POLITECNICO DI TORINO

Corso di Laurea
in Ingeneria Matematica

Tesi di Laurea Magistrale

Price Discrimination on the Spokes Model with Data Sales

Relatori
Prof. Carlo Cambini
Dott. Flavio Pino

Candidato
Domenico Bianco

Anno Accademico 2022-2023

Summary

This Master's Thesis examines the strategies employed by a Data Broker (DB) who sells information to firms operating in a non-localized spatial competition environment represented by the Spokes Model. Each firm perceives itself as competing in various submarkets, categorized based on the availability of consumers' first and second preferred brands. In certain submarkets, the firm acts as a monopolist, serving captive consumers with no alternative brand option. In other submarkets, the firm operates as a duopolist, competing against an alternative brand. The focus is on a market characterized by low-utility valued products, where uninformed firms are motivated to carve out a portion of captive consumers facing an elastic monopolistic demand.

The DB offers a data segment of equal length to all firms through a Take-It-or-Leave-It (TIOLI) arrangement, maximizing firms' willingness to pay while leveraging the threat of an Outside Option. The data enable first-degree price discrimination, specifically designed for competitive and monopolistic markets, while allowing informed firms to serve at a unique basic price all unidentified consumers. Data have mainly three effects: firstly, they reduce the scope of non-discriminatory markets, thereby facilitating monopolistic market coverage; secondly, they intensify competition among firms for consumers in competitive sub-markets; thirdly, they grant total surplus extraction from captive consumers. In contrast to well-known oligopolistic spatial price discrimination models (i.e. Hotelling), even when the data broker offers overlapping segments that include non-exclusive information about all consumers in competitive markets, each informed firm does not lower its basic price to zero but adjusts it to serve unidentified captive consumers.

When the number of firms is sufficiently low and there is a large turf of captive consumers, firms strive to leverage the rent extraction from monopolistic sub-markets and avoid basic prices war for unidentified consumers in competitive sub-markets. Consequently, they target a specific kink of the demand function and grant full market coverage with marginal captive consumers' surplus at zero. In this scenario, the primary focus of the DB is to enable every informed company to maximize surplus extraction from monopolistic sub-markets. This involves selling equal-sized data partition, almost in its entirety, to all firms involved. However, when there is a low market concentration informed firms succeed in maximizing the rent extraction from monopolistic sub-markets only when they are not forced to defend their non-discriminatory consumers in each duopolistic segment. Indeed, when data segments do not overlap there is a fringe of unidentified consumers in competitive sub-markets that triggers a price war. We find that the DB finds it convenient to mitigate competition among firms by providing a low quantity of non-exclusive information. Furthermore, even in cases where duopolistic segments dominate, the presence of a small group of captive consumers, where informed firms can fully exploit surplus, significantly hampers the performance of the uninformed firm. As a result, even in this scenario, the DB chooses to sell equal-sized partitions encompassing nearly all available data to maximize the threat posed by being uninformed and ensures full market coverage.

Contents

List of Tables 4

List of Figures 5

1 The Role of Data 7
- 1.1 Discrimination without strategic interactions 10
- 1.2 Discrimination thorough DB 15

2 Price discrimination on Spokes 21
- 2.1 The Spokes Model: baseline framework 21
- 2.2 Model Setup 27
  - 2.2.1 Selling mechanism and timing 28
  - 2.2.2 All Firms informed 29
  - 2.2.3 One firm uninformed 32
- 2.3 Equilibrium Prices and DB's strategies 33
  - 2.3.1 Monopolistic Sub-Markets Prevalence 35
  - 2.3.2 Competitive Sub-Markets Prevalence 45
  - 2.3.3 DB's Optimum Strategy 53
- 2.4 Conclusions 56

A Proofs and Definitions 57

List of Tables

1.1 Strategic Choice of Spatial Price Policy in normal form 13

List of Figures

1.1 Spatial pricing strategy in the Hotelling Framework . . . 12
2.1 Spokes Model for $N = 5, n = 3$ . . . 22
2.2 Spokes Model for $N = 5, n = 3, v \in [\frac{3}{2}, 2]$ . . . 27
2.3 Spatial pricing strategy when all firms (W) hold d data . . . 31
2.4 Spatial pricing strategy when only one firm (L) holds no information . . . 33
2.5 Inside Option profits and basic prices in the region of M.S.P., for $d \in [0,1]$, $n \in [2, n_2^I(v) = 53]$, $v = \frac{19}{10}$ and $N = 100$ . . . 36
2.6 Prices and profits comparison between Hotelling and Spokes Model when all firms have the same amount of information . . . 38
2.7 Outside Option profits and prices in the region of M.S.P. when the DB offers non-overlapping segments ($d \in [0, \frac{1}{2})$) for $n \in [2, n_2^I(v) = 53]$, $v = \frac{19}{10}$ and $N = 100$ . . . 40
2.8 Outside Option profits and prices when the DB offers overlapping segments ($d \in [\frac{1}{2}, 1]$) for $n \in [2, 99]$, $N = 100$ and $v = \frac{19}{10}$ . . . 42
2.9 Outside Option profits and basic prices in the region of M.S.P., for $d \in [0,1]$, $v = \frac{19}{10}$ and $N = 100$ . . . 43
2.10 DB's Profits in the region of Monopolistic Sub-Markets Prevalence for $d \in [0,1]$, $n \in [2, n_2^I(v) = 53]$, $v = \frac{19}{10}$ and $N = 100$ . . . 44
2.11 Inside Option prices and profits in the region of Competitive Sub-Markets Prevalence for $d \in [0,1]$, $N = 100$ and $v = \frac{19}{10}$ . . . 46
2.12 Outside Option profits and prices in the region of Competitive Sub-Markets Prevalence for $d \in [0,1]$, $N = 100$, $v = \frac{19}{10}$ . . . 49
2.13 DB's Profits in the region of Competitive Sub-Markets Prevalence for $d \in [0,1]$, $v = \frac{19}{10}$ and $N = 100$ . . . 51
2.14 DB's Optimal Strategy in the Region of Competitive Sub-Markets Prevalence and in the exclusive information scenario for $N = 100$ and $v \in [\frac{42}{25}, \frac{9}{5}]$ . . . 52
2.15 DB's Profits and Optimal Strategies for $N = 100$, $v = \frac{19}{10}$ . . . 55

Chapter 1

The Role of Data

Digital economics refers to the branch of economics that focuses on the study of economic phenomena in the context of the digital world. It examines how digital technologies and the internet impact economic activities, markets, industries, and overall economic systems. Digital economics encompasses various areas, including digital markets, online platforms, e-commerce, digital currencies, digital labor, and the overall digital transformation of the economy. It investigates the impact of digitalization on market structures, competition, pricing, and consumer behavior exploring the dynamics of online-market places.

