t. } & \pi_0(\bar{c}) + \int_c^{\bar{c}} \boldsymbol{q}(z) \,\mathrm{d}z \geq \pi_0(c), \, \forall c \in C, \end{aligned} \tag{2.29}$$

where $\mathcal{Q}$ is the collection of nonincreasing functions in $[0,1]^C$.

Proof. Consider any incentive feasible mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ under regime $\mathcal{P}$, I will first show that there exists $\boldsymbol{q} \in [0,1]^C$ such that the mechanism $(\bar{\sigma}, \tau^{\boldsymbol{q}}, \boldsymbol{\gamma}^{\boldsymbol{q}})$ generates weakly higher revenue for the data broker and is incentive feasible, where

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

and $\tau^{\boldsymbol{q}}$ is the transfer determined by $(\bar{\sigma}, \boldsymbol{\gamma}^{\boldsymbol{q}})$ when the constant is chosen so that $U(\overline{c}) = \pi_0(\overline{c})$ according to Lemma 2.13. Then, I will show that maximizing revenue across the family of incentive feasible mechanisms $(\bar{\sigma}, \tau^{\boldsymbol{q}}, \boldsymbol{\gamma}^{\boldsymbol{q}})$ is equivalent to solving (2.29). Finally, the existence of the optimal mechanism can then be ensured by the existence of the solution of (2.29).

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

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

By Lemma 2.13, incentive compatibility of $(\sigma, \tau, \boldsymbol{\gamma})$ implies that $\boldsymbol{q} \in [0,1]^C$ is nonincreasing and, by (2.25), for any $c \in C$,

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

Thus, by Lemma 2.14, $(\bar{\sigma}(c), \boldsymbol{\gamma}^{\boldsymbol{q}}(c))$ solves the problem (2.26) with the quantity constraint being $\boldsymbol{q}(c)$ and hence, since $(\sigma(c), \boldsymbol{\gamma}(c))$ is also feasible in this problem,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} pD(p) \boldsymbol{\gamma}(\mathrm{d}p|D, c) \right) \sigma(\mathrm{d}D|c) \leq \int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} pD(p) \boldsymbol{\gamma}^{\boldsymbol{q}}(\mathrm{d}p|D, c) \right) \bar{\sigma}(\mathrm{d}D|c) = \bar{R}(\boldsymbol{q}(c)). \tag{2.30}$$

As a result,

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

where the first and the third equalities follows from the definition of $\boldsymbol{q}(c)$ and from (2.25), and the inequality and the second equality follows from (2.30). Moreover, by (2.25), since $\boldsymbol{q}$ is nonincreasing, the function

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

is nonincreasing. Together with Lemma 2.13 and individual rationality of $(\sigma, \tau, \boldsymbol{\gamma})$, for any $c \in C$,

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

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

Now notice that by (2.25) and Lemma 2.14, for any $\boldsymbol{q} \in [0,1]^C$ and for any $c \in C$,

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

On the other hand, by (2.25) and by Lemma 2.14, $(\bar{\sigma}, \tau^{\boldsymbol{q}}, \boldsymbol{\gamma}^{\boldsymbol{q}})$ is incentive feasible if and only if $\boldsymbol{q}$ is nonincreasing and

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

Therefore, maximizing revenue among all incentive feasible mechanism is equivalent to solving (2.29).

Finally, notice that for the maximization problem (2.29), endow the set of nonincreasing functions with the L^1 norm. Helly's selection theorem and the Lebesgue dominated convergence theorem then imply that this set is compact. Furthermore, since for any sequence $\{q_n\}$ that converges to q in the L^1 norm, there exists a subsequence $\{q_{n_k}\}$ that converges to q pointwise, by the Lebesgue dominated convergence again, the objective function of (2.29) is continuous and the feasible set is a closed subset of a compact set, and hence is itself compact. Together, the problem (2.29) has a solution. This completes the proof. ∎

With Lemma 2.15, the data broker's revenue maximization problem can be solved explicitly.

Proof of Theorem 2.1. Recall that $\mathcal{Q} \subset [0,1]^C$ denotes the set of nonincreasing functions $[0,1]^C$. Using (2.15) and Lemma 2.15, rewrite the data broker's problem (2.29) as

$$\begin{aligned} &\sup_{\boldsymbol{q} \in \mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \phi_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \\ &\text{s.t. } \pi_0(\overline{c}) + \int_c^{\overline{c}} \boldsymbol{q}(z)\,\mathrm{d}z \geq \pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\boldsymbol{p}_0(z))\,\mathrm{d}z. \end{aligned} \tag{2.31}$$

Let R^* be the value of (2.31) and write objective function of (2.31) as

$$R(\boldsymbol{q}) := \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \phi_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) - \pi_0(\overline{c}),\ \forall q \in \mathcal{Q}.$$

Consider the dual problem of (2.31). That is, for any Borel measure μ on C, let

$$d(\mu) := \sup_{\boldsymbol{q}\in\mathcal{Q}} \left[R(\boldsymbol{q}) + \int_C \left(\int_c^{\overline{c}} (\boldsymbol{q}(z) - D_0(\overline{\boldsymbol{p}}_0(z)))\,\mathrm{d}z \right) \mu(\mathrm{d}c) \right]$$

and let

$$d^* := \inf_{\mu} d(\mu),$$

where the infimum is taken over all Borel measures on C. Then clearly

$$d^* \geq R^*.$$

Moreover, if there exists a Borel measure μ^* on C and a feasible choice $\boldsymbol{q}^* \in \mathcal{Q}$ of the primal problem (2.31) such that $d(\mu^*) = R(\boldsymbol{q}^*)$, then

$$R^* \leq d^* \leq d(\mu^*) = R(\boldsymbol{q}^*) \leq R^*,$$

and hence $\boldsymbol{q}^* \in \mathcal{Q}$ is a solution of the primal problem (2.31). As a result, to solve (2.31), it suffices to find a Borel measure μ^* and a feasible $\boldsymbol{q}^* \in \mathcal{Q}$ such that $\boldsymbol{q}^*$ is a solution of

$$\sup_{\boldsymbol{q}\in\mathcal{Q}} \left[R(\boldsymbol{q}) + \int_C \left(\int_c^{\overline{c}} (\boldsymbol{q}(z) - D_0(\overline{\boldsymbol{p}}_0(z)))\,\mathrm{d}z \right) \mu^*(\mathrm{d}c) \right] \tag{2.32}$$

and that

$$\int_C \left(\int_c^{\overline{c}} (\boldsymbol{q}^*(z) - D_0(\overline{\boldsymbol{p}}_0(z)))\,\mathrm{d}z \right) \mu^*(\mathrm{d}c) = 0. \tag{2.33}$$

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

$$M^*(c) := \lim_{z \downarrow c} g(z)(\phi_G(z) - \overline{\boldsymbol{p}}_0(z))^+, \ \forall c \in C. \tag{2.34}$$

By definition, M^* is right-continuous. Also, by Assumption 2.1, M^* is nondecreasing and hence M^* a CDF. Let μ^* be the Borel measure induced by M^*. Notice that $\mathrm{supp}(\mu^*) = [c^*, \overline{c}]$, where

$$c^* := \inf\{c \in C : \phi_G(c) > \overline{\boldsymbol{p}}_0(c)\}.$$

Notice that for any $q \in \mathcal{Q}$, by interchanging the order of integrals,

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

and therefore for any $\boldsymbol{q} \in \mathcal{Q}$,

$$\begin{aligned}
&R(\boldsymbol{q}) + \int_C \left(\int_c^{\overline{c}} (\boldsymbol{q}(z) - D_0(\overline{\boldsymbol{p}}_0(z))) \,\mathrm{d}z \right) \mu^*(\mathrm{d}c) \\
&= \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \phi_G(c)) \,\mathrm{d}q \right) G(\mathrm{d}c) - \pi_0(\overline{c}) + \int_C M^*(c)(\boldsymbol{q}(c) - D_0(\overline{\boldsymbol{p}}_0(c))) \,\mathrm{d}c \\
&= \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\phi}_G(c)) \,\mathrm{d}q \right) G(\mathrm{d}c) - \pi_0(\overline{c}) - \int_C M^*(c) D_0(\overline{\boldsymbol{p}}_0(c)) \,\mathrm{d}c,
\end{aligned}$$

where $\overline{\phi}_G := \min\{\phi_G, \overline{\boldsymbol{p}}_0\}$. As only the first term depends on the choice variable $\boldsymbol{q}$, the solution of (2.32) is the same as the solution of

$$\sup_{\boldsymbol{q} \in \mathcal{Q}} \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\phi}_G(c)) \,\mathrm{d}q \right) G(\mathrm{d}c). \tag{2.35}$$

To solve (2.35), consider first the case when G is regular so that ϕ_G is nondecreasing. In this case, notice that for any $\boldsymbol{q} \in \mathcal{Q}$ and for all $c \in C$,

$$\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\phi}_G(c))\,\mathrm{d}q \le \int_0^{\boldsymbol{q}^{\overline{\phi}_G}(c)} (D_0^{-1}(q) - \overline{\phi}_G(c))\,\mathrm{d}q,$$

where for any function ψ,

$$\boldsymbol{q}^{\psi}(c) := \sup\{y \in [0,1] : D_0^{-1}(q) \ge \psi(c)\}.$$

Moreover, the function $\overline{\phi}_G$ is nondecreasing since both ϕ_G and $\overline{\boldsymbol{p}}_0$ are nondecreasing, also, the function D_0^{-1} is nonincreasing. As a result, $\boldsymbol{q}^{\overline{\phi}_G} \in \mathcal{Q}$ is a solution of (2.35) and thus is a solution of (2.32). Furthermore, by definition, for any $c \in C$,

$$D_0(\overline{\phi}_G(c)) = \boldsymbol{q}^{\overline{\phi}_G}(c)$$

In particular, since $\overline{\phi}_G \le \overline{\boldsymbol{p}}_0$, $\boldsymbol{q}^{\overline{\phi}_G}$ is feasible in the primal problem (2.31). That is, for any $c \in C$,

$$\begin{aligned}\pi_0(\overline{c}) + \int_c^{\overline{c}} \boldsymbol{q}^{\overline{\phi}_G}(z)\,\mathrm{d}z \ge& \pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\phi}_G(z))\,\mathrm{d}z \\ \ge& \pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z.\end{aligned}$$

Finally, notice that since $\overline{\phi}_G(c) = \overline{\boldsymbol{p}}_0(c)$ for all $c \in (c^*, \overline{c}]$ and since $M^*(c) = 0$ for all $c \in [\underline{c}, c^*)$, the complementary slackness condition (2.33) also holds for $\overline{\phi}_G$ and μ^*. That is,

$$\begin{aligned}\int_C M^*(c)(\boldsymbol{q}^{\overline{\phi}_G}(c) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c =& \int_{c^*}^{\overline{c}} M^*(c)(D_0(\overline{\phi}_G(c)) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c \\ =& \int_{c^*}^{\overline{c}} M^*(c)(D_0(\overline{\boldsymbol{p}}_0(c)) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c \\ =& 0.\end{aligned}$$

Together, when G is regular, $\boldsymbol{q}^{\overline{\phi}_G} \equiv D_0 \circ \overline{\phi}_G$ solves the primal problem (2.31).

Now consider the case for a general G, to solve (2.35) Let φ_G be the ironed virtual cost. That is, φ_G is defined by the following procedure: Let $h : [0,1] \to \mathbb{R}_+$ be defined as

$$h(q) := \phi_G(G^{-1}(q)) = G^{-1}(q) + \frac{q}{g(G^{-1}(q))}.$$

and define $H : [0,1] \to \mathbb{R}_+$, $K : [0,1] \to \mathbb{R}_+$ as

$$H(q) := \int_0^q h(s)\,\mathrm{d}s$$

and

$$K := \mathrm{co}(H).$$

Finally, for every $q \in [0,1]$ let $k(q) := K'(q)$. φ_G is then defined as

$$\varphi_G(c) := k(G(c)).$$

Also, let $\overline{\varphi}_G := \min\{\varphi_G, \overline{\boldsymbol{p}}_0\}$. With this definition, notice that for any $\boldsymbol{q} \in \mathcal{Q}$,

$$\begin{aligned}
&\int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\phi}_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) \\
&= \int_C \left(\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\varphi}_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) + \int_C (\overline{\varphi}_G(c) - \overline{\phi}_G(c))\boldsymbol{q}(c) G(\mathrm{d}c). \qquad (2.36)
\end{aligned}$$

Moreover, using integration by parts, since $K(0) = H(0)$ and $K(1) = H(1)$,

$$\begin{aligned}
\int_C (\overline{\varphi}_G(c) - \overline{\phi}_G(c))\boldsymbol{q}(c) G(\mathrm{d}c) &= \int_C (\varphi_G(c) - \phi_G(c))\boldsymbol{q}(c) G(\mathrm{d}c) \\
&= -\int_C (K(G(c)) - H(G(c)))\boldsymbol{q}(\mathrm{d}c) \qquad (2.37) \\
&\leq 0,
\end{aligned}$$

where the first equality follows from the observation that $\overline{\phi}_G(c) = \overline{\varphi}_G(c) = \phi_G(c) = \varphi_G(c) = \overline{\boldsymbol{p}}_0(c)$ for all $c \geq c^*$, which is due to Assumption 2.1, and the inequality follows from the fact that $K = \text{co}(H)$ and that $\boldsymbol{q}$ is nonincreasing for any $\boldsymbol{q} \in \mathcal{Q}$.

On the other hand, notice that for any $\boldsymbol{q} \in \mathcal{Q}$ and for all $c \in C$,

$$\int_0^{\boldsymbol{q}(c)} (D_0^{-1}(q) - \overline{\varphi}_G(c))\,\mathrm{d}q \leq \int_0^{\boldsymbol{q}^{\overline{\varphi}_G}(c)} (D_0^{-1}(q) - \overline{\varphi}_G(c))\,\mathrm{d}q,$$

and hence

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

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

$$\int_C (\overline{\varphi}_G(c) - \overline{\phi}_G(c))\boldsymbol{q}^{\overline{\varphi}_G}(c) G(\mathrm{d}c) = -\int_C (K(G(c)) - H(G(c)))\boldsymbol{q}^{\overline{\varphi}_G}(\mathrm{d}c) = 0. \tag{2.38}$$

Together with (2.36) and (2.37), for any $\boldsymbol{q} \in \mathcal{Q}$,

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

Also, since $\overline{\varphi}_G$ is nondecreasing by definition, $\boldsymbol{q}^{\overline{\varphi}_G}$ is indeed a solution of (2.35) and hence a solution of (2.32).

Moreover, by definition, for any $c \in C$,

$$\boldsymbol{q}^{\overline{\varphi}_G}(c) = D_0(\overline{\varphi}_G(c)). \tag{2.39}$$

Thus, since $\overline{\varphi}_G \leq \overline{\boldsymbol{p}}_0$,

$$\begin{aligned}\pi_0(\overline{c}) + \int_c^{\overline{c}} \boldsymbol{q}^{\overline{\varphi}_G}(z)\,\mathrm{d}z =&\pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\varphi}_G(c))\,\mathrm{d}z \\ \geq&\pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in C\end{aligned}$$

and hence $\boldsymbol{q}^{\overline{\varphi}_G} \in \mathcal{Q}$ is feasible choice in the primal problem (2.31). Furthermore, since $M^*(c) = 0$ for all $c \in [\underline{c}, c^*)$ and since $\overline{\varphi}_G(c) = \overline{\boldsymbol{p}}_0(c)$ for all $c \in (c^*, \overline{c}]$, the complementary slackness condition (2.33) is also satisfied. That is,

$$\begin{aligned}\int_C M^*(c)(\boldsymbol{q}^{\overline{\varphi}_G}(c) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c =& \int_C M^*(c)(D_0(\overline{\varphi}_G(c)) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c \\ =& \int_{c^*}^{\overline{c}} M^*(c)(D_0(\overline{\boldsymbol{p}}_0(c)) - D_0(\overline{\boldsymbol{p}}_0(c)))\,\mathrm{d}c \\ =&0.\end{aligned}$$

Together, $\boldsymbol{q}^{\overline{\varphi}_G} \equiv D_0 \circ \overline{\varphi}_G$ is indeed a solution of (2.31).

Finally, by (2.39), it then follows that

$$\begin{aligned}R^* =& \int_C \left(\int_0^{\boldsymbol{q}^{\overline{\varphi}_G}(c)} (D_0^{-1}(q) - \phi_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \\ =& \int_C \left(\int_0^{D_0(\overline{\varphi}_G(c))} (D_0^{-1}(q) - \phi_G(c))\,\mathrm{d}q \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \\ =& \int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \pi_0(\overline{c}).\end{aligned}$$

The see that any solution under regime $\mathcal{P}$ must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G almost all $c \in C$, consider any optimal mechanism $(\sigma, \tau, \boldsymbol{\gamma})$ under regime $\mathcal{P}$. By optimality, it must be that $\mathbb{E}_G[\tau(c)] = R^*$ and that the indirect utility of the producer with marginal cost $\overline{c}$ is $\pi_0(\overline{c})$. Also, by Lemma 2.13, it must be that

$$\begin{aligned}R^* =&\mathbb{E}_G[\tau(c)]\\ =&\int_C\left(\int_{\mathcal{D}}\left(\int_{\mathbb{R}_+}(p-\phi_G(c))D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\right)\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)-\pi_0(\overline{c}),\end{aligned}$$

which implies that

$$\int_C\left(\int_{\mathcal{D}}\left(\int_{\mathbb{R}_+}(p-\phi_G(c))D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\right)\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)=\int_C\left(\int_{\{v\geq\overline{\varphi}_G(c)\}}(v-\phi_G(c))D_0(\mathrm{d}v)\right)G(\mathrm{d}c) \tag{2.40}$$

Thus,

$$\begin{aligned}&\int_C\left(\int_{\mathcal{D}}\left(\int_{\mathbb{R}_+}(p-\overline{\varphi}_G(c))D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\right)\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)\\ &+\int_C(\overline{\varphi}_G(c)-\phi_G(c))\left(\int_{\mathcal{D}}\left(\int_{\mathbb{R}_+}D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\right)\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c)\\ =&\int_C\left(\int_{\mathcal{D}}\left(\int_{\mathbb{R}_+}(p-\phi_G(c))D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c)\right)\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c) \qquad (2.41)\\ =&\int_C\left(\int_{\{v\geq\overline{\varphi}_G(c)\}}(v-\phi_G(c))D_0(\mathrm{d}v)\right)G(\mathrm{d}c)\\ =&\int_C\left(\int_{\{v\geq\overline{\varphi}_G(c)\}}(v-\overline{\varphi}_G(c))D_0(\mathrm{d}v)\right)G(\mathrm{d}c)+\int_C(\overline{\varphi}_G(c)-\phi_G(c))D_0(\overline{\varphi}_G(c))G(\mathrm{d}c)\end{aligned}$$

where the second equality follows from (2.40). Moreover, since for any $c \in C$,

$$\begin{aligned}
&\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \overline{\varphi}_G(c)) D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}D|c) \\
&\leq \int_{\mathcal{D}} \max_{p \in \mathbb{R}_+} [(p - \overline{\varphi}_G(c)) D(p)] \sigma(\mathrm{d}D|c) \qquad (2.42) \\
&\leq \int_V (v - \overline{\varphi}_G(c))^+ D_0(\mathrm{d}v),
\end{aligned}$$

it must be that

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

Together with (2.37) and (2.38), we have

$$\begin{aligned}
&\int_C (\overline{\phi}_G(c) - \phi_G(c)) \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) \\
&\geq \int_C (\overline{\varphi}_G(c) - \phi_G(c)) \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, c) \right) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) \\
&\geq \int_C (\overline{\varphi}_G(c) - \phi_G(c)) D_0(\overline{\varphi}_G(c)) G(\mathrm{d}c) \qquad (2.43) \\
&= \int_C (\overline{\phi}_G(c) - \phi_G(c)) D_0(\overline{\varphi}_G(c)) G(\mathrm{d}c).
\end{aligned}$$

Furthermore, since $\overline{\phi}_G(c) = \overline{\boldsymbol{p}}_0(c) \leq \phi_G(c)$ for all $c \in (c^*, \overline{c}]$ and $\overline{\phi}_G(c) = \phi_G(c)$, by the definition of M^* given by (2.34), and by using integration by parts, (2.43) is equivalent to

$$\int_C \left(\int_c^{\overline{c}} \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \boldsymbol{\gamma}(\mathrm{d}p | D, z) \right) \sigma(\mathrm{d}D|z) - D_0(\overline{\boldsymbol{p}}_0(z)) \right) \mathrm{d}z \right) M^*(\mathrm{d}c) \leq 0$$

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

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

Thus,

$$\int_C \left(\int_c^{\bar{c}} \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,z) \right) \sigma(\mathrm{d}D|z) - D_0(\overline{\boldsymbol{p}}_0(z)) \right) \mathrm{d}z \right) M^*(\mathrm{d}c) = 0$$

and hence

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

which, together with (2.41), implies that

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

Moreover, by (2.42), it then follows that for G-almost all $c \in C$,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - \overline{\varphi}_G(c)) D(p)\boldsymbol{\gamma}(\mathrm{d}p|D,c) \right) \sigma(\mathrm{d}D|c) = \int_V (v - \overline{\varphi}_G(c))^+ D_0(\mathrm{d}v), \tag{2.44}$$

which implies that for G-almost all $c \in C$, $(\sigma(c), \boldsymbol{\gamma}(c))$ must induce perfect price discrimination for the economy where the producer's marginal cost is $\overline{\varphi}_G(c)$, or equivalently, $(\sigma(c), \boldsymbol{\gamma}(c))$ must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G-almost all $c \in C$. This completes the proof. ■

Proof of Lemma 2.1

Proof of Lemma 2.1. For necessity, consider any incentive compatible mechanism (σ, τ). First notice that, by Lemma 2.4, $\pi_D : C \to \mathbb{R}_+$ is convex and continuous on C for any $D \in \mathcal{D}$ and

$$\pi'_D(c) = -D(\boldsymbol{p}_D(c))$$

for any $\boldsymbol{p} \in \boldsymbol{P}$ and for almost all $c \in C$. Moreover, since for any $D \in \mathcal{D}$ and for any $\boldsymbol{p} \in \boldsymbol{P}$

$$\left|\pi'_D(c)\right| = \left|D(\boldsymbol{p}_D(c))\right| \leq 1,$$

for almost all $c \in C$, the order of integral and differential can be interchanged. That is, for any $c, c' \in C$,

$$\frac{\mathrm{d}}{\mathrm{d}c} \int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c') = \int_{\mathcal{D}} \pi'_D(c)\sigma(\mathrm{d}D|c') = -\int_{\mathcal{D}} D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c'). \tag{2.45}$$

As such, if Π is defined as

$$\Pi(c, c') := \int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c'),\ \forall c, c' \in C,$$

then $\Pi(\cdot, c')$ is convex for all $c' \in C$. Moreover, (2.45) implies that

$$\Pi_1(c, c') = -\int_{\mathcal{D}} D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c'), \tag{2.46}$$

for any $c' \in C$, any $\boldsymbol{p} \in \boldsymbol{P}$, and for almost all $c \in C$. Now let $u(c, c') := \Pi(c, c') - t(c')$ for all $c, c' \in C$ be a producer's interim expected profit if her report is c' and marginal cost is c. By the Lebesgue dominated convergence theorem, $u(\cdot, c')$ is convex and absolutely continuous on C for all $c' \in C$ as π_D is absolutely continuous for all $D \in \mathcal{D}$. Furthermore, since the mechanism (σ, τ) is incentive compatible,

$$U(c) := u(c, c) \geq u(c, c'),\ \forall c, c' \in C.$$

By the envelope theorem again,

$$U(c) = U(\overline{c}) - \int_c^{\overline{c}} u_1(z, z)\,\mathrm{d}z.$$

Moreover, since for any $\boldsymbol{p} \in \boldsymbol{P}$ and for almost all $c \in C$,

$$u_1(c,c) = \Pi_1(c,c) = -\int_{\mathcal{D}} D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c), \tag{2.47}$$

it must be that for any $\boldsymbol{p} \in \boldsymbol{P}$ and for any $c \in C$,

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

Rearranging, and use the definition of Π, it follows that

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

which proves assertion 1. Furthermore, by incentive compatibility, for any $c, c' \in C$,

$$\begin{aligned}
0 \leq & U(c) - u(c,c') \\
= & U(c) - (\Pi(c,c') - \tau(c')) \\
= & U(c) - (\Pi(c,c') - \Pi(c',c')) - U(c') \\
= & (U(c) - U(c')) + (\Pi(c',c') - \Pi(c,c')) \\
= & \int_c^{c'} (u_1(z,z) + \Pi_1(z,c'))\,\mathrm{d}z \\
= & \int_c^{c'} \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(z))\sigma(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\boldsymbol{p}_D(z))\sigma(\mathrm{d}D|c') \right) \mathrm{d}z \\
= & \int_c^{c'} \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(z))(\sigma(\mathrm{d}D|z) - \sigma(\mathrm{d}D|c')) \right) \mathrm{d}z,
\end{aligned}$$

for any $\boldsymbol{p} \in \boldsymbol{P}$, where the forth equality follows from the fundamental theorem of calculus, and the last equality follows from (2.46) and (2.47). This proves assertion 2.

Conversely, suppose that a mechanism (σ, τ) satisfies assertions 1 and 2. Then, for any

$c, c' \in C$,

$$
\begin{aligned}
&(\Pi(c,c) - \tau(c)) - (\Pi(c,c') - \tau(c')) \\
=&\Pi(c,c) - \tau(c) - (\Pi(c,c') - \Pi(c',c')) - (\Pi(c',c') - \tau(c')) \\
=&\int_c^{c'} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) + \Pi_1(z,c') \right) \mathrm{d}z \\
=&\int_c^{c'} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) - \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|c') \right) \mathrm{d}z \\
=&\int_c^{c'} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))(\sigma(\mathrm{d}D|z) - \sigma(\mathrm{d}D|c')) \right) \mathrm{d}z \\
\geq&0,
\end{aligned}
$$

where the second equality follows from assertion 1 and the fundamental theorem of calculus, the last equality follows from (2.46), and the last inequality follows from assertion 2. As such, the mechanism (σ, τ) is incentive compatible. ■

Proof of Lemma 2.2

Proof of Lemma 2.2. Given any nondecreasing function $\psi \in \mathbb{R}_+^C$, and any ψ-quasi-perfect segmentation $\sigma \in \mathcal{S}$, suppose that for any $c \in C$,

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

for Lebesgue almost all $z \in [\underline{c}, c]$ and for all $D \in \text{supp}(\sigma(c))$. Then, for any $c, c' \in C$ with $c < c'$,

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

where the second equality follows from assertion 3 of Lemma 2.10, the inequality follows from (2.5), and the third equality follows from $\sigma(z) \in \mathcal{S}$ for all $z \in [c, c']$. Together with Lemma 2.11, this implies that

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

for all $c, c' \in C$. Therefore, by Lemma 2.1, there exists a transfer τ such that (σ, τ) is incentive compatible, as desired. ■

Proof of Lemma 2.3

Proof of Lemma 2.3. I first prove the lemma for D_0 being a step function. Consider step function $D \in \mathcal{D}$ and any $\psi \in \mathbb{R}_+^C$ such that $c \leq \psi(c) \leq \overline{\boldsymbol{p}}_D(c)$ for all $c \in C$ and fix any $c \in C$, let

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

and let

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

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

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

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

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

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

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

By construction,

$$\sum_{v \in V^+} \alpha^*(v) = \sum_{v \in V^+} \beta^*(v|v') = 1 \tag{2.48}$$

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

$$\sum_{v \in V^+} \hat{m}^v(v') = m^D(v'), \forall v' \in \text{supp}(D). \tag{2.49}$$

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

$$\sum_{v \geq \psi(c)} (v-\hat{c})m^D(v) \geq \sum_{v \geq \overline{\boldsymbol{p}}_D(\hat{c})} (v-\hat{c})m^D(v) \geq (\overline{\boldsymbol{p}}_D(\hat{c})-\hat{c})D(\overline{\boldsymbol{p}}_D(\hat{c})) = (\underline{\boldsymbol{p}}_D(\hat{c})-\hat{c})D(\underline{\boldsymbol{p}}_D(\hat{c})). \tag{2.50}$$

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

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

where the first equality follows from (2.49), the first inequality follows from (2.50), the second inequality follows from the fact that $\underline{\boldsymbol{p}}_D(\hat{c}) \in \boldsymbol{P}_D(\hat{c})$, and the last inequality follows from $\underline{\boldsymbol{p}}_D(\hat{c}) \geq \hat{c}$. As such, for any $v' \in \text{supp}(D)$ with $\underline{\boldsymbol{p}}_D(\hat{c}) \leq v' < \psi(c)$ and for any $v \in V^+$, if

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

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

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

which in turn implies that

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

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

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

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

$$(v'-\hat{c}) \sum_{\hat{v}\geq v'} \hat{m}^v(\hat{v}) \leq (v-\hat{c})\hat{m}^v(v). \tag{2.51}$$

Together with (2.48), this also ensures that

$$\alpha^* \in \Delta(V^+) \tag{2.52}$$

and

$$\beta^*(v') \in \Delta(V^+), \tag{2.53}$$

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

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

$$\sum_{\hat{v} \geq v'} \hat{m}^v(\hat{v}) = \alpha^*(v) \sum_{v' \leq \hat{v} < \underline{\boldsymbol{p}}_D(\hat{c})} m^D(\hat{v}) + \sum_{\hat{v} \geq \underline{\boldsymbol{p}}_D(\hat{c})} \hat{m}^v(\hat{v}) = \alpha^*(v) \sum_{\hat{v} \geq v'} m^D(\hat{v}). \tag{2.54}$$

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

$$\begin{aligned}
(v' - \hat{c}) \sum_{\hat{v} \geq v'} \hat{m}^v(\hat{v}) =& \alpha^*(v)(v' - \hat{c}) D(v') \\
\leq& \alpha^*(v)(\underline{\boldsymbol{p}}_D(\hat{c}) - \hat{c}) D(\underline{\boldsymbol{p}}_D(\hat{c})) \\
=& (\underline{\boldsymbol{p}}_D(\hat{c}) - \hat{c}) \sum_{\hat{v} \geq \underline{\boldsymbol{p}}_D(\hat{c})} \hat{m}^v(\hat{v}) \\
\leq& (v - \hat{c}) \hat{m}^v(v),
\end{aligned} \tag{2.55}$$

where both equalities follow from (2.54), the first inequality follows from the fact that $\underline{\boldsymbol{p}}_D(\hat{c}) \in \boldsymbol{P}_D(\hat{c})$, and the last inequality follows from (2.51) by taking $v' = \underline{\boldsymbol{p}}_D(\hat{c})$.

Moreover, by (2.54), for any $z \in [\underline{c}, \hat{c})$, and any $v \in V^+$, since $\overline{\boldsymbol{p}}_D(z) \leq \underline{\boldsymbol{p}}_D(\hat{c})$, it must be that for all $v' \leq \overline{\boldsymbol{p}}_D(z)$,

$$\begin{aligned}
(v' - z) \sum_{\hat{v} \geq v'} \hat{m}^v(\hat{v}) =& \alpha^*(v)(v' - z) D(v') \\
\leq& \alpha^*(v)(\underline{\boldsymbol{p}}_D(z) - z) D(\underline{\boldsymbol{p}}_D(z)) \\
=& (\underline{\boldsymbol{p}}_D(z) - z) \sum_{\hat{v} \geq \underline{\boldsymbol{p}}_D(z)} \hat{m}^v(\hat{v}).
\end{aligned} \tag{2.56}$$

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

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

and

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

where

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

Again,

$$\sum_{v \geq \overline{\boldsymbol{p}}_D(\underline{c})} \alpha^*(v) = 1 \tag{2.57}$$

and hence

$$\sum_{v \in V^+} \hat{m}^v(v') = m^D(v'),\, \forall v' \in V. \tag{2.58}$$

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

$$(v' - \underline{c}) \sum_{\hat{v} \geq v'} \hat{m}^v(v') = \alpha^*(v)(v' - \underline{c}) D(v') \leq \alpha^*(v)(\overline{\boldsymbol{p}}_D(\underline{c}) - \underline{c}) D(\overline{\boldsymbol{p}}_D(\underline{c})) \leq (v - \underline{c}) \hat{m}^v(v). \tag{2.59}$$

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

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

and let $D_v(p) := m^v([p, \overline{v}])$. By (2.52), (2.53) and (2.57), in each case, $D_v \in \mathcal{D}$ for all

$v \in V^+$. Now define $\sigma(c) \in \Delta^f(\mathcal{D})$ by

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

By (2.49) and (2.58), in each case, $\sigma(c) \in \mathcal{S}_D$. Furthermore, since m^v is proportional to $\hat{m}^v$ for all $v \in V^+$, (2.51), (2.55) and (2.59) ensure that in each case, $\sigma(c)$ is a $\psi(c)$-quasi-perfect segmentation for $\hat{c}$. Meanwhile, since $\hat{c} \leq c$, $\sigma(c)$ is also a $\psi(c)$-quasi-perfect segmentation for c. Finally, since m^v is proportional to $\hat{m}^v$, (2.56) implies that for any $z \in [\underline{c}, \hat{c})$,

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

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

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

In particular, $\sigma(c)$ is also a $\psi(c)$-quasi-perfect segmentation for c. Together with the fact that ψ is nondecreasing and that $\psi \leq \overline{\boldsymbol{p}}_D$, it then follows that for any $z \in [\underline{c}, c]$ and for any $D \in \operatorname{supp}(\sigma(c))$, , $\psi(z) \leq \overline{\boldsymbol{p}}_D(z)$. Since $c \in C$ is arbitrary, this ensures that there exists a ψ-quasi-perfect scheme $\sigma \in \mathcal{S}_D^C$ that satisfies (2.5).