According to the Report UE on "Competition Policy for the Digital Era" (Crémer et al. [2019]), this new branch of economics is characterized by the following three main features:

- **Extremes Returns to Scale**: the cost of production is much less than proportional to the number of customers served. While this has always been true to some extent, as bigger factories or retailers are often more efficient than smaller ones, the digital world pushes this phenomenon to the extreme. Once created, information can be transmitted to a large number of people at a very low cost.

- **Network Externalities**: the usefulness for each user of using a technology or a service increases as the number of users increases. They can be characterized by their two-sided nature, which means that they create benefits or costs for two distinct groups of users or participants in a platform or network.

- **The Role of Data**: advancements in technology have facilitated the accumulation, retention, and utilization of vast volumes of data by companies. This transformative capability has already brought about significant alterations to market dynamics and is expected to do so in the future as well. Data serves as a fundamental component for artificial intelligence (AI) and intelligent online services, playing a vital role in production processes, logistics, and targeted marketing. The capacity to leverage data and create novel, inventive applications and products has become a pivotal factor in competitiveness, and its significance is poised to grow further.

This work concerns the role of data as information that possesses inherent value for the entities that exercise control over them, whether directly or indirectly. This value lies in the potential for enabling more informed decision-making by economic actors. From an individual's perspective, the value of information corresponds to the amount they are willing to pay to enhance the quality of their decisions.

Goldfarb and Tucker [2019] describe data as enablers of five types of downward shifting costs. The reduction in search and replication costs has far-reaching effects on prices, price dispersion, product variety, and media availability. It influences matches across various domains, from labor markets to dating, and has led to the rise of platform-based businesses, impacting the organizational structure of firms. Digital interactions and transactions are readily captured and preserved, as digital activity can be easily recorded and stored. In practice, numerous web servers automatically store data, and companies must actively choose to discard it. The decline in tracking costs facilitates personalized experiences and the emergence of one-to-one markets. As a result, there is a renewed focus on economic models that involve asymmetric information and differentiated products. This includes areas such as auctions and advertising models, which leverage personalization and targeted approaches made possible by the abundance of digital data.

As outlined by Pino [2022], when revising the literature on digital economics, the assumptions regarding the process of data collection are a strong driver for the models' market outcomes, regardless of the specific data use. We will concentrate on the following two types of data acquisition:

- **No Strategic Interactions**: firms obtain data through non-strategic means, without engaging in deliberate interactions with other actors. This can be observed in scenarios where data is readily available to firms from external sources, or when firms incur a marginal cost specifically for data acquisition.

- **Data Intermediaries**: with the proliferation of online platforms and the vast amount of information generated and shared by users, data has become a valuable currency in the digital landscape. Data intermediaries play a pivotal role in this ecosystem, facilitating the exchange of information between various stakeholders and giving increasing importance to information markets in economic activity and welfare. Every day, consumers participate in a wide range of online and offline actions that expose personal information about themselves. These activities can include using mobile devices, searching for a new home or car, subscribing to magazines, making purchases either in-store or through catalogs, browsing the internet, completing surveys to receive coupons, utilizing social media platforms, subscribing to online news sites, or entering sweepstakes. During these interactions, the entities with which consumers engage gather information about them and often share or sell that information to data brokers or attention platforms:

  - **Data Brokers**: According to the FTC Report (2014), DBs typically do not get their data directly from consumers but rather, they collect data from numerous other sources, which fall into three categories: government sources, other publicly available sources and commercial sources. Furthermore, the nine data brokers studied in the FTC report obtain most of their data from other data brokers rather than directly from an original source. In theoretical models, data brokers are commonly depicted as entities that either possess consumer data or acquire it by incurring a marginal cost. The strategic interaction in these models typically occurs between the data broker and downstream firms, while consumers are not directly involved in the strategic dynamics. The most common examples of such interactions are represented by the audience segments sold by Nielsen, Acxiom, and Epsilon.

  - **Attention Platforms**: Attention platforms operate by attracting users to their platform and encouraging them to spend time and engage in various activities on the platform, both practical and social. The value of user attention stems from two key factors. Firstly, the usage data generated provides the attention broker with valuable proprietary information regarding the search and communication behaviors of users. Secondly, the attention broker can leverage this user attention by selling targeted advertising space to firms operating in industries related to the user's interests, such as refrigerator makers or local plumbers, which we refer to as the retail product industry (Prat and Valletti [2022]).

Relying on targeting and tracking technologies, firms can engage in price discrimination: examples can be found in industries such as personalized healthcare services, professional consulting, customized luxury goods, or some online platforms that utilize personalized pricing based on user data and preferences. Even the first-degree price discrimination, which was nothing more than a mirage, now is becoming a reality. In this form of price discrimination, the seller maximizes their profit by charging each customer the highest price they are willing to pay for a product or service and capturing the entire surplus.

The focus of this work will be on data as a tool that allows customer identification, enabling price discrimination.

Within the context of price discrimination, privacy issues are a significant concern in terms of:

- **Data collection and consent**: the collection and use of data may not always be transparent or accompanied by clear consent mechanisms. Consumers may be unaware of the extent of data being collected and how it is being used for price discrimination purposes;
- **Lack of control and transparency**: consumers may have limited control over how their data is collected, stored, and shared in digital markets. They may not be aware of the specific entities accessing their data or how it is being used to determine prices. This lack of transparency undermines consumer trust and their ability to make informed decisions about their data;
- **Secondary use of data**: personal data collected for one purpose, such as completing a transaction, can be used for unrelated activities without explicit consent. This secondary use of data raises concerns about privacy and fairness, as consumers may not anticipate or approve of their data being used for price discrimination or other purposes;

To address these privacy concerns, regulatory frameworks such as the GDPR in the European Union and similar data protection laws in other jurisdictions aim to protect individuals' privacy rights. These regulations emphasize the need for clear consent, data minimization, purpose limitation, data accuracy, and enhanced security measures.

1.1 Discrimination without strategic interactions

When companies have external access to data and use them for price discrimination, the economics literature finds it intriguing to examine spatial competition settings characterized by oligopoly or duopoly, where the possession of such data plays a crucial role as a competitive advantage. These frameworks of differentiated oligopolistic competition (vertically or horizontally) aim to measure the impact of data in terms of firm profits, consumer surplus, and endogenous entry addressing the following questions:

- How would a firm with data price its product?
- How does a firm's pricing strategy depend on whether competitors also possess detailed profiles about some customers?
- How does pricing strategy affect product variety and firms' profits?
- Do customers, given that price targeting requires access to customers' profiles, find it convenient to protect their information or are they better off with discrimination?

To provide some initial intuition regarding the forces at play in spatial models of price discrimination, we will consider the Hotelling [1929] model on the unit interval. Firm A (located at $x = 0$) competes with firm B (located at $x = 1$). The consumer located at $x \in [0,1]$ has utility $v - p_A - tx$ if he buys from A and $v - p_B - t(1-x)$ if he buys from B.
The indifferent consumer is located at: $\hat{x} = \frac{1}{2} + \frac{p_B - p_A}{2t}$.
Each firm maximizes its profit and the results are the following:

- $p_A^* = p_B^* = t$ and $\hat{x}^* = \frac{1}{2}$;
- $\pi_A^* + \pi_B^* = t$;
- $CS = v - \frac{5t}{4}$.

In microeconomics, there are several types of price discrimination strategies that firms might employ to maximize their profits based on varying consumer preferences and willingness to pay. These include:

- **First-Degree Price Discrimination**: This involves charging each customer a unique price based on their individual willingness to pay. This type of discrimination results in capturing the entire consumer surplus for the firm;
- **Second-Degree Price Discrimination**: Firms offer different prices based on the quantity of a good or service purchased. This encourages customers to buy more by offering lower prices for larger quantities. Examples include bulk discounts and tiered pricing;

- **Third-Degree Price Discrimination**: Firms divide the market into different segments based on characteristics such as age, location, income, etc. Each segment is charged a different price that aligns with their willingness to pay. This is commonly seen in student discounts, senior discounts, and regional pricing.

If one of the two firms holds information about certain consumers and can engage in first-degree price discrimination, it possesses precise knowledge of the exact willingness to pay for each individual identified customer. As a result, it will establish a customized price that considers its competitors' prices, to render the consumer equally inclined towards either option. In particular, in the Hotelling framework, data as a competitive advantage can be seen as a quantity $d \in [0,1]$ of the unit interval (but can be any spatial configuration of consumer types) where firms know the exact willingness to pay of each consumer.

As in Thisse and Vives [1988], we grant to both (D, D) companies or one of them (U, D) the ability to engage in first-degree price discrimination on a quantity $d \in [0,1]$ of consumers while allowing the company without data to set a uniform price.

When both firms have a quantity $d = 1$ of data (D, D), prices will be driven downward to match the difference in transportation costs as follows:

$$\begin{cases} p_A^*(x) = t(1-2x) \mid p_B^*(x) = 0 \mid\mid x \leq \frac{1}{2} \\ p_A^*(x) = 0 \mid p_B^*(x) = t(2x-1) \mid\mid x > \frac{1}{2} \end{cases}$$

In the (D, D) setting, the results are the following:

- $\pi_A^* + \pi_B^* = \frac{t}{2}$;
- $CS = v - \frac{3t}{4}$.

It is important to observe that when comparing the results between a scenario where companies charge equal prices (U, U), and individualized pricing (D, D), all consumers experience improved outcomes. Consumers located at points 0 and 1 are offered identical prices under both conditions and are indifferent to the change. However, other consumers are strictly better off in the (D, D) situation: instead of competing for the marginal consumer, firms compete for each consumer individually. As a result, prices decrease, and some of the profits previously earned by firms are transferred to consumers.

In the (D, U) setting only one firm can discriminate (A) and the firm without data (B) is the price leader, so firm A will set its discriminatory price in a second stage.

Given the uniform price set by firm B, firm A makes the consumer indifferent between the two options until the discriminatory price is greater than zero:

$$p_A^* = p_B + t(1-2x) \mid\mid 0 \leq x \leq d$$

This implies the existence of a threshold value of d beyond which the data no longer have any impact on prices.This value corresponds to the condition $p_B + t(1-2x) = 0$, that is to say $d = \frac{p_B+t}{2t}$.

When firm A discriminate on a quantity $d = \frac{p_B+t}{2t}$ of the unit interval and firm B has no information about consumer types, the results are the following:

(a) Both Informed (b) One Informed

Figure 1.1. Spatial pricing strategy in the Hotelling Framework

- $p_B^* = \frac{t}{2}$, $p_A^* = p_B^* + t(1-2x)$
- $\pi_A^* = \frac{9t}{16}, \pi_B^* = \frac{t}{8}$;
- $CS = v - t$.

Prices are lower when consumer information is available compared to the scenario where no information is present. Consequently, the consumer surplus reaches its peak when both firms possess consumer information, and it reaches its lowest point when neither firm has access to such information. Additionally, a firm with consumer information is more profitable than an uninformed competitor. Furthermore, in the setting (D, U), the informed firm generates higher profits compared to all previous cases examined.

These settings will result in a game under simultaneous moves which can be represented in the normal form (Table 1.1) and matches the strategic choice of spatial price policy game analyzed by Thisse and Vives [1988].

The equilibrium of the game represented in Table 1.1 arises when both firms adopt first-degree price discrimination. When the rival firm does not commit to price discrimination, a firm's profits are enhanced by implementing price discrimination. Conversely, if the rival commits to price discrimination, the firm minimizes its losses by also committing to it. Essentially, firms face a prisoner's dilemma situation where, although not engaging in price discrimination would be more advantageous, it becomes their optimal response to the rival's strategy.