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

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

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

for all $D \in \text{supp}(\sigma_n(c))$ and for all $z \in [\underline{c}, c]$. Furthermore, according to Helly's selection theorem, by possibly taking a subsequence,[3] $\{\sigma_n\} \to \sigma$ for some $\sigma : C \to \Delta(\mathcal{D})$. By Lemma 2.12, $\sigma \in \mathcal{S}^C$ and is a ψ-quasi-perfect scheme.

It then remains to show that σ satisfies (2.5). To this end, fix any $c \in C$ and consider any $D \in \text{supp}(\sigma(c))$, by definition, for any $\delta > 0$, $\sigma(N_\delta(D)|c) > 0$.[4] Furthermore, since $\sigma(c)$ has at most countably many atoms, there exists a sequence $\{\delta_k\} \subset (0, 1]$ such that $\{\delta_k\} \to 0$, $\sigma(N_{\delta_k}(D)|c) > 0$ and $\sigma(\partial N_{\delta_k}(D)|c) = 0$ for all $k \in \mathbb{N}$. As a result, since $\{\sigma_n(c)\} \to \sigma(c)$ under the weak-* topology, $\lim_{n\to\infty} \sigma_n(N_{\delta_k}(D)|c) = \sigma(N_{\delta_k}(D)|c) > 0$ for all $k \in \mathbb{N}$. Thus, for each $k \in \mathbb{N}$, there exists $n_k \in \mathbb{N}$ such that $\sigma_{n_k}(N_{\delta_k}(D)|c) > 0$. Moreover, since $\sigma_n(c)$ has finite support as D_n is a step function and $\sigma_n(c) \in \mathcal{S}_{D_n}$, there must be some $D_{n_k} \in N_{\delta_k}(D)$ such that $D_{n_k} \in \text{supp}(\sigma_{n_k}(c))$. Notice that for the subsequence $\{n_k\}$, $\{D_{n_k}\} \to D$ and $D_{n_k} \in \text{supp}(\sigma_{n_k}(c))$ for all $k \in \mathbb{N}$. As a result, together with Lemma 2.9, since σ_{n_k} satisfies (2.5) for all $k \in \mathbb{N}$, for Lebesgue almost all $z \in [\underline{c}, c]$,

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

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

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

4. $N_\delta(D)$ is the δ-ball around D under the Lévy-Prokhorov metric on $\mathcal{D}$.

Proof of Theorem 2.2

Proof of Theorem 2.2. I first show that the data broker's optimal revenue under regime $\mathcal{I}$ is R^*. To see this, since $c \leq \overline{\varphi}_G(c) \leq \overline{\boldsymbol{p}}_0(c)$ for all $c \in C$ and $\overline{\varphi}_G \in \mathbb{R}_+^C$ is nondecreasing, by Lemma 2.3, there exists a $\overline{\varphi}_G$-quasi-perfect scheme $\sigma^* \in \mathcal{S}^C$ that satisfies (2.5). Together with Lemma 2.2, there exists a transfer τ^* such that (σ^*, τ^*) is incentive compatible. Moreover, since $\sigma \in \mathcal{S}^C$ is a $\overline{\varphi}_G$-quasi-perfect scheme, by assertion 3 and assertion 4 of Lemma 2.10, for any $c \in C$,

$$\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma^*(\mathrm{d}D|c) = \int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v). \tag{2.60}$$

Also, by possibly adding a constant to the transfer τ^*, the indirect utility of the producer with cost $\overline{c}$, $U(\overline{c})$, equals to $\pi_0(\overline{c})$ under the mechanism (σ^*, τ^*). Thus, for any $c \in C$,

$$\begin{aligned}\int_{\mathcal{D}} \pi_D(c)\sigma^*(\mathrm{d}D|c) - \tau(c) =& U(\overline{c}) + \int_c^{\overline{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*(\mathrm{d}D|z) \right) \mathrm{d}z \\ =& \pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\varphi}_G(z)) \, \mathrm{d}z \\ \geq& \pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z)) \, \mathrm{d}z \\ =& \pi_0(c),\end{aligned}$$

where the first equality follows from Lemma 2.1, the second equality follows from assertion 3 of Lemma 2.10, the inequality follows from $\overline{\varphi}_G \leq \overline{\boldsymbol{p}}_0$ and the last equality follows from

(2.2). As a result, (σ^*, τ^*) is individually rational and, together with (2.60) and Lemma 2.1,

$$\begin{aligned}\mathbb{E}[\tau^*(c)] =& \int_C \left(\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c) - \phi_G(c)) D(\overline{\boldsymbol{p}}_D(c)) \sigma^*(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \\ =& \int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \\ =& R^*,\end{aligned}$$

as desired.

Since the data broker's optimal revenue is R^* and since (2.60) holds for any $\overline{\varphi}_G$-quasi-perfect scheme σ, by Lemma 2.1, any incentive feasible $\overline{\varphi}_G$-quasi-perfect mechanism must give revenue R^* and hence is optimal.

Conversely, to see why any optimal mechanism must be a $\overline{\varphi}_G$-quasi-perfect mechanism, consider any optimal mechanism (σ, τ). As it is optimal and incentive compatible, by Lemma 2.1,

$$R^* = \mathbb{E}[\tau(c)] = \int_C \left(\int_{\mathcal{D}} (\boldsymbol{p}_D(c) - \phi_G(c)) D(\boldsymbol{p}_D(c)) \sigma(\mathrm{d}D|c) \right) G(\mathrm{d}c) - \pi_0(\overline{c}), \tag{2.61}$$

for any $\boldsymbol{p} \in \boldsymbol{P}$. Also, since (σ, τ) is incentive compatible, for any $\boldsymbol{p} \in \boldsymbol{P}$, the function $\boldsymbol{D}_{\boldsymbol{p}}^{\sigma} \in [0,1]^C$, defined as

$$\boldsymbol{D}_{\boldsymbol{p}}^{\sigma}(c) := \int_{\mathcal{D}} D(\boldsymbol{p}_D(c)) \sigma(\mathrm{d}D|c), \, \forall c \in C$$

is nonincreasing.[5] Thus, by (2.37),

$$\int_C \overline{\phi}_G(c)\left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c) \geq \int_C \overline{\varphi}_G(c)\left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{d}c). \tag{2.62}$$

Moreover, since (σ, τ) is individually rational, by Lemma 2.1, it must be that

$$\int_c^{\overline{c}} \left(\int_{\mathcal{D}} D(\boldsymbol{p}_D(z))\sigma(\mathrm{d}D|z)\right)\mathrm{d}z \geq \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z))\,\mathrm{d}z,\ \forall c \in C. \tag{2.63}$$

Furthermore, since σ^* is a $\overline{\varphi}_G$-quasi-perfect scheme, Lemma 2.10 implies that, for all $c \in C$,

$$\boldsymbol{D}_{\overline{\boldsymbol{p}}}^{\sigma^*}(c) = \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c))\sigma^*(\mathrm{d}D|c) = D_0(\overline{\varphi}_G(c)).$$

Together with (2.38) and (2.39), we have

$$\int_C \overline{\varphi}_G(c) D_0(\overline{\varphi}_G(c))G(\mathrm{d}c) = \int_C \overline{\phi}_G(c) D_0(\overline{\varphi}_G(c))G(\mathrm{d}c). \tag{2.64}$$

Now suppose that (σ, τ) is not a $\overline{\varphi}_G$-quasi-perfect mechanism or it does not induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for a positive G-measure of c, then there exists $\boldsymbol{p} \in \boldsymbol{P}$, a positive G-measure of c and a positive $\sigma(c)$-measure of $D \in \mathcal{D}$ such that either $\boldsymbol{p}_D(c) < \overline{\boldsymbol{p}}_D(c)$, or $D(c) > 0$ and either $\#\{v \in \mathrm{supp}(D) : v \geq \overline{\varphi}_G(c)\} \neq 1$ or $\max(\mathrm{supp}(D)) \notin \boldsymbol{P}_D(c)$, which imply that there is a positive G-measure of c and a positive $\sigma(c)$-measure of D such

5. To see this, notice that U is convex since it is a pointwise supremum of convex functions, which is because $\pi_D(c)$ is convex for all D. Lemma 2.1 implies that the derivative of U is $-\boldsymbol{D}_{\boldsymbol{p}}^{\sigma}$ and thus $\boldsymbol{D}_{\boldsymbol{p}}^{\sigma}$ must be nonincreasing.

that

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

with at least one inequality being strict. Therefore,

$$\int_C\left(\int_{\mathcal{D}}(\boldsymbol{p}_D(c)-\overline{\varphi}_G(c))D(\boldsymbol{p}_D(c))\sigma(\mathrm{d}D|c)\right)G(\mathrm{dc})<\int_C\left(\int_V(v-\overline{\varphi}_G(c))^+D_0(\mathrm{d}v)\right)G(\mathrm{dc}).\tag{2.65}$$

Moreover, by (2.61), since

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

together with, (2.65), it must be that

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

where the first inequality follows from (2.62) and the equality follows from (2.64). Furthermore, since $\overline{\phi}_G(c) = \phi_G(c)$ for all $c \in [\underline{c}, c^*]$ and $\overline{\phi}_G(c) = \overline{\varphi}_G(c) = \overline{\boldsymbol{p}}_0(c)$ for all $c \in (c^*, \overline{c}]$, it then follows that

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

Using integration by parts, this is equivalent to

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

where M^* is defined in (2.34). However, by (2.63) and by the fact that M^* is a CDF of a Borel measure, which is due to Assumption 2.1,

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

a contradiction. Therefore, σ must be a $\overline{\varphi}_G$-quasi-perfect scheme and must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G-almost all $c \in C$. Together with Lemma 2.1, and the fact that $U(\overline{c}) = \pi_0(\overline{c})$ under any optimal mechanism, (σ, τ) must be a $\overline{\varphi}_G$-quasi-perfect mechanism. This completes the proof. ■

Proof of Theorem 2.3

Proof of Theorem 2.3. Notice that from the proof of Lemma 2.3, for any step function $D \in \mathcal{D}$ that is regular, by replacing ψ with $\overline{\varphi}_G$, the resulting scheme $\sigma \in \mathcal{S}_D^C$ must take form of

$$\sigma\left(D_v^{\overline{\varphi}_G(c)}\middle|c\right) = \frac{m^D(v)}{D(\overline{\varphi}_G(c))} \tag{2.66}$$

for all $c \in C$ and for all $v \in [\overline{\varphi}_G(c), \overline{v}]$, where $D_v^{\overline{\varphi}_G(c)}$ is defined by (2.6) with D_0 being replaced by D.

Now, for any regular $D_0 \in \mathcal{D}$, take a sequence of regular step functions $\{D_n\} \subseteq \mathcal{D}$ such that $\{D_n\} \to D_0$ and take the associated segmentation scheme σ_n defined by (2.66) for each $n \in \mathbb{N}$. By the proof of Lemma 2.3, $\{\sigma_n\} \to \sigma$ for some $\overline{\varphi}_G$-quasi-perfect scheme $\sigma \in \mathcal{S}^C$ that satisfies (2.5). Moreover, by the same argument as in the proof of Lemma 2.3, for all $c \in C$ and for all $D \in \text{supp}(\sigma(c))$, there exists a subsequence $\{D_{n_k}\}$ such that $D_{n_k} \in \text{supp}(\sigma_{n_k}(c))$ and $\{D_{n_k}\} \to D$. This then implies that $D(p) = D_0(p)$ for every $p \in [\underline{v}, \overline{\varphi}_G(c))$ at which D_0 is continuous. Finally, since $\sigma(c)$ is a $\overline{\varphi}_G(c)$-quasi-perfect segmentation for c, it must be that for any $p \in [\overline{\varphi}_G(c), \overline{v}]$, $D(p) \in \{D_0(\overline{\varphi}_G(c)), 0\}$ and for any $p \in [\underline{v}, \overline{\varphi}_G(c))$, $D(p) = D_0(p)$, which means that $D = D_v^{\overline{\varphi}_G(c)}$ for some $v \in [\overline{\varphi}_C(c), \overline{v}]$ and

$$\sigma\left(\left\{D_v^{\overline{\varphi}_G(c)} : v \geq p\right\}\middle|c\right) = \frac{D_0(p)}{D_0(\overline{\varphi}_G(c))} \tag{2.67}$$

for any $p \in [\overline{\varphi}_G(c), \overline{v}]$, where $D_v^{\overline{\varphi}_G(c)}$ is defined by (2.6)

Finally, by the proof of Lemma 2.3, σ is a $\overline{\varphi}_G$-quasi-perfect scheme satisfying (2.5) and hence, by Lemma 2.2, there exists a transfer scheme τ such that (σ, τ) is incentive feasible. Thus, by Theorem 2.2, (σ, τ) is optimal, and by (2.67), (σ, τ) is the canonical $\overline{\varphi}_G$-quasi perfect mechanism. ■

Proof of Theorem 2.4

Proof of Theorem 2.4. By Theorem 2.1 and Theorem 2.2, both the data broker's optimal revenue under regime $\mathcal{P}$ is R^*. Furthermore, for any optimal mechanism (σ, τ) under regime $\mathcal{I}$ and any optimal mechanism $(\hat{\sigma}, \hat{\tau}, \hat{\boldsymbol{\gamma}})$ under regime $\mathcal{P}$, both of them must induce $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for G-almost all $c \in C$. In particular, for G-almost all $c \in C$, all the consumers with $v \geq \overline{\varphi}_G(c)$ buys the product by paying their values and all the consumers with $v < \overline{\varphi}_G(c)$ do not buy the product. Thus, the consumer surplus and the allocation of the product induced by (σ, τ) and $(\hat{\sigma}, \hat{\tau}, \hat{\boldsymbol{\gamma}})$ are the same.

In addition, for the optimal mechanism (σ, τ) under regime $\mathcal{I}$, Theorem 2.2 implies that σ must be a $\overline{\varphi}_G$-quasi-perfect scheme and hence by Lemma 2.10, for G-almost all $c \in C$,

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

and

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

Therefore, for Lebesgue almost all $c \in C$, by Lemma 2.1,

$$\begin{aligned}\int_{\mathcal{D}} \pi_D(c)\sigma(\mathrm{d}D|c) - \tau(c) =& \pi_0(\overline{c}) + \int_c^{\overline{c}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma(\mathrm{d}D|z) \right) \mathrm{d}z \\ =& \pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\varphi}_G(z)) \, \mathrm{d}z. \end{aligned} \tag{2.68}$$

On the other hand, for the optimal mechanism $(\hat{\sigma}, \hat{\tau}, \hat{\boldsymbol{\gamma}})$ under regime $\mathcal{P}$, since it induces $\overline{\varphi}_G(c)$-quasi-perfect price discrimination for almost all $c \in C$, it must be that,

$$\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p)\hat{\boldsymbol{\gamma}}(\mathrm{d}p|D, c) \right) \hat{\sigma}(\mathrm{d}D|c) = D_0(\overline{\varphi}_G(c)),$$

for G-almost all $c \in C$. Together with (2.44), we have

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

Therefore, by Lemma 2.13, for any $c \in C$,

$$\begin{aligned}
&\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} (p - c) D(p) \hat{\boldsymbol{\gamma}}(\mathrm{d}p|D, c) \right) \hat{\sigma}(\mathrm{d}D|c) - \hat{\tau}(c) \\
=& \pi_0(\overline{c}) + \int_c^{\overline{c}} \left(\int_{\mathcal{D}} \left(\int_{\mathbb{R}_+} D(p) \hat{\boldsymbol{\gamma}}(\mathrm{d}p|D, z) \right) \hat{\sigma}(\mathrm{d}D|z) \right) \mathrm{d}z \\
=& \pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\varphi}_G(z)) \, \mathrm{d}z.
\end{aligned}$$

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

Proof of Proposition 2.1

Proof of Proposition 2.1. Let (σ, τ) be any optimal mechanism. By Theorem 2.2, (σ, τ) must be a $\overline{\varphi}_G$-quasi-perfect mechanism and induces $\overline{\varphi}_G$-quasi-perfect price discrimination. Therefore, for G-almost all $c \in C$ and for $\sigma(c)$-almost all $D \in \mathcal{D}$, $D(p) = 0$ for any $p > \overline{\boldsymbol{p}}_D(c) = \max(\mathrm{supp}(D))$ and consumer surplus under (σ, τ) is

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

Using integration by parts, it then follows that

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

for G-almost all $c \in C$. Together,

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

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

$$=0.$$

This completes the proof. ∎

Proof of Proposition 2.2

Proof of Proposition 1. By Theorem 2.2, the data broker's optimal revenue is

$$R^* = \int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \pi_0(\overline{c}).$$