To better understand the outcomes of this game, we need to delve deeper into the two forces that drive spatial price discrimination models (Liu and Serfes [2004]):

- **Surplus Extraction Effect**: Firms can customize their targeted offers based on consumers' willingness to pay: this allows them to extract higher profits from consumers;
- **Intensified Competition Effect**: since both firms can send targeted offers to all consumers, they anticipate their rival strategy and engage in price wars. Firms thus lower their tailored prices until they match the difference in consumers' willingness to pay between the two firms: this effect dissipates profits

In the above-stylized examples, we have assumed that all firms are capable of perfectly discriminating each consumer. However, by studying the impact of different levels of information quality (Liu and Serfes [2004]), it can be shown that when information quality is extremely low, equilibrium is reached with firms unilaterally committing to a uniform

A \ B	U	D
U	$\frac{t}{2}, \frac{t}{2}$	$\frac{t}{8}, \frac{9t}{16}$
D	$\frac{9t}{16}, \frac{t}{8}$	$\frac{t}{4}, \frac{t}{4}$

Table 1.1. Strategic Choice of Spatial Price Policy in normal form

price, even if the cost of information is zero. However, beyond a certain threshold of information quality, such a commitment is no longer an equilibrium and the game becomes a Prisoner's Dilemma. Acquiring information becomes the dominant strategy for each firm, leading to lower profits compared to those obtained under a uniform pricing rule. Consequently, as the quality of information regarding customers' preferences improves, firms are expected to completely abandon policies that aim to restrict the practice of price discrimination.

In the work of Taylor and Wagman [2014] the answers to some of the above questions and the balancing between the Intensified Competition Effect and the Surplus Extraction Effect emerge by examining four fundamental models that are commonly used in the literature: a linear city model (LCM), a circular city model (CCM), a vertical differentiation model (VDM), and a multi-unit symmetric demand model (MSDM). In particular, these models are studied under privacy enforcement and without privacy. In the privacy setting all the firms have no discriminatory power and set uniform prices. In the no-privacy configuration, all firms know all consumer types and can engage in first-degree price discrimination.

The impact of price discrimination (no-privacy enforcement) differs among models, leading to varying outcomes in terms of efficiency and deadweight loss. Generally, incorporating privacy measures often results in less efficiency and higher deadweight loss. Additionally, consumer preferences regarding privacy tend to vary: consumers with higher demand parameters for a particular product tend to prioritize privacy, while those with lower demand parameters lean towards not having privacy.

A setting of spatial price discrimination on the Salop model is outlined by Abrardi et al. [2022]: on a circular city of length 1, n symmetric and equally spaced firms (indexed by $i \in \{0, \ldots, n-1\}$) compete for a mass 1 of consumers uniformly distributed on the circle (located at $x \in [0,1)$ in counter-clockwise order). Each firm i performs first-degree price discrimination on the arch d_i (centered on the location of each firm i). A firm i offers location-specific tailored prices $p_i(x)$ to the consumer x in the identified segment and a basic price p_i^B to the unidentified consumer. Indifferent consumers, firm i best response function, and profit are as follows:

$$\hat{x}_{i-1,i} = \frac{2i-1}{n} + \frac{p_i^B - p_{i-1}^B}{2t} \quad \hat{x}_{i,i+1} = \frac{2i+1}{n} + \frac{p_{i+1}^B - p_i^B}{2t}$$

$$p_i^{B^{BR}}(d_i, p_{-i}^B) = \frac{t}{2n} - \frac{td_i}{2} + \frac{p_{i+1}^B + p_{i-1}^B}{4}$$

$$\pi_i = \int_{\frac{i}{n}-\frac{d_i}{2}}^{\frac{i}{n}+\frac{d_i}{2}} p_i(x)\,dx + p_i^B(\hat{x}_{i,i+1} - \hat{x}_{i-1,i} - d_i) - F$$

Jointly considering these equations, we find the two aforementioned effects of data on firm's profit. On one hand, a firm sets tailored prices extracting surplus from identified consumers (the first term of the profit function). On the other hand, as firm i acquires more data, its anonymous consumers are on average farther from its location and its basic price gets lower (through the term $-\frac{td_i}{2}$).

The aforementioned effects of privacy regulation on market outcomes are valid also in a duopolistic setting (Chen et al. [2020]) where each firm has an arbitrary segment of data and consumers engage in identity management. In this setting, an active consumer can potentially choose from three prices (a personalized price from the targeting firm and two uniform prices) and a passive consumer faces at most two prices (a personalized price from the targeting firm and the uniform price from the rival firm).

First, the implementation of personalized pricing enables firms to defend their target segments more effectively compared to when they charge prices at higher levels of aggregation. This allows firms to defend their targeted consumers individually and, if necessary, even lower the price for the marginal consumer to zero. Second, firms display more aggression in poaching targeted consumers from their rivals when their targeted consumers are passive rather than active. This is because passive targeted consumers do not have the option to choose the firm's uniform poaching price, which enables the firm to disconnect its personalized prices from the poaching price. As a result, the cost of poaching increases, and more active consumers lead to a softer competition environment.

Market asymmetries can also influence the surplus extraction effect and intensified competition effect, leading to diverse outcomes. When two firms exogenously have data on all consumers and are vertically differentiated in a spatial competition setting, the high-quality firm can expand its market share, increasing profits (Shaffer and Zhang [2002]). Instead, if only one firm has data, semi-collusive behavior can arise through a first-mover advantage (Gu et al. [2019]): the informed firm sets a high price, enabling his rival to undercut him and thus avoid a price war.

In a homogeneous duopolistic product setting (Belleflamme et al. [2020]), for given uniform prices, personalized prices decrease as firms' profiling technologies become more symmetric, which is due to the higher intensity of competition between firms for consumers profiled by both firms. As a result, when both firms profile the same set of consumers, or only one firm profiles consumers, marginal cost pricing arises. At the same time, they achieve positive expected profits at the expense of consumers in the presence of different tracking abilities.

1.2 Discrimination thorough DB

The importance of information markets in economic activity and welfare has increased significantly, partly due to the abundance of available data sources. However, trading information goes beyond simply selling access to a database. The ability to collect, analyze, and extract insights from large datasets presents opportunities for exchanging information in various forms such as predictions, ratings, recommendations, and customization of products and services. Simultaneously, the mechanisms involved in trading information give rise to new challenges related to privacy, the market power of information intermediaries, and potential distortions in the information sector and other sectors.

The most recent and steadily growing literature strand focuses on data intermediaries: these firms collect massive amounts of data that they then sell in downstream markets. Their upstream position allows them to consider data externalities to their full extent: while an individual firm aims at maximizing its profits, a data intermediary's goal is to maximize the value of data in the downstream market that it can then extract from purchasing firms. In particular, the literature has been split into two major strands: one regarding data brokers and the other regarding attention platforms. These intermediaries shape the market for information along two key dimensions (FTC, 2014):

- **Only Information**: the data broker can provide the data buyer with a new list of prospects or append information about an individual (or a group) that the buyer has already identified. According to the FTC (2014) report, marketing, and lead-generation companies, as well as providers of financial data like Bloomberg, primarily sell original lists. An original list typically represents a customer segment, which is a group of potential consumers with specific characteristics. Instead, data appends involve revealing additional information about a company's existing or potential customers. In the marketing context, companies like Nielsen Catalina Solutions and Oracle Datalogix connect an individual's offline and online purchases with the digital media they consume.
- **Direct Access to Consumer**: In various markets, information is not only directly sold but also indirectly offered through customized goods and services. One example of this is the sale of contextualized original lists to gain access to consumers. A prime illustration of such a market is sponsored-search advertising, as seen on platforms like Google or Bing. In this scenario, the search engine possesses information in the form of user search queries. Using this information, the search engine can potentially provide recommendations or predictions to advertisers regarding user preferences. When considering multiple users, this can be seen as purchasing an original list of selected consumers. However, it's important to note that search engines employ a different and more profitable strategy for monetizing their information. They grant access to a targeted audience by selling advertising slots specifically tied to keyword searches.

The focus of this work is on data as enablers of price discrimination mechanisms. In this setting, looking at the consumer as an agent who reveals the information either directly or indirectly through their past behavior and purchases, explains why in the market for data there is room for data intermediaries to make profits (Bergemann and Bonatti [2019]). As the information is ultimately utilized for price differentiation, individual consumers demand compensation in exchange for sharing their information. However, since each consumer's information encompasses both their unique preferences and the overall demand patterns, the individual consumer can only request compensation based on marginal value. On the other hand, the data intermediary has the ability to charge the seller for the complete value of the demand information. Consequently, there exists a conflict between marginal pricing concerning the consumer and average pricing concerning the producer. This conflict creates opportunities for an intermediary with market power to inefficiently utilize and transfer information (Bergemann and Bonatti [2019]).

Our objective is to understand how the pricing and selling mechanisms within the market for information influence competition in the downstream markets.

In this section, we will dive deep into the role of data brokers. These intermediaries in theoretic models are usually modeled as actors who already have consumer data or collect them by paying a marginal cost: the strategic interaction only happens between the data broker and downstream firms, not consumers.

Compared to settings where firms obtain data without strategic interactions, the primary distinction lies in the fact that data brokers take into account the consequences of selling data to one company on other firms. This external effect affects the value of data for these companies and, consequently, the profits of the data broker.

The part of the literature that we are interested in for the purposes of this work focuses on data brokers who sell strategic information to firms operating in spatial competition settings: the models of spatial duopolistic and oligopolistic price discrimination described in the previous section are incorporated into a sequential game where the initial stages are focused on the sale of data.

Consider a setting, as in Thisse and Vives [1988], where consumers have horizontal preferences uniformly distributed on the unit line and they benefit from the intensified competition that arises when all firms perform first-degree price discrimination across the entire market.

Assume the presence of two distinct consumer groups referred to as "new" and "old" consumers (Montes et al. [2019]). In the case of the first group, individual information cannot be obtained, however, for the second group, information may be accessible through the option to purchase from a data supplier (each group is uniformly distributed on a Hotelling line). Two firms (located at the extremes of the line) can acquire information on consumers' tastes from a monopolistic intermediary that sells all data in one block.

The Data Broker sells information through an auction with externalities as in Jehiel and Moldovanu [2000]. If the DB sells exclusively to a single firm the maximum price, denoted as T_1, that he can set is determined by the difference between the profits of the winning and losing firms in the auction. Specifically, T_1 represents the difference between the firm's profits when it possesses consumer information and the firm's profits when its rival possesses that information. When the DB sells to both firms, then T_2 represents the difference between the case when both firms can use consumer information and when only one firm can use this information. The DB profits are:

$$\pi_{DB} = \max\{T_1, 2T_2\}$$

Using an auction, the data intermediary leverages the negative externality related to the threat of being uninformed, which increases the willingness to pay of a prospective buyer.

In this duopoly setting, the result is influenced by the informational framework, specifically which firms obtain information about old consumers. When the dataset is sold as a single unit, the data broker opts to sell consumer data exclusively to one firm (Montes et al. [2019]). Subsequently, the uninformed firm sets a price below the Hotelling price, while the informed firm adopts a less aggressive approach toward acquiring new consumers. As for the old consumers, the informed firm matches the benefits they could receive by purchasing from the other firm but does not seize the entire market.

If old consumers can pay for a price to avoid being discriminated against, prices are expected to be higher compared to the Hotelling case, and they will decrease as the cost of privacy increases (Montes et al. [2019]). This is because, with a small privacy cost, the majority of old consumers prefer purchasing from the anonymous market. Consequently, the uninformed firm can generate significant profits by focusing on this market and setting a higher price. As the privacy cost increases, the size and preferences of consumers in the anonymous market change, leading to a more aggressive approach from both firms. Also in this case, as long as the data are sold in one block, the DB prefers to deal exclusively with one firm.

Weakening the assumption that the DB sells all the data, Bounie et al. [2021] build a model where a monopolistic data broker can sell information that partitions the Hotelling unit line into segments of arbitrary sizes to one or two competing firms (indexed $\theta = 1{,}2$). Thinner segments give more precise information, but are more costly to collect.

The information structure in this model is characterized by dividing the unit line into n segments of flexible sizes. These segments are formed by combining elementary segments, each with a size of $\frac{1}{k}$, where k is a predetermined integer representing the quality of information. Denote S the set comprising the $k-1$ endpoints of the segments of size $\frac{1}{k}$. A bijection $(M : S \to \mathbb{P})$ that associates to any subset $\{\frac{s_1}{k}, \ldots, \frac{s_{n-1}}{k}\} \in S$ a partition $\{[0, \frac{s_1}{k}], [\frac{s_1}{k}, \frac{s_2}{k}], \ldots, [\frac{s_{n-1}}{k}, 1]\}$, outlines the sigma-field $\mathbb{P}$ of all possible partitions of the unit line.

The DB can sell any partition P_θ to firm θ. In fact, starting from any pair of partitions, when the DB sells information to both firms, it will sell the same partition. Denote with I an informed firm and with NI a firm without information. As $\pi_{P,\theta}^{NI,I} = \pi_{P,\theta}^{I,NI}$, it follows that there are three possible configurations of profits:$\{\pi_{P,\theta}^{NI,NI}, \pi_{P,\theta}^{I,NI}, \pi_{P,\theta}^{I,I}\}$.