On the other hand, since $\boldsymbol{P}_0(c)$ is a singleton for (Lebesgue)-almost all $c \in C$ and since G is absolutely continuous, consumer surplus under uniform pricing is

$$\int_C \left(\int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \overline{\boldsymbol{p}}_0(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c).$$

Furthermore, for any $c \in C$,

$$\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) = \int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) + \int_{[\overline{\varphi}_G(c), \overline{\boldsymbol{p}}_0(c))} (v - \phi_G(c)) D_0(\mathrm{d}v). \tag{2.69}$$

Notice first that for any $c \in C$,

$$\begin{aligned}
&\int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) - \int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \overline{\boldsymbol{p}}_0(c)) D_0(\mathrm{d}v) - \pi_0(\overline{c}) \\
=&(\overline{\boldsymbol{p}}_0(c) - \phi_G(c)) D_0(\overline{\boldsymbol{p}}_0(c)) - \pi_0(\overline{c}) \\
=&\pi_0(c) - \frac{G(c)}{g(c)} D_0(\overline{\boldsymbol{p}}_0(c)) - \pi_0(\overline{c}) \\
=&\int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z)) \, \mathrm{d}z - \frac{G(c)}{g(c)} D_0(\overline{\boldsymbol{p}}_0(c)),
\end{aligned}$$

and thus

$$\begin{aligned}
&\int_C \left(\int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) - \int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \overline{\boldsymbol{p}}_0(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \\
=&\int_C \left(\int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z)) \, \mathrm{d}z - \frac{G(c)}{g(c)} D_0(\overline{\boldsymbol{p}}_0(c)) \right) G(\mathrm{d}c) \qquad (2.70) \\
=&\int_C G(c)(D_0(\overline{\boldsymbol{p}}_0(c)) - D_0(\overline{\boldsymbol{p}}_0(c))) \, \mathrm{d}c \\
=&0.
\end{aligned}$$

On the other hand, for any $c \in C$,

$$\begin{aligned}
&\int_{[\overline{\varphi}_G(c), \overline{\boldsymbol{p}}_0(c))} (v - \phi_G(c)) D_0(\mathrm{d}v) \\
=&\int_{[\overline{\varphi}_G(c), \overline{\boldsymbol{p}}_0(c))} (v - \varphi_G(c)) D_0(\mathrm{d}v) + (\varphi_G(c) - \phi_G(c))(D_0(\overline{\varphi}_G(c)) - D_0(\overline{\boldsymbol{p}}_0(c))),
\end{aligned}$$

and thus,

$$
\begin{aligned}
&\int_C \left(\int_{[\overline{\varphi}_G(c), \overline{\boldsymbol{p}}_0(c))} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) \\
=& \int_C \left(\int_{[\overline{\varphi}_G(c), \overline{\boldsymbol{p}}_0(c))} (v - \varphi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) \\
&+ \int_C (\varphi_G(c) - \phi_G(c))(D_0(\overline{\varphi}_G(c)) - D_0(\overline{\boldsymbol{p}}_0(c)))G(\mathrm{d}c) \\
\geq& \int_C (\varphi_G(c) - \phi_G(c))(D_0(\overline{\varphi}_G(c)) - D_0(\overline{\boldsymbol{p}}_0(c)))G(\mathrm{d}c) \qquad (2.71) \\
=& \int_C (\varphi_G(c) - \phi_G(c)) D_0(\overline{\varphi}_G(c)) G(\mathrm{d}c) - \int_C (\varphi_G(c) - \phi_G(c)) D_0(\overline{\boldsymbol{p}}_0(c))) G(\mathrm{d}c) \\
\geq& 0,
\end{aligned}
$$

where the first inequality follows from $\overline{\varphi}_G = \min\{\varphi_G, \overline{\boldsymbol{p}}_0\}$ and the second inequality follows from the fact that $\overline{\boldsymbol{p}}_0$ is nondecreasing, (2.37) and (2.38). Together, combining (2.69), (2.70) and (2.71),

$$
\begin{aligned}
&\int_C \left(\int_{\{v \geq \overline{\varphi}_G(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \pi_0(\overline{c}) - \int_C \left(\int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \overline{\boldsymbol{p}}_0(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) \\
=& \int_C \left(\int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \phi_G(c)) D_0(\mathrm{d}v) - \int_{\{v \geq \overline{\boldsymbol{p}}_0(c)\}} (v - \overline{\boldsymbol{p}}_0(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) - \pi_0(\overline{c}) \\
&+ \int_C \left(\int_{[\overline{\varphi}_G(c), \overline{\boldsymbol{p}}_0(c))} (v - \phi_G(c)) D_0(\mathrm{d}v) \right) G(\mathrm{d}c) \\
\geq& 0,
\end{aligned}
$$

as desired. ■

Proof of Theorem 2.5

Proof of Theorem 2.5. For each $\theta \in \Theta$, write $\text{supp}(D_\theta)$ as $[l(\theta), u(\theta)]$. Also, for any $p \in V$, let θ_p be the unique θ such that $p \in (l(\theta), u(\theta)]$. Notice that since $\{(l(\theta), u(\theta))\}_{\theta \in \Theta}$ is disjoint, for any $\beta \in \Delta(\Theta)$, any $\theta \in \Theta$, and any $p \in \text{supp}(D_\beta)$,

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

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

As a result, the construction in the proof of Lemma 2.3 is still valid, with the demands being replaced by D_β. Specifically, for any $\beta \in \Delta(\Theta)$ and any $c \in C$, there exists $\{\beta_i\}_{i=1}^n \subseteq \Delta^f(V)$ such that:

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

2. For each $i \in \{1, \dots, n\}$, the set

$$\{\theta \in \text{supp}(\beta_i) | u(\theta) \geq \overline{\boldsymbol{p}}_{D_{\beta_i}}(c)\}$$

is nonempty and is a singleton.

3. For each $i \in \{1, \dots, n\}$,

$$\boldsymbol{P}_{D_{\beta_i}}(c) \bigcap \text{supp}(D_{\bar{\theta}_{\beta_i}}) \neq \emptyset,$$

where $\bar{\theta}_{\beta_i} := \max\{u(\theta) : \theta \in \text{supp}(\beta_i)\}$.

4. For each $i \in \{1, \ldots, n\}$ and any $z \in [\underline{c}, c]$,

$$\overline{\boldsymbol{p}}_{D_{\beta_i}}(z) \geq \overline{\boldsymbol{p}}_{D_\beta}(z).$$

This further implies that, by Lemma 2.10, and by the same argument as in the proof of Lemma 2.2, for any $\beta \in \Delta(\Theta)$, there exists $\sigma^\beta \in \Delta^f(\Delta(\Theta))^C$ such that

5. For any $c \in C$,

$$\sum_{\beta' \in \text{supp}(\sigma^\beta(c))} (\overline{\boldsymbol{p}}_{D_{\beta'}}(c) - \overline{\boldsymbol{p}}_{D_\beta}(c)) D_{\beta'}(\overline{\boldsymbol{p}}_{D'_\beta}(c)) \sigma^\beta(\beta'|c)$$
$$= \sum_{\{\theta: u(\theta) \geq \theta(\overline{\boldsymbol{p}}_{D_\beta}(c))\}} (\overline{\boldsymbol{p}}_{D_\theta}(c) - \overline{\boldsymbol{p}}_{D_\beta}(c)) D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) \beta(\theta).$$

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

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

8. For any $\beta' \in \text{supp}(\sigma^\beta(c'))$,

$$\overline{\boldsymbol{p}}_{D_{\beta'}}(c) \geq \overline{\boldsymbol{p}}_{D_\beta}(c),$$

   for any $c, c' \in C$ such that $c < c'$ and

$$\sum_{\beta' \in \text{supp}(\sigma^\beta(c))} D_{\beta'}(\overline{\boldsymbol{p}}_{D_{\beta'}}(c)) \sigma^\beta(\beta'|c) \geq D(\overline{\boldsymbol{p}}_{D_\beta}(c)),$$

   for any $c, c' \in C$ such that $c > c'$.

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

$$\{\theta \in \text{supp}(\beta) : u(\theta) > u(\theta_{\boldsymbol{p}_{D_\beta}(c)})\} \neq \emptyset. \tag{2.72}$$

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

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

where the second equality follows from 5 and 6 and the inequality is strict due to (2.72).

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

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

As (2.72) holds with positive G-measure of $c \in C$, the induced segmentation scheme $\hat{\sigma} \in \Delta^f(\Delta(\Theta))^C$ strictly improves the data broker's revenue. Finally, as argued in the proof of Theorem S1 in the Supplemental Material, assertions 6 and 8 above and Lemma 2.1 ensure that there exists a transfer $\hat{\tau}$ such that $(\hat{\sigma}, \hat{\tau})$ is incentive compatible and individually rational and strictly improves the data broker's revenue.

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

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

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

Proof of Theorem 2.6

To prove Theorem 2.6, I first introduce two useful lemmas.

Lemma 2.16. *For any $c \in C$, any $\nu \geq c$ and any segmentation $s \in \Delta^f(\Delta(\Theta))$,*

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

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

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

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

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

for any $\theta' \in \mathrm{supp}(\beta)$ and for any $\theta \in \mathrm{supp}(\beta)$ with $u(\theta) \geq \overline{\boldsymbol{p}}_{D_\beta}(c)$. Notice that by construction, $\beta \in \mathrm{co}(\{\hat{\beta}^\theta\}_{\theta \geq \overline{\boldsymbol{p}}_{D_\beta}(c)})$ and hence there exists $K^\beta \in \Delta^f(\Delta(\Theta))$ such that $\beta = \sum_{\hat{\beta}} K^\beta(\hat{\beta})$. Therefore, by splitting every β according to K^β, and by the same arguments as in the proof of Lemma 2.3, the resulting segmentation $\hat{s} \in \Delta^f(\Theta)$ must be such that for any $\hat{\beta} \in \mathrm{supp}(\hat{s})$, $\overline{\boldsymbol{p}}_{D_{\hat{\beta}}}(c)$ is in the interval described by $\max(\mathrm{supp}(\hat{\beta}))$. Moreover, since $\{(l(\theta), u(\theta))\}_{\theta \in \Theta}$ is disjoint, it follows that $\overline{\boldsymbol{p}}_{D_{\bar{\theta}_{\hat{\beta}}}}(c) = \overline{\boldsymbol{p}}_{D_{\hat{\beta}}}(c)$. Furthermore, since for any $\beta \in \text{supp}(s)$,

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

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

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

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

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

as desired. ∎

Lemma 2.17. *Suppose that D_0 is regular. For any $c \in C$ and for any $\nu \in [c, \overline{\boldsymbol{p}}_0(c)]$,*

$$D_0(\overline{\boldsymbol{p}}_0(c)) \leq \sum_{\{\theta: u(\theta) \geq \nu\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c))\beta_0(\theta) \tag{2.73}$$

and

$$D_0(\overline{\boldsymbol{p}}_0(c)) \geq \sum_{\{\theta: l(\theta) \geq \overline{\boldsymbol{p}}_0(c)\}} D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c))\beta_0(\theta). \tag{2.74}$$

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

$p \in (l(\theta), u(\theta)]$,

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

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

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

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

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

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

and

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

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

Consequently,

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

which proves (2.74). On the other hand, for any $\nu \in [c, \overline{\boldsymbol{p}}_0(c)]$

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

which proves (2.73) ■

Proof of Theorem 2.6. To prove Theorem 2.6, first notice that Lemma 2.1 still applies and hence the data broker's maximization problem can be written as

$$\begin{aligned}
\max_{\sigma} \; & \int_C \left(\int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \phi_G(c)) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c)) \sigma(\mathrm{d}\beta|c) \right) G(\mathrm{d}c) \\
\text{s.t.} \; & \int_c^{c'} \left(D_\beta(\overline{\boldsymbol{p}}_\beta(z))(\sigma(\mathrm{d}\beta|z) - \sigma(\mathrm{d}\beta|c')) \right) \mathrm{d}z \geq 0, \, \forall c, c' \in C \\
& \pi_0(\overline{c}) + \int_c^{\overline{c}} \left(\int_{\Delta(\Theta)} D_\beta(\overline{\boldsymbol{p}}_\beta(z)) \sigma(\mathrm{d}\beta|z) \right) \mathrm{d}z \geq \pi_0(\overline{c}) + \int_c^{\overline{c}} D_0(\overline{\boldsymbol{p}}_0(z)) \, \mathrm{d}z, \, \forall c \in C,
\end{aligned} \tag{2.75}$$

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

Consider first a relaxed problem of (2.75) where the first constraint is relaxed to $\boldsymbol{D}_\sigma \in [0,1]^C$ being nonincreasing, where

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

for all $c \in C$. By the same duality argument as in the proof of Lemma 2.15, it suffices to find a feasible σ^* and a Borel measure μ^* on C such that

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

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

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

To this end, let M^* be defined as

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

Since $c \mapsto g(c)(\phi_G(c) - \widehat{\boldsymbol{p}}_0(c))^+$ in nondecreasing, M^* is nondecreasing and right-continuous and hence induced a Borel measure μ^* with $\mathrm{supp}(\mu^*) = [c^*, \overline{c}]$ for some $c^* \leq \overline{c}$. Then, by the same arguments as in the proof of Lemma 2.15 (in particular, the definition of $\widehat{\varphi}_G$, (2.37) and (2.38)),

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

is equivalent to

$$\max_{\sigma \in \Sigma} \int_C \left(\int_{\Delta(\Theta)} (\overline{\boldsymbol{p}}_{D_\beta}(c) - \widehat{\varphi}_G(c)) D_\beta(\overline{\boldsymbol{p}}_{D_\beta}(c)) \sigma(\mathrm{d}\beta | c) \right) G(\mathrm{d}c). \tag{2.76}$$

To solve (2.76), notice that for any $c \in [\underline{c}, c^*)$,

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

which is due to $\widehat{\varphi}_G(c) = \varphi_G(c) \leq \widehat{\boldsymbol{p}}_0(c) \leq \overline{\boldsymbol{p}}_0(c)$ and (2.74). On the other hand, for any $c \in (c^*, \overline{c}]$, there exists a unique $\lambda(c)$ such that

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

which is due to the fact that $\widehat{\varphi}_G(c) = \widehat{\boldsymbol{p}}_0(c)$ for all $c \in (c^*, \overline{c}]$ and (2.73). Furthermore, Since

D_0 is regular, for any $\theta \in \Theta$ such that $u(\theta) \geq \widehat{\varphi}_G(c)$ and for any $p \leq l(\theta_{\widehat{\varphi}_G(c)})$,

$$
\begin{aligned}
(p-c)D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}(p) =& \sum_{\{\theta' : u(\theta') \geq u(\theta_p)\}} (p-c)D_{\theta'}(p)\beta^{\theta}_{\widehat{\varphi}_G(c)}(\theta') \\
=&(p-c)D_0(p) \\
\leq&(l(\theta_{\widehat{\varphi}_G(c)}) - c)D_0(l(\theta_{\widehat{\varphi}_G(c)})) \\
\leq&(l(\theta) - c)D_0(l(\theta_{\widehat{\varphi}_G(c)})) \\
=&(l(\theta) - c) \sum_{\{\theta' : u(\theta') \geq \widehat{\varphi}_G(c)\}} \beta^{\theta}_{\widehat{\varphi}_G(c)}(\theta') \\
=&(l(\theta) - c)D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}(l(\theta)) \\
=&(\overline{\boldsymbol{p}}_{D_\theta}(c) - c)D_{\beta^{\theta}_{\widehat{\varphi}_G(c)}}(\overline{\boldsymbol{p}}_{D_\theta}(c)),
\end{aligned}
\tag{2.77}
$$

where $\beta^{\theta}_{\widehat{\varphi}_G(c)}$ is defined in (2.11). In addition, by the same construction as in the proof of Lemma 2.3, for any $c \in (c^*, \overline{c}]$, there exists a segmentation $\tilde{\sigma}(c) \in \Delta^f(\Delta(\Theta))$ such that $\text{supp}(\tilde{\sigma}(c)) = \{\tilde{\beta}^{\theta}_{\widehat{\boldsymbol{p}}_0(c)} : l(\theta) \geq \widehat{\boldsymbol{p}}_0(c)\}$, with $\tilde{\beta}^{\theta}_{\widehat{\boldsymbol{p}}_0(c)}$ satisfying (2.12) and (2.13) and that

$$
(p-c)D_{\tilde{\beta}^{\theta}_{\widehat{\boldsymbol{p}}_0(c)}}(p) \leq (l(\theta) - c)D_\theta(l(\theta)) = (\overline{\boldsymbol{p}}_{D_\theta}(c) - c)D_\theta(\overline{\boldsymbol{p}}_{D_\theta}(c)) \tag{2.78}
$$

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

$$
\overline{\boldsymbol{p}}_{D_{\tilde{\beta}^{\theta}_{\widehat{\boldsymbol{p}}_0(c)}}}(z) \geq \overline{\boldsymbol{p}}_{D_0}(z) \geq \widehat{\boldsymbol{p}}_0(z) \tag{2.79}
$$

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

Now define σ^* as follows.

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

where

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

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

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

for all $c \in (c^*, \overline{c}]$. It then follows that, by (2.77) and (2.78),

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

which, together with Lemma 2.16, implies that σ^* is a solution of (2.76).

Furthermore, for any $c > c^*$, by the definition of $\sigma_2(c)$ and $\lambda(c)$, by (2.77) and (2.78), and by the fact that $\widehat{\varphi}_G(c) = \widehat{\boldsymbol{p}}_0(c)$,

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

Therefore,

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

Finally, by definition of $\widehat{\varphi}_G$ and by Lemma 2.17,

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

for all $c \in [\underline{c}, c^*)$. Together with monotonicity of $\widehat{\varphi}_G$, $\sigma^* \in \Sigma$ and is a solution of the relaxed problem of (2.75).

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

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

then it must be that

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

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

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

for all $z \in [\underline{c}, c]$. Together with (2.79), by the same argument as the proof of Lemma 2.2, σ^* is indeed implementable. This completes the proof. ■

Proof of Theorem 2.7

Before proving Theorem 2.7. I first introduced the counterpart of Lemma 2.1 when targeting is available. The proof of this result is entirely analogous to the proof of Lemma 2.1 and therefore omitted.

Lemma 2.18. *A mechanism (σ, τ, q) is incentive compatible if and only if there exists constants $\{\overline{\tau}_j\}_{j \in \mathcal{J}} \subset \mathbb{R}$ such that for all $j \in \mathcal{J}$ and for all $c_j, c_j' \in C_j$,* *1.*

$$
\begin{aligned}
&\mathbb{E}_{c_{-j}}[\tau_j(c)]\\
=&\mathbb{E}_{c_{-j}}\left[\sum_{i\in\mathcal{I}}\int_{\mathcal{D}}\pi_D(c_j)\sigma_{ij}(\mathrm{d}D|c)q_{ij}(c)\right]\\
&-\int_{c_j}^{\overline{c}_j}\mathbb{E}_{c_{-j}}\left[\sum_{i\in\mathcal{I}}\int_{\mathcal{D}}D(\overline{\boldsymbol{p}}_D^j(z))\sigma_{ij}(\mathrm{d}D|z,c_{-j})q_{ij}(z,c_{-j})\right]\mathrm{d}z-\overline{\tau}_j,
\end{aligned}
$$

2.

$$
\begin{aligned}
&\int_{c_j}^{c_j'}\left(\mathbb{E}_{c_{-j}}\left[\sum_{i\in\mathcal{I}}\int_{\mathcal{D}}D(\overline{\boldsymbol{p}}_D^j(z))\sigma_{ij}(\mathrm{d}D|z,c_{-j})q_{ij}(z,c_{-j})\right]\right)\mathrm{d}z\\
&\quad-\int_{c_j}^{c_j'}\left(\mathbb{E}_{c_{-j}}\left[\sum_{i\in\mathcal{I}}\int_{\mathcal{D}}D(\overline{\boldsymbol{p}}_D^j(z))\sigma_{ij}(\mathrm{d}D|c_j',c_{-j})\right]q_{ij}(c_j',c_{-j})\right)\mathrm{d}z\geq 0
\end{aligned}
$$

With the characterization given by Lemma 2.18, the data broker's expected revenue can again be written as

$$
\int_C\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\left(\int_{\mathcal{D}}(\overline{\boldsymbol{p}}_D^j(c_j)-\phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D^j(c_j))\sigma_{ij}(\mathrm{d}D|c)q_{ij}(c)\right)G(\mathrm{d}c)-\sum_{j\in\mathcal{J}}\pi_0(\overline{c})_j,
$$

where $\pi_0(\overline{c})_j:=\pi_{D_0^j}(\overline{c}_j)$. Using this observation,

Proof of Theorem 2.7. Existence of solutions is ensured by compactness of the feasible set and continuity of the objective function, which rely on Lemma 2.7, Tychonoff's theorem, and the Lebesgue dominate convergence theorem.

Now consider any mechanism (σ,τ,q). Suppose that the consumers retain positive surplus. That is

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\int_C\int_{\mathcal{D}}(v-\overline{\boldsymbol{p}}_D(c_j))D(\mathrm{d}v)\sigma_{ij}(\mathrm{d}D|c)q_{ij}(c)G(\mathrm{d}c)>0.$$

It then suffices to show that there exists an incentive feasible mechanism $(\hat{\sigma},\hat{\tau},\hat{q})$ that strictly improves the data broker's revenue.

Notice that under (σ,τ,q), the data broker's revenue is

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\int_C\int_{\mathcal{D}}(\overline{\boldsymbol{p}}_D(c_j)-\phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D(c_j))\sigma_{ij}(\mathrm{d}D|c)q_{ij}(c)G(\mathrm{d}c)-\sum_{j\in\mathcal{J}}U_j(\bar{c}_j), \tag{2.80}$$

where U_j is the indirect utility of producer j. On the other hand, notice that for any $i\in\mathcal{I}$, $j\in\mathcal{J}$, since the segmentation scheme $\sigma_{ij}\in\mathcal{S}^C_{D_0^{ij}}$ is measurable, the mapping $\overline{\sigma}_{ij}:C_j\to\mathcal{D}$, defined as

$$\overline{\sigma}_{ij}(c_j):=\mathbb{E}_{c_{-j}}[\sigma_{ij}(c_j,c_{-j})],\,\forall c_j\in C_j$$

is also measurable and thus is also in $\mathcal{S}^C_{D_0^{ij}}$. As a result, as shown in the proof of Theorem S1 in the Supplemental Material, for any $j\in\mathcal{J}$ and any $i\in\mathcal{I}$, there exists a measurable function $\tilde{\sigma}_{ij}:C_j\to\mathcal{D}$ such that

$$\int_{\mathcal{D}}D\tilde{\sigma}_{ij}(\mathrm{d}D|c_j)=D_0^{ij},\,\forall c_j\in C_j, \tag{2.81}$$

and that

$$
\begin{aligned}
&\int_{C_j}\int_{\mathcal{D}}(\overline{\boldsymbol{p}}_D(c_j)-\phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D(c_j))\tilde{\sigma}_{ij}(\mathrm{d}D|c_j)G_j(\mathrm{d}c) \\
\geq &\int_{C_j}\int_{\mathcal{D}}(\overline{\boldsymbol{p}}_D(c_j)-\phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D(c_j))\overline{\sigma}_{ij}(\mathrm{d}D|c_j)G_j(\mathrm{d}c_j) \\
&+\int_{C_j}\int_{\mathcal{D}}\int_{\{v\geq\overline{\boldsymbol{p}}_D(c_j)\}}(v-\overline{\boldsymbol{p}}_D(c_j))D(\mathrm{d}v)\overline{\sigma}_{ij}(\mathrm{d}D|c_j)G_j(\mathrm{d}c_j) \\
\geq &\int_{C_j}\int_{\mathcal{D}}(\overline{\boldsymbol{p}}_D(c_j)-\phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D(c_j))\overline{\sigma}_{ij}(\mathrm{d}D|c)G_j(\mathrm{d}c_j).
\end{aligned} \tag{2.82}
$$

By (2.80), there exists $i^*\in\mathcal{I}$ and $j^*\in\mathcal{J}$ such that

$$
\int_{C_j}\int_{\mathcal{D}}\int_{\{v\geq\overline{\boldsymbol{p}}_D(c_j)\}}(v-\overline{\boldsymbol{p}}_D(c_j))D(\mathrm{d}v)\overline{\sigma}_{ij}(\mathrm{d}D|c_j)\overline{q}_{ij}(c_j)G_j(\mathrm{d}c_j)>0,
$$

where $\overline{q}_{ij}(c_j):=\mathbb{E}_{c_{-j}}[q_{ij}(c_j,c_{-j})]$ for all $i\in\mathcal{I}$, $j\in\mathcal{J}$, $c_j\in C_j$. As such, since $q_{i^*j^*}\in[0,1]$, we must have

$$
\int_{C_j^*}\int_{\mathcal{D}}\int_{\{v\geq\overline{\boldsymbol{p}}_D(c_j^*)\}}(v-\overline{\boldsymbol{p}}_D(c_j^*))D(\mathrm{d}v)\overline{\sigma}_{i^*j^*}(\mathrm{d}D|c_{j^*})\overline{q}_{i^*j^*}(c_j^*)G(\mathrm{d}c_j^*)>0
$$

and hence the last inequality in (2.82) must be strict inequality for some $i\in\mathcal{I}$, $j\in\mathcal{J}$. Therefore, if $\hat{\sigma}\in\mathcal{D}^C$ is defined as

$$
\hat{\sigma}_{ij}(c):=\tilde{\sigma}_{ij}(c_j),\ \forall i\in\mathcal{I},j\in\mathcal{J},c\in C,
$$

then

$$
\begin{aligned}
&\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\int_C\int_{\mathcal{D}}(\overline{\boldsymbol{p}}_D(c_j)-\phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D(c_j))\hat{\sigma}_{ij}(\mathrm{d}D|c)q_{ij}(c)G(\mathrm{d}c) \\
>&\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\int_C\int_{\mathcal{D}}(\overline{\boldsymbol{p}}_D(c_j)-\phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D(c_j))\sigma_{ij}(\mathrm{d}D|c)q_{ij}(c)G(\mathrm{d}c).
\end{aligned} \tag{2.83}
$$

On the other hand, as shown in the proof of Theorem S1 in the Supplemental Material, such $\{\tilde{\sigma}_{ij}\}$ are such that for any $i \in \mathcal{I}$ and $j \in \mathcal{J}$, for Lebesgue-almost all $c_j \in C_j$,

$$\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c_j))\tilde{\sigma}_{ij}(\mathrm{d}D|c_j) = \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c_j))\overline{\sigma}_{ij}(\mathrm{d}D|c_j). \tag{2.84}$$

Moreover, for all $c_j, c'_j \in C_j$ with $c'_j > c_j$,

$$\int_{c_j}^{c'_j} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\tilde{\sigma}_{ij}(\mathrm{d}D|z) \right) \mathrm{d}z \geq \int_{c_j}^{c'_j} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\tilde{\sigma}_{ij}(\mathrm{d}D|c_j) \right) \mathrm{d}z, \tag{2.85}$$

and,

$$\int_{c_j}^{c'_j} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\tilde{\sigma}_{ij}(\mathrm{d}D|z) \right) \leq \int_{c_j}^{c'_j} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\tilde{\sigma}_{ij}(\mathrm{d}D|c'_j) \right). \tag{2.86}$$

These imply that, as $\overline{q}_{ij}(c_j) \in [0,1]$ for all $c_j \in C_j$, for all $j \in \mathcal{J}$,

$$\begin{aligned}
&\int_{c_j}^{c_j'} \left(\mathbb{E}_{c_{-j}} \left[\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \hat{\sigma}_{ij}(\mathrm{d}D | z, c_{-j}) q_{ij}(z, c_{-j}) \right] \right) \mathrm{d}z \\
&\quad - \int_{c_j}^{c_j'} \left(\mathbb{E}_{c_{-j}} \left[\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \hat{\sigma}_{ij}(\mathrm{d}D | c_j', c_{-j}) q_{ij}(c_j', c_{-j}) \right] \right) \mathrm{d}z \\
&= \int_{c_j}^{c_j'} \left(\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \tilde{\sigma}_{ij}(\mathrm{d}D | z) \overline{q}_{ij}(z) \right) \mathrm{d}z \\
&\quad - \int_{c_j}^{c_j'} \left(\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \tilde{\sigma}_{ij}(\mathrm{d}D | c_j') \overline{q}_{ij}(c_j') \right) \mathrm{d}z \qquad (2.87) \\
&\geq \int_{c_j}^{c_j'} \left(\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \overline{\sigma}_{ij}(\mathrm{d}D | z) \overline{q}_{ij}(z) \right) \mathrm{d}z \\
&\quad - \int_{c_j}^{c_j'} \left(\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \overline{\sigma}_{ij}(\mathrm{d}D | c_j') \overline{q}_{ij}(c_j') \right) \mathrm{d}z \\
&= \int_{c_j}^{c_j'} \left(\mathbb{E}_{c_{-j}} \left[\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma_{ij}(\mathrm{d}D | z, c_{-j}) q_{ij}(z, c_{-j}) \right] \right) \mathrm{d}z \\
&\quad - \int_{c_j}^{c_j'} \left(\mathbb{E}_{c_{-j}} \left[\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma_{ij}(\mathrm{d}D | c_j', c_{-j}) q_{ij}(c_j', c_{-j}) \right] \right) \mathrm{d}z \\
&\geq 0,
\end{aligned}$$

for all $j \in \mathcal{J}$, $c_j, c_j' \in C_j$, where the first equality follows from the definitions of $\hat{\sigma}_{ij}$ and $\overline{q}_{ij}$, the first inequality follows from (2.84), (2.85) and (2.86), the second equality follows from the definition of $\overline{\sigma}_{ij}$ and the last inequality follows from the incentive compatibility of

(σ, τ, q). Furthermore, for any $j \in \mathcal{J}$, $c_j \in C_j$,

$$\begin{aligned}
&\int_{c_j}^{\bar{c}_j} \left(\mathbb{E}_{c_j} \left[\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \hat{\sigma}_{ij}(\mathrm{d}D|z, c_{-j}) q(z, c_{-j}) \right] \right) \mathrm{d}z \\
&= \int_{c_j}^{\bar{c}_j} \left(\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \tilde{\sigma}_{ij}(\mathrm{d}D|z) \overline{q}_{ij}(z) \right) \mathrm{d}z \\
&= \int_{c_j}^{\bar{c}_j} \left(\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \overline{\sigma}_{ij}(\mathrm{d}D|z) \overline{q}_{ij}(z) \right) \mathrm{d}z \qquad (2.88) \\
&= \int_{c_j}^{\bar{c}_j} \left(\mathbb{E}_{c_j} \left[\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma_{ij}(\mathrm{d}D|z, c_{-j}) q(z, c_{-j}) \right] \right) \mathrm{d}z \\
&\geq \pi_j(c_j, m_j^*) - \pi_0(\bar{c})_j,
\end{aligned}$$

where the first equality follows from the definition of $\hat{\sigma}_{ij}$ and $\overline{q}_{ij}$, the second equality follows from (2.84), the third equality follows from the definition of $\overline{\sigma}_{ij}$ and the last inequality follows from the individual rationality of (σ, τ, q) and Lemma 2.18.