Firms acquire information at a price that depends on the extent to which information increases their profits. This value of information varies according to whether the competitor purchases the information. The data broker extracts all surplus from competing firms and maximizes the difference between the profits of an informed firm and those of an uninformed firm. The data broker profit function can be written as:

$$\Pi = \begin{cases} \Pi_1 = \max_{P\in\mathbb{P}}\{\pi_{P,\theta}^{I,NI} - \pi_{P,\theta}^{NI,NI}\} \\ \text{if the DB sells information to only one firm} \\ \Pi_2 = \max_{P\in\mathbb{P}}\{\pi_{P,\theta}^{I,I} - \pi_{P,\theta}^{NI,I}\} \\ \text{if the DB sells information to both firms} \end{cases}$$

Firms simultaneously set their prices on the unit line when they have no information or on each segment of the partition when they are informed. For any partition P composed of n segments, firm θ maximizes its profits with respect to the prices on each segment, denoted by the vector $p_\theta = (p_{\theta 1}, \dots, p_{\theta n}) \in \mathbb{R}^n_+$:

$$\pi_{P,\theta} = \sum_{i=1}^{n} d_{\theta i}(p_\theta, p_{-\theta})p_{\theta i}$$

In this framework (Bounie et al. [2021]), two important results on DB's strategy arise:

- When information is sold exclusively to one firm, the DB sells a partition P with all segments up to a point $\frac{j}{k}$ and leaves a large segment of unidentified consumers after that point. The DB adopts this strategy to soften the competitive pressure that occurs against the informed firm when its uninformed competitor lowers the basic price on the entire unit line;
- When information is sold to both firms, the data broker sells the same information structure to both firms, that is, $\frac{j}{k} = \frac{j'}{k}$. The remaining consumers are unidentified: segments that would allow firms to poach consumers are not sold.

When firms acquire information through a second-price auction with negative externalities as in Montes et al. [2019], the DB finds it convenient to offer an exclusive sale. Furthermore, in cases where the DB doesn't exploit the negative externalities, and instead keeps both companies uninformed after one firm declines its proposal, it can be demonstrated that the DB's profits are greater in the scenario without exclusive sales. Hence, the assumption regarding the sales mechanism plays a pivotal role in comprehending the process of information acquisition

Bounie et al. [2022] use the same hotelling framework to study the impact of different selling mechanisms: list prices, sequential bargaining, and auctions. Each one of these offers different outside options for potential buyers. A firm that chooses not to acquire information may encounter a negative externality when competing with an informed competitor. The informational structure sold to one or both firms is the same as in Bounie et al. [2021] and it's outlined in the previous two bullet points.

Assume that the DB wants to sell data to only one firm.

With list prices, the data intermediary proposes an information partition to Firm 1 (P_1^{lp}) that accepts or rejects the offer. If Firm 1 declines the offer, both firms remain uninformed obtaining profit π.

The second mechanism, sequential bargaining, allows the data intermediary to propose information to Firm 2 if Firm 1 declines the offer, and so on until one of the firms acquires information. The DB will offer to Firm 1 a partition P_1^{sb} that maximizes $\pi_1(P_1^{sb}, k)$ and a partition $P_2^{sb} = P_1^{sb}$ to Firm 2 creating a credible threat on Firm 1.

The third selling mechanism is an auction with a negative externality: a firm that loses the auction faces an informed competitor, similar to sequential bargaining. In order to maximize the price of information, the data intermediary designs two simultaneous auctions with a reserve price, and only the partition with the highest bid will be sold. If the DB wants to sell information only to Firm 1, it will sell the partition tailored to Firm

1 (P_1^a) in the first auction and it will exert the maximal level of threat for a firm that does not purchase information offering the reference partition P_k (with all segments included) in the second auction.

The price for information will be:

$$p = \begin{cases} p_{lp}(P_1^{lp}) = \pi_1(P_1^{lp}, k) - \pi \\ \text{with List Prices} \\ p_{sb}(P_1^{sb}) = \pi_1(P_1^{sb}, k) - \bar{\pi}_1(P_2^{sb}, k) \\ \text{with Sequential Bargaining} \\ p_a(P_1^a) = \pi_1(P_1^a, k) - \bar{\pi}_1(P_k, k) \\ \text{with Auctions} \end{cases}$$

Since, for these three selling mechanisms, the outside option partition offered to Firm 2 does not depend on the inside option partition offered to Firm 1, the DB will sell $P_1^{lp} = P_1^{sb} = P_1^a$. As a consequence, the impact on DB profit and consumer surplus will be driven by the partition offered in outside option. In the case of auctions, the data intermediary can maximize the value of the outside option threat and maximize Firm 1's willingness to pay. Conversely, with list prices, when a firm rejects the data intermediary's offer, both firms remain uninformed, leading to a lower willingness to pay for information.

When the DB chooses to sell information to both firms, it can be shown that the three selling mechanisms are equivalent as the outside option for each firm is the same.

When utilizing auctions to sell information to a single firm, the data intermediary can capitalize on the negative externality associated with the risk of remaining uninformed. This increases the prospective buyer's willingness to pay. Conversely, in sequential bargaining, the partition of information sold to Firm 2, in the event of Firm 1 declining the offer, is chosen to maximize Firm 2's profit rather than exerting the maximum negative externality on Firm 1. In the case of list prices, the data intermediary lacks the ability to threaten Firm 1 if it chooses not to purchase information. Consequently, the data intermediary prefers to employ sequential bargaining and list prices to sell information to both firms, while utilizing auctions to sell information to only one firm (Bounie et al. [2022]).

Following the analysis on data selling strategies with spatial price discrimination, Abrardi et al. [2022] investigates how the presence of DB affects firms' entry and competition in a Salop setting. The DB sells consumer data to the downstream market through a mechanism of simultaneous auctions with reserve prices and negative externalities. They analyze two DB's potential optimal strategies depending on whether all entering firms or an alternating sequence of them hold information. Under the sale to alternating firms, a portion of them are given a partitioned dataset, while the remaining firms receive it entirely but their auctions will not be fulfilled. Offering the complete dataset to a firm's direct competitors acts as a deterrent and increases their willingness to pay for the data, as they want to avoid competing against well-informed rivals.

The DB's optimal strategy involves reducing competition through two methods. Firstly, the DB limits market access by selling data to alternating firms. This creates local data monopolies where well-informed firms compete against uninformed competitors, thereby maximizing the value that firms are willing to pay for data. Secondly, the DB decreases the number of firms entering the market compared to a scenario where data is not available. This creates a barrier to entry as the cost of acquiring data reduces firms' profits and limits their capacity to enter the market.

However, differently from the auction mechanism, under the Take It Or Leave It offers the DB prefers to sell data to all entering firms and leaves space for a higher consumer surplus (consistently with the findings in Bounie et al. [2022]).

Chapter 2

Price discrimination on Spokes

The core contribution of this work lies in the extension of spatial price discrimination frameworks with data sales to a model of non-localized competition with differentiated products. All the spatial competition settings revised in the previous chapter pertain to the class of localized competition. Hotelling [1929] and Salop [1979] models are standard tools of this type of oligopoly analysis: in particular, in the Salop model a small change in a firm's price only affects its two neighbors, not the rest of the firms. Under non-localized competition (Chamberlin [1933]), each firm competes against the market and a price change by one firm affects all other firms (more or less) equally. Moreover, our model allows introducing an elastic monopolistic demand which expands or shrinks the total market output and can be useful to analyze the effect of data on market coverage.

2.1 The Spokes Model: baseline framework

Although there have been advancements in the economics literature concerning product differentiation, the analysis of oligopoly competition with product differentiation in a spatial model with nonlocalized competition had been lacking until the introduction of the Spokes model by Chen and Riordan [2007] and Caminal and Claici [2007].

Geographically, the market is represented by a line of unit length. From the midpoint of this line, additional lines of half the length, called spokes, radiate outward in a radial network. Each spoke (denoted as l_i) terminates at the center and originates at the other end. There are N distinct possible varieties of a product, labeled as $i = 1,2,\ldots,N$. Each variety is located at the origin of a specific spoke l_i. There are n (with $n \leq N$) firms in the market. Each firm produces a single variety or brand. These brands are physically identical but differentiated by their unique locations.

Consumers are uniformly distributed along the network of spokes. Each consumer is located on a particular spoke l_i. A consumer travels to a firm to purchase the firm's brand and incurs transportation costs (or utility losses due to imperfect preference matching). For a consumer located on l_i, brand i is her first preferred brand (or local brand), and each of the other $N-1$ brands is equally likely to be her second preferred brand. The consumer has value v for one unit of either her first or second preferred brands, and zero value for additional units or for other brands.

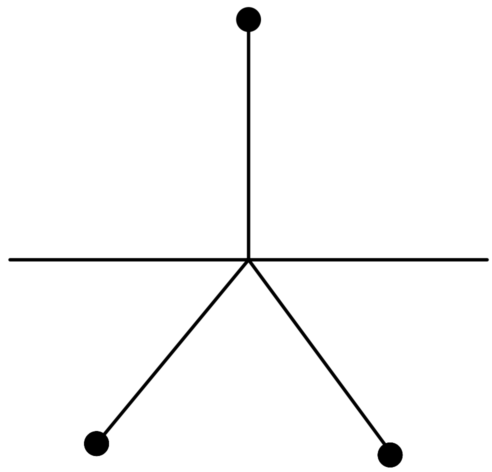


Figure 2.1. Spokes Model for $N = 5, n = 3$

Comparing this framework to the Salop model, three main differences arise:

- The Spokes model maintains symmetry among all brands and firms. This means that there is no need to adjust the locations of existing firms when new firms enter the market. Each brand and firm is treated equally within the model;
- Despite each consumer having a fixed number of preferred varieties, each firm competes directly with all other firms in the market. This means that firms must consider the actions and strategies of all competitors, even if individual consumers are only interested in a specific subset of product varieties;
- Unlike the Salop model, the total output in the Spokes model is not predetermined but depends on equilibrium prices and the number of firms present in the market. This means that the entry of new firms into the market can have a market expansion effect, leading to changes in output and overall market dynamics.

In the Spokes model, each consumer location is characterized by a vector (l_i, x_i) meaning that the consumer is on l_i at a distance x_i to variety i (the origin of l_i). Considering transportation cost normalized to 1, the distance between a consumer (x_i, l_i) and any of the variety i' is spoke independent and is equal to $\frac{1}{2} - x_i + \frac{1}{2} = 1 - x_i$.

In Chen and Riordan [2007], variety i is consumer (l_i, x_i)'s first preferred brand (or local brand), of which her value for one unit is v; she also has a second preferred brand, which is any $i' \neq i$ chosen by nature with probability $\frac{1}{N-1}$, and of which her value for one unit is also v. Supposing alternatively that consumer (l_i, x_i) valued equally all varieties other than variety i, the existence of a symmetric pure strategy Nash equilibrium would not be guaranteed.

For any consumer located on l_j or on l_k, denoted as (l_j, x_j) or (l_k, x_k), for $j,\ k \in \{1, \dots, n\}$, both variety j and variety k are her desired brands with conditional probability $\frac{1}{N-1}$. Such a consumer is indifferent between variety j and k if $p_j + x_j = p_k + (1 - x_j)$ or $p_j + (1 - x_k) = p_k + x_k$. The marginal consumer between j and k is a distance:

$$\hat{x} = \max\left[\min\left(\frac{1}{2} + \frac{p_k - p_j}{2}, 1\right), 0\right] \tag{2.1}$$

from firm j. The number of such consumers served by firm j is:

$$\frac{2}{N}\frac{1}{N-1}\sum_{k\neq j, k\in 1,\dots,n} \max\left[\min\left(\frac{1}{2} + \frac{p_k - p_j}{2}, 1\right), 0\right] \tag{2.2}$$

where $\frac{2}{N}$ is the density of consumer on l_j and on l_k.

For any consumer on l_j with probability $\frac{1}{N-1}$ variety i is her second preferred brand where $i \notin \{1, \dots, n\}$. Such a consumer prefers purchasing from firm j to not purchase if $p_j + x_j \leq v$. Firm j's demand from this second category of the consumer is:

$$\frac{N-n}{N-1}\frac{2}{N}\min\left[\max(0, v - p_j), \frac{1}{2}\right] \tag{2.3}$$

where $\frac{2}{N}$ is the density of consumer on l_j, and $N - n$ varieties are unavailable.