Together, by (2.87), (2.88), (2.81) and Lemma 2.18, there exist transfers $\{\hat{\tau}_j\}$ such that $(\hat{\sigma}, \hat{\tau}, q)$ is an incentive compatible and individually rational mechanism. Moreover, by (2.83), this mechanism improves the data broker's revenue. As such, (σ, τ, q) cannot be optimal. This completes the proof. ■

Proof of Theorem 2.8

Proof of Theorem 2.8. First notice that by Lemma 2.18, for any incentive feasible mechanism (σ, τ, q), the data broker's expected revenue is at most

$$\sum_{j \in \mathcal{J}} \left[\int_{C_j} \left(\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c_j) - \phi_{G_j}(c_j)) D(\overline{\boldsymbol{p}}_D(c_j)) \overline{\sigma}_{ij}(\mathrm{d}D|c_j) \overline{q}_{ij}(c) \right) G_j(\mathrm{d}c_j) - \pi_0(\bar{c})_j \right],$$

where for any $i \in \mathcal{I}$, $j \in \mathcal{J}$, $c_j \in C_j$,

$$\overline{\sigma}_{ij}(c_j) := \mathbb{E}_{c_j}[\sigma_{ij}(c)], \quad \overline{q}_{ij}(c_j) := \mathbb{E}_{c_{-j}}[q_{ij}(c)].$$

Furthermore, notice that for any $i \in \mathcal{I}$, $j \in \mathcal{J}$ and any $c_j \in C_j$,

$$\begin{aligned}\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c_j) - \phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D(c_j))\overline{\sigma}_{ij}(\mathrm{d}D|c_j) &\leq \int_{\mathcal{D}} \max_{p\in\mathbb{R}_+}(p - \phi_{G_j}(c_j))D(p)\overline{\sigma}_{ij}(\mathrm{d}D|c_j) \\ &\leq \int_V (v - \phi_{G_j}(c_j))^+ D_0^{ij}(\mathrm{d}v).\end{aligned}$$

Therefore,

$$\begin{aligned}&\sum_{j\in\mathcal{J}} \left[\int_{C_j} \left(\sum_{i\in\mathcal{I}} \int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c_j) - \phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D(c_j))\overline{\sigma}_{ij}(\mathrm{d}D|c_j)\overline{q}_{ij}(c)\right) G_j(\mathrm{d}c_j)\right] \\ &\leq \sum_{j\in\mathcal{J}} \left[\int_{C_j} \left(\sum_{i\in\mathcal{I}} \int_{\mathcal{D}} \max_{p\in\mathbb{R}_+}(p - \phi_{G_j}(c_j))D(p)\overline{\sigma}_{ij}(\mathrm{d}D|c_j)\overline{q}_{ij}(c)\right) G_j(\mathrm{d}c_j)\right] \\ &\leq \sum_{j\in\mathcal{J}} \left[\int_{C_j} \left(\int_V (v - \phi_{G_j}(c_j))^+ D_0^{ij}(\mathrm{d}v)\right) \overline{q}_{ij}(c_j) G_j(\mathrm{d}c_j)\right] \\ &= \int_C \sum_{j\in\mathcal{J}} \left(\sum_{i\in\mathcal{I}} \left(\int_V (v - \phi_{G_j}(c_j))^+ D_0^{ij}(\mathrm{d}v)\right) q_{ij}(c)\right) G(\mathrm{d}c).\end{aligned} \tag{2.89}$$

Since $\sum_{i\in\mathcal{I}} q_{ij} \leq 1$ for all $j \in \mathcal{J}$, the data broker's revenue is bounded from above by

$$\bar{R} := \sum_{j\in\mathcal{J}} \int_C \left[\max_{i\in\mathcal{I}} \left(\int_V (v - \phi_{G_j}(c_j))^+ D_0^{ij}(\mathrm{d}v)\right) G_j(\mathrm{d}c_j)\right] - \sum_{j\in\mathcal{J}} \pi_0(\bar{c})_j.$$

I first show that there exists an incentive feasible mechanism that attains the upper bound $\bar{R}$. To see this, notice that for any $i \in \mathcal{I}$, $j \in \mathcal{J}$, since $c_j \leq \phi_{G_j}(c_j) \leq \overline{\boldsymbol{p}}_{D_0^{ij}}(c_j)$ for all $c_j \in C_j$, as shown in the proof of Theorem 2.2, there exists a segmentation scheme

$\sigma^*_{ij} : C_j \to \mathcal{D}$ such that for all $c_j \in C_j$ and for any $p \in V$,

$$\int_{\mathcal{D}} D(p)\sigma^*_{ij}(\mathrm{d}D|c_j) = D_0^{ij}(p). \tag{2.90}$$

Moreover, for G_j-almost all $c_j \in C_j$,

$$\int_{\mathcal{D}} (\overline{\boldsymbol{p}}_D(c_j) - \phi_{G_j}(c_j))D(\overline{\boldsymbol{p}}_D(c))\sigma^*_{ij}(\mathrm{d}D|c_j) = \int_V (v - \phi_{G_j}(c_j))^+ D_0^{ij}(\mathrm{d}v). \tag{2.91}$$

Furthermore, for Lebesgue-almost all $c_j \in C_j$,

$$\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c_j))\sigma^*_{ij}(\mathrm{d}D|c_j) = D_0^{ij}(\phi_{G_j}(c_j)), \tag{2.92}$$

and for all $c_j, c'_j \in C_j$ with $c_j < c'_j$, then

$$\int_{c_j}^{c'_j} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z))\sigma^*_{ij}(\mathrm{d}D|c'_j) \right) \mathrm{d}z \leq \int_{c_j}^{c'_j} D_0^{ij}(\phi_{G_j}(z)) \, \mathrm{d}z, \tag{2.93}$$

while for Lebesgue-almost all $c_j, c'_j \in C_j$ with $c'_j < c_j$,

$$\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(c_j))\sigma^*_{ij}(\mathrm{d}D|c'_j) = D_0^{ij}(\phi_{G_j}(c'_j)). \tag{2.94}$$

Now let q^* be defined as

$$q^*_{ij}(c) := \frac{1}{|W_j(c_j)|} \mathbf{1} \left\{ \int_V (v - \phi_{G_j}(c_j))^+ D_0^{ij}(\mathrm{d}v) = \max_{i' \in \mathcal{I}} \int_V (v - \phi_{G_j}(c_j))^+ D_0^{i'j}(\mathrm{d}v) \right\},$$

where

$$W_j(c_j) := \left\{ i \in \mathcal{I} \middle| \int_V (v - \phi_{G_j}(c_j))^+ D_0^{ij}(\mathrm{d}v) = \max_{i' \in \mathcal{I}} \int_V (v - \phi_{G_j}(c_j))^+ D_0^{i'j}(\mathrm{d}v) \right\}.$$

Then the data broker's revenue under σ^* and q^* attains the upper $\bar{R}$. Furthermore, notice that since the function

$$z \mapsto \int_{\{v \geq z\}} (v-z) D(\mathrm{d}v)$$

is nonincreasing for any $D \in \mathcal{D}$ and since ϕ_{G_j} is nondecreasing, q^*_{ij} is nonincreasing for each $i \in \mathcal{I}$. As a result, for each $j \in \mathcal{J}$, for any $c_j, c'_j \in C_j$

$$\begin{aligned}
&\int_{c_j}^{c'_j} \left(\mathbb{E}_{c_{-j}} \left[\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma^*_{ij}(\mathrm{d}D|z, c_{-j}) q^*_{ij}(z, c_{-j}) \right] \right) \mathrm{d}z \\
&\quad - \int_{c_j}^{c'_j} \left(\mathbb{E}_{c_{-j}} \left[\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma^*_{ij}(\mathrm{d}D|c'_j, c_{-j}) \right] q^*_{ij}(c'_j, c_{-j}) \right) \mathrm{d}z \\
&= \int_{c_j}^{c'_j} \left(\int_{\mathcal{D}} \sum_{i \in \mathcal{I}} D(\overline{\boldsymbol{p}}_D(z)) \sigma^*_{ij}(\mathrm{d}D|z) q^*_{ij}(z) - \int_{\mathcal{D}} \sum_{i \in \mathcal{I}} D(\overline{\boldsymbol{p}}_D(z)) \sigma^*_{ij}(\mathrm{d}D|c'_j) q^*_{ij}(c'_j) \right) \mathrm{d}z \\
&\geq \int_{c_j}^{c'_j} \sum_{i \in \mathcal{I}} \left(D_0^{ij}(\phi_{G_j}(z)) q^*_{ij}(z) - D_0^{ij}(\phi_{G_j}(c'_j)) q^*_{ij}(c'_j) \right) \mathrm{d}z \\
&\geq 0,
\end{aligned} \tag{2.95}$$

where the equality follows from the fact that σ^*_{ij} and q^*_{ij} do not depend on c_{-j}, the first inequality follows from (2.92), (2.93), and (2.94); and the last inequality follows from monotonicity of ϕ_{G_j} and q^*_{ij} for each $i \in \mathcal{I}$.

Finally, notice that since for each $j \in \mathcal{J}$, $\{D_0^{ij}\}_{i \in \mathcal{I}}$ is ordered by pointwise dominance, for any $c_j \in C_j$, any $i, i' \in \mathcal{I}$

$$\begin{aligned}
&\int_V (v - \phi_{G_j}(c_j))^+ D_0^{ij}(\mathrm{d}v) \geq \int_V (v - \phi_{G_j}(c_j))^+ D_0^{i'j}(\mathrm{d}v) \\
&\iff m^{D_0^{ij}} \geq_{\text{FOSD}} m^{D_0^{i'j}} \\
&\iff D_0^{ij} \geq D_0^{i'j}.
\end{aligned} \tag{2.96}$$

As a result, for any $j \in \mathcal{J}$, for any $c_j \in C_j$,

$$\begin{aligned}
&\int_{c_j}^{\bar{c}_j} \left(\mathbb{E}_{c_{-j}} \left[\sum_{i \in \mathcal{I}} \int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma^*_{ij}(\mathrm{d}D|z, c_{-j}) q^*_{ij}(z, c_{-j}) \right] \right) \mathrm{d}z \\
&= \int_{c_j}^{\bar{c}_j} \sum_{i \in \mathcal{I}} \left(\int_{\mathcal{D}} D(\overline{\boldsymbol{p}}_D(z)) \sigma^*_{ij}(\mathrm{d}D|z) q^*_{ij}(z) \right) \mathrm{d}z \\
&= \int_{c_j}^{\bar{c}_j} \sum_{i \in \mathcal{I}} D^{ij}_0(\phi_{G_j}(z)) q^*_{ij}(z) \, \mathrm{d}z \\
&= \int_{c_j}^{\bar{c}_j} \sum_{i \in W_j(z)} D^{ij}_0(\phi_{G_j}(z)) \frac{1}{|W_j(z)|} \, \mathrm{d}z \\
&\geq \int_{c_j}^{\bar{c}_j} \sum_{i \in \mathcal{I}} D^{ij}_0(\phi_{G_j}(c_j)) \frac{1}{I} \, \mathrm{d}z \\
&\geq \int_{c_j}^{\bar{c}_j} D^{j}_0(\overline{\boldsymbol{p}}_{D^j_0}(z)) \, \mathrm{d}z,
\end{aligned} \tag{2.97}$$

where the first equality follows from the fact that σ^*_{ij} and q^*_{ij} do not depend on c_{-j} for all $i \in \mathcal{I}$, the second equality follows from (2.92), the third equality is from the definition of $\{q^*_{ij}\}$, the fourth equality follows from the definition of q^* and from (2.96), and the last equality follows from that fact that $\phi_{G_j} \leq \overline{\boldsymbol{p}}_{D^j_0}$ for all $j \in \mathcal{J}$.

Together, from (2.95) and (2.97), there exists transfers $\{\tau^*_j\}_{j \in \mathcal{J}}$ such that (σ^*, τ^*, q^*) is incentive compatible. Moreover, for each $j \in \mathcal{J}$, by taking $\tau^*_j(\bar{c}_j)$ as $\int_{\{v \geq \bar{c}_j\}} \sum_{i \in \mathcal{I}} (v - \bar{c}_j) D^{ij}_0(\mathrm{d}v) q^*_{ij}(\bar{c}_j) - \pi_0(\bar{c})_j$, together with (2.91), (2.97) and Lemma 2.1, the mechanism (σ^*, τ^*, q^*) is indeed incentive feasible and attains the upper bound $\bar{R}$.

Finally, it remains to show that the producers' gross expected profit and the allocation of the product under any optimal mechanism are the same under both regime $\mathcal{P}$ and regime $\mathcal{I}$. Since the optimal mechanism (σ^*, τ^*, q^*) constructed above attains the upper bound $\bar{R}$, all the inequalities in (2.89) are binding, by exactly the same arguments as in the proof of Theorem 2.1 and Theorem 2.4 and by noticing that any optimal mechanism (σ, τ, q) under regime $\mathcal{I}$ and any optimal mechanism $(\sigma, \tau, q, \boldsymbol{\gamma})$ under regime $\mathcal{P}$ for the data broker must entail

$$q_{ij}(c) > 0 \iff \int_V (v - \phi_{G_j}(c_j)) D_0^{ij}(\mathrm{d}v) = \max_{i' \in \mathcal{I}} \int_V (v - \phi_{G_j}(c_j)) D_0^{i'j}(\mathrm{d}v),$$

under any optimal mechanism of either regime $\mathcal{I}$ and $\mathcal{P}$, the allocation of the product must be such that for each product j, all the consumers in group $i(j)$ buys product j by paying their values and the rest of the consumers do not buy, where $i(j)$ is the group that prefers j the most (i.e. $i(j)$ is such that $D_0^{i(j)j} \geq D_0^{ij}$ for all $i \in \mathcal{I}$), while each producer $j \in \mathcal{J}$ must have expected profit

$$\max_{i \in \mathcal{I}} \int_{C_j} \int_{\{v \geq \phi_{G_j}(c_j)\}} (v - c_j) D_0^{ij}(\mathrm{d}v),$$

which are exactly the allocation and the gross profit producer $j \in \mathcal{J}$ earns under both regimes. This completes the proof. ■

Appendix for Chapter 3

A Envelope Characterizations

Proof of Lemma 3.1

Proof of Lemma 3.1. For necessity, Given an incentive compatible mechanism $(\boldsymbol{D}, \tau, \boldsymbol{\gamma})$, let

$$U(c, c') := \int_{\mathbb{R}_+} (p - c)\boldsymbol{D}(p|c')\boldsymbol{\gamma}(\mathrm{d}p|c') - \tau(c') \tag{3.5}$$

for all $c, c' \in C$ be the seller's net profit when her true cost is c reports c'. Incentive compatibility implies that

$$U(c) := \max_{c' \in C} U(c, c') = U(c, c), \ \forall c \in C.$$

Notice that for any $c' \in C$, the function $U(\cdot, c')$ is affine and thus absolutely continuous, and the family $\{U(\cdot, c')\}_{c' \in C}$ is uniformly bounded. As such, by the envelope theorem (Milgrom and Segal, 2002), U is absolutely continuous and

$$U(c) = U(\overline{c}) + \int_c^{\overline{c}} \left(\int_{\mathbb{R}_+} \boldsymbol{D}(p|z)\boldsymbol{\gamma}(\mathrm{d}p|z) \right) \mathrm{d}z. \tag{3.6}$$

Furthermore, since U is a pointwise supremum of a family of affine functions, U is convex and therefore its subgradiant, $-\int_{\mathbb{R}_+} D(p|c)\boldsymbol{\gamma}(\mathrm{d}p|c)$, is nondecreasing. This establishes assertion 2. On the other hand, using (3.6) and rearrange (3.5), we have

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

This then establishes assertion 1 after defining $\bar{\tau} := -U(\overline{c})$.

For sufficiency, suppose that a mechanism $(\boldsymbol{D}, \tau, \boldsymbol{\gamma})$ satisfies assertion 1 and 2. Again, let $U(c,c')$ be defined as (3.5) and let

$$\Pi(c,c') := \int_{\mathbb{R}_+} (p-c)\boldsymbol{D}(p|c')\boldsymbol{\gamma}(\mathrm{d}p|c').$$

for all $c, c' \in C$ and notice that $\Pi(\cdot, c')$ is differentiable for all $c' \in C$ and its derivative is $\Pi_1(c,c') = -\int_{\mathbb{R}_+} \boldsymbol{D}(p|c)\boldsymbol{\gamma}(\mathrm{d}p|c)$. Then for any $c, c' \in C$,

$$\begin{aligned}
&U(c,c) - U(c,c') \\
=&\int_c^{c'} \left(\int_{\mathbb{R}_+} \boldsymbol{D}(p|z)\boldsymbol{\gamma}(\mathrm{d}p|z)\right)\mathrm{d}z - (\Pi(c,c') - \Pi(c',c')) - \int_{c'}^{\overline{c}} \left(\int_{\mathbb{R}_+} \boldsymbol{D}(p|z)\boldsymbol{\gamma}(\mathrm{d}p|z)\right)\mathrm{d}z \\
=&\int_c^{c'} \left(\int_{\mathbb{R}_+} \boldsymbol{D}(p|z)\boldsymbol{\gamma}(\mathrm{d}p|z)\right)\mathrm{d}z - (\Pi(c,c') - \Pi(c',c')) \\
=&\int_c^{c'} \left(\int_{\mathbb{R}_+} \boldsymbol{D}(p|z)\boldsymbol{\gamma}(\mathrm{d}p|z)\right)\mathrm{d}z - \int_c^{c'} \Pi_1(z,c')\,\mathrm{d}z \\
=&\int_c^{c'} \left(\int_{\mathbb{R}_+} \boldsymbol{D}(p|z)\boldsymbol{\gamma}(\mathrm{d}p|z) - \int_{\mathbb{R}_+} \boldsymbol{D}(p|c')\boldsymbol{\gamma}(\mathrm{d}p|c')\right)\mathrm{d}z \\
\geq&0,
\end{aligned}$$

where the first equality follows from assertion 1, the third equality follows from the fundamental theorem of calculus and the last inequality follows from assertion 2. This completes the proof. ■

Proof of Lemma 3.2

Proof of Lemma 3.2. For necessity, consider any incentive compatible and individually rational mechanism $(\alpha, \boldsymbol{D}, \tau)$. Let

$$\Pi(c,c') := \alpha(c') \cdot \max_{p \geq 0} (p-c)\boldsymbol{D}(p|c')$$

be the seller's expected profit under the information structure $\boldsymbol{D}(c') \in \mathcal{D}_0$, publicizing policy $\alpha(c') \in \{0, 1\}$ and cost c. By the envelope theorem (Milgrom and Segal, 2002), since the function

$$c \mapsto \alpha(c')(p - c)\boldsymbol{D}(p|c')$$

is absolutely continuous with uniformly bounded (almost everywhere) derivative for any fixed $p \geq 0$ and $c' \in C$, $\Pi(\cdot, c')$ is absolutely continuous for all $c' \in C$ and its derivative exists and equals to

$$\Pi_1(c, c') = -\alpha(c')\boldsymbol{D}(\boldsymbol{p}_{\boldsymbol{D}(c')}(c)|c') \tag{3.7}$$

for any selection $\boldsymbol{p} \in \boldsymbol{P}$ and for (Lebesgue) almost all $c \in C$.

Now let

$$U(c, c') := \Pi(c, c') - \tau(c')$$

be the seller's profit net of transfer if the cost is c and the (mis)report is c'. Incentive compatibility then implies

$$U(c) := U(c, c) = \max_{c' \in C} U(c, c').$$

Since $\Pi(\cdot, c')$ is absolutely continuous and uniformly bounded, by the envelope theorem again,

$$U(c) = U(\overline{c}) - \int_c^{\overline{c}} \Pi_1(z, z)\, \mathrm{d}z = U(\overline{c}) + \int_c^{\overline{c}} \boldsymbol{D}(\boldsymbol{p}_{\boldsymbol{D}(z)}(z)|z)\, \mathrm{d}z.$$

Rearranging, we have:

$$\tau(c) = -U(\overline{c}) + \alpha(c) \cdot \max_{p \geq 0}(p - c)\boldsymbol{D}(p|c) - \int_c^{\overline{c}} \alpha(z)\boldsymbol{D}(\boldsymbol{p}_{\boldsymbol{D}(z)}(z)|z)\, \mathrm{d}z,$$

which establishes assertion 1.

In addition, by assertion 1, for any $c, c' \in C$,

$$
\begin{aligned}
&\int_c^{c'} [\alpha(z)\boldsymbol{D}(\boldsymbol{p}_{\boldsymbol{D}(z)}(z)|z) - \alpha(c')\boldsymbol{D}(\boldsymbol{p}_{\boldsymbol{D}(c')}(z)|c')]\,\mathrm{d}z \\
=&\int_c^{c'} \alpha(z)\boldsymbol{D}(\boldsymbol{p}_{\boldsymbol{D}(z)}(z)|z)\,\mathrm{d}z - \int_c^{c'} \Pi_1(z, c')\,\mathrm{d}z \\
=&\int_c^{c'} \alpha(z)\boldsymbol{D}(\boldsymbol{p}_{\boldsymbol{D}(z)}(z)|z)\,\mathrm{d}z - (\Pi(c, c') - \Pi(c', c')) \\
=&\int_c^{\overline{c}} \alpha(z)\boldsymbol{D}(\boldsymbol{p}_{\boldsymbol{D}(z)}(z)|z)\,\mathrm{d}z - (\Pi(c, c') - \Pi(c', c')) - \int_{c'}^{\overline{c}} \alpha(z)\boldsymbol{D}(\boldsymbol{p}_{\boldsymbol{D}(z)}(z)|z)\,\mathrm{d}z \\
=&U(c, c) - U(c, c') \\
\geq&0,
\end{aligned}
\tag{3.8}
$$

where the first equality follows from (3.7), the second equality follows from the fundamental theorem of calculus and the last equality follows from assertion 1. This establishes assertion 2.

Conversely, consider any mechanism $(\alpha, \boldsymbol{D}, \tau)$ that satisfies assertions 1 and 2. Again, let $\Pi(c, c') := \max_{p \geq 0} \alpha(c')(p - c)\boldsymbol{D}(p|c')$ and let $U(c, c') := \Pi(c, c') - \tau(c')$. By assertions 1 and 2, (3.7) and (3.8), for any $c, c' \in C$,

$$U(c, c) - U(c, c') = \int_c^{c'} [\alpha(z)\boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(z)}(z)|z) - \alpha(c')\boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c')}(z)|c')]\,\mathrm{d}z \geq 0,$$

where the inequality follows from condition 2. This completes the proof. ■

Proof of Theorem 3.1

Proof of Theorem 3.1. By Lemma 3.1, the intermediary's revenue maximization problem under regime $\mathcal{P}$ can be rewritten as:

$$\sup_{\boldsymbol{D},\boldsymbol{\gamma}} \int_C \left(\int_{\mathbb{R}_+} (p - \phi_G(c)) \boldsymbol{D}(p|c) \boldsymbol{\gamma}(\mathrm{d}p|c) \right) G(\mathrm{d}c) \tag{3.9}$$

$$\text{s.t. } c \mapsto \int_{\mathbb{R}_+} \boldsymbol{D}(p|c) \boldsymbol{\gamma}(\mathrm{d}p|c) \text{ is nonincreasing} \tag{3.10}$$

To solve (3.9), a pointwise maximization approach can be adopted. Consider any $c \in C$, any $p \geq 0$ and any $D \in \mathcal{D}_0$. Notice that

$$\begin{aligned} & (p - \phi_G(c)) D(p) \\ \leq & (p - \phi_G(c)) D(p) + \int_p^\infty D(v) \, \mathrm{d}v \\ \leq & \int_{\phi_G(c)}^\infty D(v) \, \mathrm{d}v \\ \leq & \int_{\phi_G(c)}^\infty D_0(v) \, \mathrm{d}v. \end{aligned} \tag{3.11}$$

Furthermore, a binary information structure $\boldsymbol{D}_\mathrm{b}(c) \in \mathcal{D}_0$, defined as:

$$\boldsymbol{D}_\mathrm{b}(p|c) := \begin{cases} 0, & \text{if } p \in [0, \mathbb{E}_{m^0}[v | v \leq \phi_G(c)]) \\ D_0(\phi_G(c)), & \text{if } p \in [\mathbb{E}_{m^0}[v|v \leq \phi_G(c)], \mathbb{E}_{m^0}[v|v \geq \phi_G(c)]] \\ 1, & \text{if } p \in (\mathbb{E}_{m^0}[v|v \geq \phi_G(c)], \infty) \end{cases},$$

together with the pricing scheme $\boldsymbol{\gamma}_{\rm b}(c)$ that assigns probability one on $v(c) := \mathbb{E}_{m^0}[v|v \geq \phi_G(c)]$, attains this upper bound. That is,

$$(v(c) - \phi_G(c))\boldsymbol{D}_{\rm b}(v(c)|c) = (v(c) - \phi_G(c))D_0(\phi_G(c)) = \int_{\phi_G(c)}^{\infty} D_0(v)\,{\rm d}v.$$

As such, for each $c \in C$, the information structure $\boldsymbol{D}_{\rm b}(c)$ and the price $\boldsymbol{\gamma}_{\rm b}(c)$ is a solution of

$$\sup_{\gamma \in \Delta(\mathbb{R}_+), D \in \mathcal{D}_0} \int_{\mathbb{R}_+} (p - \phi_G(c))D(p)\gamma({\rm d}p).$$

Furthermore, consider the mechanism $(\boldsymbol{D}_{\rm b}, \tau_{\rm b}, \boldsymbol{\gamma}_{\rm b})$, where $\tau_{\rm b}$ is defined by assertion 1 of Lemma 3.1. Since ϕ_G is increasing by regularity, $\int_{\mathbb{R}_+} \boldsymbol{D}_{\rm b}(p|c)\boldsymbol{\gamma}_{\rm b}({\rm d}p|c) = D_0(\phi_G(c))$ is nonincreasing, by Lemma 3.1, the mechanism $(\boldsymbol{D}_{\rm b}, \tau_{\rm b}, \boldsymbol{\gamma}_{\rm b})$ is incentive compatible and individually rational. Together, $(\boldsymbol{D}_{\rm b}, \tau_{\rm b}, \boldsymbol{\gamma}_{\rm b})$ maximizes the intermediary's revenue under regime $\mathcal{P}$ and yields revenue R^*, as desired. ■

Proof of Theorem 3.2

Proof of Theorem 3.2. Since the optimal revenue is weakly higher for the intermediary under regime $\mathcal{P}$ than under regime $\mathcal{I}$, by Theorem 3.1, it suffices to show that the mechanism $(\alpha_{\rm u}, \boldsymbol{D}_{\rm u}, \tau_{\rm u})$ is incentive feasible and generates revenue R^*.

To this end, recall that

$$R^* := \int_C \left(\int_{\phi_G(c)}^{\infty} D_0(v)\,{\rm d}v \right) G({\rm d}c) = \int_C (v(c) - \phi_G(c))D_0(\phi_G(c))G({\rm d}c),$$

where the second equality follows from integration by parts. Furthermore, notice that under the upper censorship mechanism $(\alpha_{\rm u}, \boldsymbol{D}_{\rm u}, \tau_{\rm u})$, if each truthful-reporting seller whose cost is

$c \in C$ sets price optimally at $v(c)$, then for all $c \in C$,

$$(\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c) - \phi_G(c))\boldsymbol{D}_{\mathrm{u}}(\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c)|c) = (v(c) - \phi_G(c))D_0(\phi_G(c)) \tag{3.12}$$

Therefore, it suffices to show that (3.12) holds for the upper censorship mechanism $(\alpha_{\mathrm{u}}, \boldsymbol{D}_{\mathrm{u}}, \tau_{\mathrm{u}})$ and that this mechanism is incentive feasible, as this would imply that the mechanism $(\alpha_{\mathrm{u}}, \boldsymbol{D}_{\mathrm{u}}, \tau_{\mathrm{u}})$ is feasible and attains the upper bound R^* of problem (3.2).

Indeed, first notice that for any $c \in C$, under the upper censorship $\boldsymbol{D}_{\mathrm{u}}(c)$, for a seller with cost c, setting prices with $p < \phi_G(c)$ gives profit

$$(p - c)D_0(p) \leq (\phi_G(c) - c)D_0(\phi_G),$$

since the function $p \mapsto (p - c)D_0(p)$ is single-peaked with a peak at $\overline{\boldsymbol{p}}_0(c)$ by regularity and since $\phi_G(c) \leq \overline{\boldsymbol{p}}_0(c)$ by (2.7). Meanwhile, setting any prices $p \in [\phi_G(c), v(c))$ must be worse than setting price at $v(c)$ since $\boldsymbol{D}_{\mathrm{u}}(\cdot|c)$ is a constant on $[\phi_G(c), v(c))$. Finally, for any $p > v(c)$, the seller gets zero profit by setting a price at p as $\boldsymbol{D}_{\mathrm{u}}(p|c) = 0$. Together, for the truthfully-reporting seller with cost c, setting price at $v(c)$ is indeed optimal.

Moreover, notice that for any $c, c' \in C$ such that $c' \leq c$, either $\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') = v(c)$, or $\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') = \overline{\boldsymbol{p}}_0(c')$. To see this, first notice that since $c' \leq c$ and $\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') = v(c)$, it must be that $\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') \leq v(c)$. Suppose that $\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') < v(c)$. Then it must be that $\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') < \phi_G(c)$, as $\boldsymbol{D}_{\mathrm{u}}(\cdot|c)$ is constant on $[\phi_G(c), v(c)]$. By definition of $\boldsymbol{D}_{\mathrm{u}}(c)$, for any $p \leq \phi_G(c)$,

$$(p - c')\boldsymbol{D}_{\mathrm{u}}(p|c) \leq (\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') - c')\boldsymbol{D}_{\mathrm{u}}(\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c')|c) = (\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') - c')D_0(\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c')|c).$$

Thus, the function $p \mapsto (p - c')D_0(p)$ has a peak at $\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') < \phi_G(c)$. By regularity of m^0, this function is single-peaked. Therefore, $\overline{\boldsymbol{p}}_{\boldsymbol{D}_{\mathrm{u}}(c)}(c') = \overline{\boldsymbol{p}}_0(c')$. Together with (2.7), this implies that

$$\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(c') \geq \overline{\boldsymbol{p}}_0(c') \geq \phi_G(c'). \tag{3.13}$$

As a result,

$$\begin{aligned}
&\int_{c'}^{c} [\boldsymbol{D}_{\mathrm{u}}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(z)}(z)|z) - \boldsymbol{D}_{\mathrm{u}}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(z)|c)]\,\mathrm{d}z \\
=&\int_{c'}^{c} [D_0(\phi_G(z)) - \boldsymbol{D}_{\mathrm{u}}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(z)|c)]\,\mathrm{d}z \\
\leq&\int_{c'}^{c} [D_0(\phi_G(z)) - \boldsymbol{D}_{\mathrm{u}}(\phi_G(z)|c)] \\
=&\int_{c'}^{c} [D_0(\phi_G(z)) - D_0(\phi_G(z))]\,\mathrm{d}z = 0,
\end{aligned}$$

where the first equality follows from the definition of $\boldsymbol{D}_{\mathrm{u}}(z)$ and the fact that optimal price for the seller with cost z under $\boldsymbol{D}_{\mathrm{u}}(z)$ is $v(z)$; the inequality follows from (3.13), and the second equality follows from the definition of $\boldsymbol{D}_{\mathrm{u}}(c)$ and the fact that $\phi_G(z) \leq \phi_G(c)$ for all $z \in [c', c]$.

On the other hand, if $c' \in (c, v(c))$, then for any $z \in [c, c']$, $\boldsymbol{D}_{\mathrm{u}}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(z)|c) = D_0(\phi_G(c)) \geq D_0(\phi_G(z))$ by construction of $\boldsymbol{D}_{\mathrm{u}}(c)$ and by monotonicity of ϕ_G. As such,

$$\int_{c'}^{c} [\boldsymbol{D}_{\mathrm{u}}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(z)}(z)|z)] - \boldsymbol{D}_{\mathrm{u}}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(z)|c)]]\,\mathrm{d}z = \int_{c}^{c'} [D_0(\phi_G(c)) - D_0(\phi_G(z))]\,\mathrm{d}z \geq 0.$$

Finally, if $c' > v(c)$, optimal price under $\boldsymbol{D}_{\mathrm{u}}(c)$ gives zero profit and thus deviation gain must be negative.

Together with Lemma 3.1, the upper censorship mechanism $(\alpha_{\mathrm{u}}, \boldsymbol{D}_{\mathrm{u}}, \tau_{\mathrm{u}})$ is indeed incentive compatible and individually rational. Moreover, under this mechanism, the intermediary's expected revenue is

$$\int_C (v(c) - \phi_G(c))D_0(c)G(\mathrm{d}c) = R^*.$$

This completes the proof. ■

Proof of Theorem 3.3

Proof of Theorem 3.3. Equivalence in the intermediary's revenue is a direct consequence of Theorem 3.1 and Theorem 3.2.

For buyer's surplus and seller's profit. First notice under regime $\mathcal{P}$, as optimal revenue is R^*, which is attained by a pointwise maximum of the objective of (3.9), for any other mechanism $(\boldsymbol{D}, \tau, \boldsymbol{\gamma})$ that attains revenue R^*, by (3.11), all the weak inequalities must hold with equality. In particular, for G-almost all $c \in C$

$$\int_{\mathbb{R}_+} \int_p^\infty \boldsymbol{D}(v|c)\,\mathrm{d}v \boldsymbol{\gamma}(\mathrm{d}p|c) = 0$$

and thus the buyer's expected surplus must be zero under any optimal mechanism. On the other hand, also from (3.11), for G-almost all $c \in C$, it must be that $\boldsymbol{D}(p|c) = \boldsymbol{D}(v(c)|c)$ for all $p \in \mathrm{supp}(\boldsymbol{\gamma}(c))$ and that

$$\int_{\phi_G(c)}^\infty \boldsymbol{D}(v|c)\,\mathrm{d}v = \int_{\phi_G(c)}^\infty D_0(v)\,\mathrm{d}v.$$

Together with $\boldsymbol{D}(c) \in \mathcal{D}_0$ for all $c \in C$, this implies that for G-almost all $c \in C$,

$$\int_{\mathbb{R}_+} \boldsymbol{D}(p|c)\boldsymbol{\gamma}(\mathrm{d}p|c) = D_0(\phi_G(c)).$$

Therefore, together with Lemma 3.1, the seller's profit must be

$$\int_C \left(\int_c^{\overline{c}} D_0(\phi_G(z))\,\mathrm{d}z \right) G(\mathrm{d}c)$$

under any optimal mechanism.

Meanwhile, under regime $\mathcal{I}$, since the optimal revenue is also R^*, for any optimal mechanism $(\alpha, \boldsymbol{D}, \tau)$, all the weak inequalities in (3.11) must be equalities. In particular, it must be that $\alpha \equiv 1$,

$$\int_{\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(c)}^{\infty} \boldsymbol{D}(v|c)\,\mathrm{d}v = 0,$$

$$\boldsymbol{D}(p|c) = \boldsymbol{D}(\phi_G(c)|c), \forall p \in [\phi_G(c), \overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(c)),$$

and

$$\int_{\phi_G(c)}^{\infty} \boldsymbol{D}(v|c)\,\mathrm{d}v = \int_{\phi_G(c)}^{\infty} D_0(v)\,\mathrm{d}v,$$

for G-almost all $c \in C$. Together with $\boldsymbol{D}(c) \in \mathcal{D}_0$ for all $c \in C$, we have

$$\boldsymbol{D}(\phi_G(c)|c) = \boldsymbol{D}(\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(c)|c) = D_0(\phi_G(c)),$$

for G-almost all $c \in C$, which in turn implies that

$$\overline{\boldsymbol{p}}_{\boldsymbol{D}(c)}(c) = v(c),$$

for G-almost all $c \in C$. Together, the buyer's expected surplus under any optimal mechanism is zero, and the seller's expected profit under any optimal mechanism is

$$\int_C \left(\int_c^{\overline{c}} D_0(\phi_G(c))\,\mathrm{d}z \right) G(\mathrm{d}c),$$

This completes the proof. ■

Proof of Theorem 3.4

Proof of Theorem 3.4. As in the proof of Theorem 3.1 and Theorem 3.2, I first find upper bounds for the objective in (3.3) and (3.4). Under regime $\mathcal{P}$, for any measurable set $A \subseteq V$, any $D \in \mathcal{D}_A$, any $c \in C$ and any $p \geq 0$,

$$\begin{aligned}
&(p - \phi_G(c))D(p) \\
\leq &(p - \phi_G(c))D(p) + \int_p^\infty D(v)\,\mathrm{d}v \\
\leq &\int_{\phi_G(c)}^\infty D(v)\,\mathrm{d}v \\
\leq &\int_{\phi_G(c)}^\infty D_A(v)\,\mathrm{d}v \\
\leq &\int_{\phi_G(c)}^\infty D_0(v)\,\mathrm{d}v.
\end{aligned} \tag{3.14}$$

On the other hand, under regime $\mathcal{I}$, for any measurable set $A \subseteq V$, any $D \in \mathcal{D}_A$, any $c \in C$ and any $\boldsymbol{p} \in \boldsymbol{P}$,

$$\begin{aligned}
&(\boldsymbol{p}_D(c) - \phi_G(c))D(\boldsymbol{p}_D(c)) \\
\leq &\max_{p \geq 0}(p - \phi_G(c))D(p) \\
\leq &\max_{p \geq 0}(p - \phi_G(c))D(p) + \int_{\boldsymbol{p}_D(c)}^\infty D(v)\,\mathrm{d}v \\
\leq &\int_{\phi_G(c)}^\infty D(v)\,\mathrm{d}v \\
\leq &\int_{\phi_G(c)}^\infty D_A(v)\,\mathrm{d}v \\
\leq &\int_{\psi(c)}^\infty D_0(v)\,\mathrm{d}v.
\end{aligned} \tag{3.15}$$

Therefore, for both regimes, R^* is still an upper bound of the intermediary's revenue. It is then clear that the intermediary's optimal revenue under these two regimes must be R^*, since the optimal mechanisms constructed in the proof of Theorem 3.1 and Theorem 3.2 are still feasible in the environment. To see the equivalence for the rest of market outcomes, consider any optimal mechanism $(\boldsymbol{A}, \boldsymbol{D}, \tau, \boldsymbol{\gamma})$ under regime $\mathcal{P}$ and any optimal mechanism $(\boldsymbol{A}, \boldsymbol{D}, \tau)$ under regime $\mathcal{I}$. For buyer's surplus and seller's profit, notice that under regime

$\mathcal{P}$ and under the optimal mechanism $(\boldsymbol{A}, \boldsymbol{D}, \tau, \boldsymbol{\gamma})$, the weak inequalities in (3.14) must hold with equality and thus, as argued in the proof of Theorem 3.3, it must be that buyer's expected surplus is zero and the seller's expected profit is

$$\int_C \left(\int_c^{\overline{c}} D_{\boldsymbol{A}(z)}(\phi_G(z))\,\mathrm{d}z \right) G(\mathrm{d}c) = \int_C \left(\int_c^{\overline{c}} D_0(\phi_G(z))\,\mathrm{d}z \right) G(\mathrm{d}c).$$

On the other hand, under regime $\mathcal{I}$, under $(\boldsymbol{A}, \boldsymbol{D}, \tau)$, all the weak inequalities in (3.15) must hold with equality and thus, as in the proof of Theorem 3.3, the buyer's expected surplus is zero and the seller's profit is

$$\int_C \left(\int_c^{\overline{c}} D_{\boldsymbol{A}(z)}(\phi_G(z)))\,\mathrm{d}z \right) G(\mathrm{d}c) = \int_C \left(\int_c^{\overline{c}} D_0(\phi_G(z))\,\mathrm{d}z \right) G(\mathrm{d}c).$$

This completes the proof. ■

C Comparative Statics

Proof of Proposition 3.1

Proof of Proposition 3.1. Notice that for each $c \in C$, probability of efficient trade is the probability of the event that trade occurs when the buyer's value is greater than the seller's cost. Since $\phi_G(c) > c$ for all $c \in C$,

$$\int_C D_0(c)G(\mathrm{d}c) > \int_C D_0(\phi_G(c))G(\mathrm{d}c),$$

which implies that the probability of efficient trade is larger when the seller has control of the information technology.

On the other hand, since ϕ_G is increasing and $\phi_G(c) > c$ for all $c \in C$,

$$\begin{aligned}
&\int_{\phi_G(c)}^{\infty} D_0(v)\,\mathrm{d}v + (\phi_G(c) - c)D_0(\phi_G(c)) \\
&< \int_{\phi_G(c)}^{\infty} D_0(v)\,\mathrm{d}v + \int_{c}^{\phi_G(c)} D_0(v)\,\mathrm{d}v \\
&= \int_{c}^{\infty} D_0(v)\,\mathrm{d}v,
\end{aligned}$$

for all $c \in C$. Thus,

$$\int_C (v(c) - c)D_0(\phi_G(c))G(\mathrm{d}c) < \int_C (v(c) - c)D_0(c)G(\mathrm{d}c).$$

This completes the proof. ■

Proof of Proposition 3.2

Proof of Proposition 3.2. For 1., recall that that under both regimes, the intermediary's revenue is

$$R^* := \int_C (v(c) - \phi_G(c))D_0(\phi_G(c))G(\mathrm{d}c),$$

and total surplus is

$$\int_C (v(c) - c)D_0(\phi_G(c))G(\mathrm{d}c),$$

and the seller's expected net profit is

$$\int_C \left(\int_c^{\overline{c}} D_0(\phi_G(z))\,\mathrm{d}z \right) G(\mathrm{d}c).$$

Thus, for any $i, j \in \{1,2\}$, let $v_i^j(c) := \mathbb{E}_{m_i^0}[v|v > \phi_{G_j}(c)]$ for all $c \in C$, and let $D_0^j(p) := m_j^0([p, \overline{v}])$ for all $p \geq 0$. As such, for any $i \in \{1,2\}$,

$$(v_1^i(c) - \phi_{G_i}(c))D_0^1(\phi_{G_i}(c)) = \int_{\phi_{G_i}(c)}^{\infty} D_0^1(v)\,\mathrm{d}v$$
$$\geq \int_{\phi_{G_i}(c)}^{\infty} D_0^2(v)\,\mathrm{d}v = (v_2^i(c) - \phi_{G_i}(c))D_0^2(\phi_{G_i}(c)),$$

and

$$(v_1^i(c) - c)D_0^1(\phi_{G_i}(c)) = \int_{\phi_{G_i}(c)}^{\infty} D_0^1(v)\,\mathrm{d}v + (\phi_{G_i}(c) - c)D_0^1(\phi_G(c))$$
$$\geq (v_1^i(c) - c)D_0^2(\phi_{G_i}(c)) = \int_{\phi_{G_i}(c)}^{\infty} D_0^2(v)\,\mathrm{d}v + (\phi_{G_i}(c) - c)D_0^2(\phi_G(c)),$$

and also

$$\int_c^{\overline{c}} D_0^1(\phi_{G_i}(z))\,\mathrm{d}z \geq \int_c^{\overline{c}} D_0^2(\phi_{G_i}(z))\,\mathrm{d}z,$$

for all $c \in C$ and therefore

$$\int_C (v_1^i(c) - \psi_i(c))D_0^1(\phi_{G_i}(c))G_i(\mathrm{d}c) \geq \int_C (v_2^i(c) - \psi_i(c))D_0^1(\phi_{G_i}(c))G_i(\mathrm{d}c)$$

and

$$\int_C (v_1^i(c) - c)D_0^1(\phi_{G_i}(c))G_i(\mathrm{d}c) \geq \int_C (v_2^i(c) - c)D_0^2(\phi_{G_i}(c))G_i(\mathrm{d}c),$$

and also

$$\int_C \left(\int_c^{\overline{c}} D_0^1(\phi_{G_i}(z))\,\mathrm{d}z \right) G_i(\mathrm{d}c) \geq \int_C \left(\int_c^{\overline{c}} D_0^2(\phi_{G_i}(z))\,\mathrm{d}z \right) G_i(\mathrm{d}c),$$

for any $i \in \{1,2\}$.

For 2., notice that by using integration by parts, for all $c \in C$, $i, j \in \{1, 2\}$,

$$\int_{\phi_{G_i}(c)}^{\infty} D_0^j(v)\,\mathrm{d}v = \int_0^{\infty} \mathbf{1}\{v \geq \phi_{G_i}(c)\} D_0^j(v)\,\mathrm{d}v = \int_0^{\infty} (v - \phi_{G_i}(c))^+ D_0^j(\mathrm{d}v).$$

Therefore, since the function $v \mapsto (v - \psi(c))^+$ is convex and since m_1^0 is a mean preserving spread of m_2^0, for any $i \in \{1, 2\}$,

$$\begin{aligned}
&\int_C (v_1^i(c) - \phi_{G_i}(c)) D_0^1(\phi_{G_i}(c)) G_i(\mathrm{d}c) \\
&= \int_C \left(\int_0^{\infty} (v - \phi_{G_i}(c))^+ D_0^1(\mathrm{d}v) \right) \\
&\geq \int_C \left(\int_0^{\infty} (v - \phi_{G_i}(c))^+ D_0^2(\mathrm{d}v) \right) \\
&= \int_C (v_2^i(c) - \psi_i(c)) D_0^2(\phi_{G_i}(c)) G_i(\mathrm{d}c)
\end{aligned}$$

For 3., first notice that the hazard rate dominance implies that $\phi_{G_1} \leq \phi_{G_2}$ and that

$$G_1(c) = \exp\left(- \int_c^{\overline{c}} \frac{1}{\phi_{G_1}(z)}\,\mathrm{d}z \right) \geq \exp\left(- \int_c^{\overline{c}} \frac{1}{\phi_{G_2}(z)}\,\mathrm{d}z \right) = G_2(c).$$

That is, G_2 first order stochastic dominates G_1. As such, for each $i \in \{1, 2\}$,

$$\begin{aligned}
&\int_C (v_i^1(c) - \phi_{G_1}(c)) D_0^i(\phi_{G_1}(c)) G_1(\mathrm{d}c) \\
&= \int_C \left(\int_{\phi_{G_1}(c)}^{\infty} D_0^i(v)\,\mathrm{d}v \right) G_1(\mathrm{d}c) \\
&\geq \int_C \left(\int_{\phi_{G_2}(c)}^{\infty} D_0^i(v)\,\mathrm{d}v \right) G_1(\mathrm{d}c) \\
&\geq \int_C \left(\int_{\phi_{G_2}(c)}^{\infty} D_0^i(v)\,\mathrm{d}v \right) G_2(\mathrm{d}c) \\
&= \int_C (v_i^2(c) - \phi_{G_2}(c)) D_0^i(\phi_{G_2}(c)) G_2(\mathrm{d}c),
\end{aligned}$$

where the first inequality follows from $\phi_{G_1} \leq \phi_{G_2}$ and the second inequality follows from the fact that G_2 first order stochastic dominates G_1 and that ϕ_{G_2} is increasing. Similarly, for each $i \in \{1, 2\}$,

$$\begin{aligned}
&\int_C (v_i^1(c) - c) D_0^i(\phi_{G_1}(c)) G_1(\mathrm{d}c) \\
&= \int_C \left(\int_{\phi_{G_1}(c)}^{\infty} D_0^i(v)\,\mathrm{d}v + (\phi_{G_1}(c) - c) D_0^i(\phi_{G_1}(c)) \right) G_1(\mathrm{d}c) \\
&\geq \int_C \left(\int_{\phi_{G_2}(c)}^{\infty} D_0^i(v)\,\mathrm{d}v + (\phi_{G_2}(c) - c) D_0^i(\phi_{G_2}(c)) \right) G_1(\mathrm{d}c) \\
&\geq \int_C \left(\int_{\phi_{G_2}(c)}^{\infty} D_0^i(v)\,\mathrm{d}v + (\phi_{G_2}(c) - c) D_0^i(\phi_{G_2}(c)) \right) G_2(\mathrm{d}c) \\
&= \int_C (v_i^2(c) - c) D_0^i(\phi_{G_2}(c)) G_2(\mathrm{d}c).
\end{aligned}$$

Finally, for the same reasons,

$$\int_C \left(\int_c^{\overline{c}} D_0^i(\phi_{G_1}(z))\,\mathrm{d}z \right) G_1(\mathrm{d}c) \geq \int_C \left(\int_c^{\overline{c}} D_0^i(\phi_{G_2}(z))\,\mathrm{d}z \right) G_2(\mathrm{d}c)$$

This completes that proof. ■

REFERENCES

AGUIRRE, I., S. COWAN, AND J. VICKERS (2010): "Monopoly Price Discrimination and Demand Curvature," *American Economic Review*, 100, 1601–1615.

BARON, D. AND R. MYERSON (1982): "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50, 911–930.

BERGEMANN, DIRK, B. B. AND S. MORRIS (2017): "Informationally Robust Optimal Auction Design," Working Paper.

BERGEMANN, D. AND A. BONATTI (2015): "Selling Cookies," *American Economic Journal: Microeconomics*, 7, 259–294.

BERGEMANN, D., A. BONATTI, AND A. SMOLIN (2018): "The Design and Price of Information," *American Economic Review*, 108, 1–45.

BERGEMANN, D., B. BROOKS, AND S. MORRIS (2015): "The Limits of Price Discrimination," *American Economic Review*, 105, 921–957.

BERGEMANN, D. AND M. PESENDORFER (2007): "Information Structures in Optimal Auctions," *Journal of Economic Theory*, 137, 580–609.

BLACKWELL, D. (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

BROOKS, B. AND S. DU (2019): "Optimal Auction Design with Common Values: An Informationally-Robust Approach," Working Paper.

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

Du, S. (2018): "Robust Mechanism under Common Valuation," *Econometrica*, 86, 1569–1588.

Dubé, J.-P. and S. Misra (2017): "Scalable Price Targeting," Working Paper.

Dworczak, P. (2020): "Mechanism Design with Aftermarkets: Cutoff Mechanisms," Working Paper.

Dworczak, P. and G. Martini (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy,* forthcoming.

Esö, P. and B. Szentes (2007): "Optimal Information Disclosure in Auctions and the Handicap Auction," *Review of Economic Studies*, 74, 705–731.

Gentzkow, M. and E. Kamenica (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review, Papers & Proceeding*, 106, 597–601.

Goldberg, P. K. (1996): "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, 104, 622–654.

Hart, S. and P. Reny (2019): "The Better Half of Selling Separately," ACM *Transactions on Economics and Computation*, 7.

Johnson, J. P. and D. P. Myatt (2006): "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review*, 96, 756–784.

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

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

Kolotilin, A., M. Li, T. Mylovanov, and A. Zapechelnyuk (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85, 1949–1964.

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

——— (1991): "Supplying Information to Facilitate Price Discrimination," *International Economic Review*, 35, 309–327.

Li, H. and X. Shi (2017): "Discriminatory Information Disclosure," *American Economic Review*, 107, 3363–3385.

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

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

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

Mikians, J., L. Gyarmati, V. Erramilli, and N. Laoutaris (2012): "Detecting price and search discrimination on the internet," *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, 79–84.

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

Monteiro, P. K. and B. F. Svaiter (2010): "Optimal Auction with a General Distribution: Virtual Valuation without Densities," *Journal of Mathematical Economics*, 46, 21–31.

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

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

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

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

Neeman, Z. (2003): “The Effectiveness of English Auctions,” *Games and Economic Behavior*, 43, 214–218.

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

Ravid, D., A.-K. Roseler, and B. Szentes (2019): “Learning Before Trading: On the Inefficiency of Ignoring Free Information,” Working Paper.

Rockafellar, R. T. (1970): *Convex Analysis*, Princeton, NJ: Princeton University Press.

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

Segura-Rodriguez, C. (2019): “Selling Data,” Working Paper.

Shi, X. (2012): “Optimal Auctions with Information Acquisition,” *Games and Economic Behavior*, 74, 666–686.

Shiller, B. and J. Waldfogel (2011): “Music for a Song: An Empirical Look at Uniform Song Pricing and its Alternatives,” *Journal of Industrial Economics*, 59, 622–654.

Shiller, B. R. (2014): "First-Degree Price Discrimination Using Big Data," Working Paper.

Skreta, V. and E. Perez-Richet (2018): "Test Design under Falsification," Working Paper.

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

Yang, K. H. (2019a): "Buyer-Optimal Information with Nonlinear Technology," Working Paper.

——— (2019b): "Equivalence in Business Models of Informational Intermediary," Working Paper.

——— (2019c): "Informationally Robust Welfare Predictions under Second-Degree Price Discrimination," Working Paper.

——— (2020): "Selling Consumer Data for Profit: Optimal Market-Segmentation Design and its Consequences," Working Paper.