Finally, for any consumer on l_i, $i \neq j$ and $i \notin \{1, \dots, n\}$, variety j is her second preferred brand with probability $\frac{1}{N-1}$. Such a consumer prefers purchasing from firm j to not purchasing if $p_j + (1 - x_i) \leq v$. Firm j's demand from this last consumer type is

$$\frac{N-n}{N-1}\frac{2}{N}\min\left[\max\left(0, v - p_j - \frac{1}{2}\right), \frac{1}{2}\right] \tag{2.4}$$

Summing up these three categories of consumers, and simplifying, we obtain the firm j's total demand as:

$$q_j = \frac{1}{N-1}\frac{2}{N}\sum_{k\neq j, k\in\{1,\dots,n\}} \max\left[\min\left(\frac{1}{2} + \frac{p_k - p_j}{2}, 1\right), 0\right] + \frac{N-n}{N-1}\frac{2}{N}\max[\min(v - p_j, 1), 0] \tag{2.5}$$

which, provided $|p_k - p_j| \leq 1$, can be rewritten as

$$q_j = \begin{cases} \dfrac{1}{N-1}\dfrac{2}{N}\displaystyle\sum_{k\neq j, k\in\{1,\dots,n\}}\left(\frac{1}{2} + \frac{p_k - p_j}{2}\right) + \dfrac{N-n}{N-1}\dfrac{2}{N}(v - p_j) & \text{if} \quad 0 \leq v - p_j \leq 1 \\ \dfrac{1}{N-1}\dfrac{2}{N}\displaystyle\sum_{k\neq j, k\in\{1,\dots,n\}}\left(\frac{1}{2} + \frac{p_k - p_j}{2}\right) + \dfrac{N-n}{N-1}\dfrac{2}{N} & \text{if} \quad v - p_j > 1 \end{cases} \tag{2.6}$$

Thus firm j essentially faces two groups of consumers: consumers who have an alternative available and consumers who do not, corresponding to the two terms of the demand function in (2.5). For $v - p_j < 1$, some consumers who have the alternative j available as their first or second choice do not purchase and the marginal purchasing consumer in this group will have zero surplus. For $v - p_j > 1$, the market is fully covered, and consumers who have at least an alternative available purchase.

Given the symmetry of the problem, Chen and Riordan [2007] focus on symmetric pure strategy Nash equilibria in which all firms set the same price p^*, serve an equal quantity of consumers q^*, and earn the same profit $\pi^* = q^* p^*$. Equilibrium symmetric will have an intrinsic dependence on product valuation and it can be shown that the equilibrium exists only if:

$$v \leq 2\frac{N-1}{n-1} + \frac{1}{2}\frac{2N-n-1}{N-n} = \bar{v}(n, N)$$

The regions are distinguished by the prevailing pattern of consumer demand, in particular, the extent to which consumers whose desired brands are available to make a purchase and obtain a positive surplus in equilibrium.

Remark 1. *(Chen & Riordan,2007) For a given n, the spokes model has a unique symmetric equilibrium. The equilibrium price is:*

$$p^* = \begin{cases} \dfrac{2N-n-1}{n-1} & \text{if} \quad 2\frac{N-1}{n-1} < v < \bar{v}(n, N) \text{ (Region I)} \\ v-1 & \text{if} \quad 2 \leq v \leq 2\frac{N-1}{n-1} \text{ (Region II)} \\ \dfrac{2(N-n)v+(n-1)}{4N-3n-1} & \text{if} \quad \frac{1}{2} + \frac{N-1}{2N-n-1} < v < 2 \text{ (Region III)} \\ v - \dfrac{1}{2} & \text{if} \quad 1 \leq v \leq \frac{1}{2} + \frac{N-1}{2N-n-1} \text{ (Region IV)} \end{cases}$$

From a technical point of view, it is also worth remarking that the prices in regions I and III are interior solutions, obtained from solving the first-order conditions; regions II and IV are corner solutions, corresponding to kinks in the demand curve.

The rationale behind the results can be explained as follows: in the context of typical oligopoly competition, which occurs in Region I, price competition among firms benefits all consumers who have access to both of their preferred brands. Unlike in other regions, the equilibrium price is determined by the number of active firms and the total number of possible brands, rather than the relatively high value placed on the product. In contrast, Region II involves firms taking advantage of captive consumers who have only one available brand in the market. The price is set in a way that the marginal consumer is indifferent between buying the least preferred brand or not purchasing anything at all. In Region III, firms cater to both consumers who have the freedom to choose and those who are captive to a single brand. The marginal consumer in the competitive segment is indifferent between the two available brands, whereas the marginal captive consumer is indifferent between buying the second preferred option or abstaining from the market altogether. The equilibrium price in this region is influenced by three parameters: v,n, and N. In Region IV, there is a distinct kink in the demand curve where only consumers whose first preferred brand is available to make purchases, and the marginal consumer is indifferent between buying or not buying the product. In this case, as well, the equilibrium price is solely dependent on the valuation.

The effects of market structure on equilibrium prices follow easily from Remark 1:

Remark 2. *(Chen & Riordan,2007) For prices and boundaries defined in Remark 1, the following statements hold:*

$$\frac{dp^*}{dn} = \begin{cases} -2\dfrac{N-1}{(n-1)^2} < 0 & \textit{(Region I)} \\ 0 & \textit{(Region II)} \\ \dfrac{2(N-1)(2-v)}{(3n-4N+1)^2} > 0 & \textit{(Region III)} \\ 0 & \textit{(Region IV)} \end{cases}$$

Generally, an increase in n has a relevant impact on consumers. This impact can be observed through an improvement in consumer satisfaction (utility) or a decrease in transportation costs. Previously, certain consumers were limited to purchasing a less preferred brand because their desired brand was not available. However, with the increased number of options, these consumers can now choose and consume their preferred brand, resulting in increased utility. Additionally, the availability of a greater variety of options allows consumers who couldn't access their desired brands before to now experience a positive surplus.

The above lemma shows that the impact of changes in market concentration on price varies across different regions of the parameter space: it exhibits a weak decreasing effect in n for $v > 2$, but a weak increasing effect in n for $v < 2$. In the kinked demand equilibria of regions II and IV, changes in the number of firms do not affect the equilibrium price. Dargaud and Reggiani [2015] link this characteristic of the equilibria to empirical evidence regarding the price impacts of horizontal mergers. Indeed, existing studies indicate that the anticipated negative effects on prices and consumer surplus, which are typically examined by antitrust authorities, do not always materialize. In some cases, significant consolidations may have minimal effects on prices. This finding suggests that mergers within the context of non-localized spatial competition can result in zero price effects when firms specifically target certain kinks in the demand function (Reggiani [2020]). On the other hand, in Region I, the standard comparative statics applies: an increase in the number of competitors in the market leads to a decrease in prices.

However, a more intriguing outcome is observed in Region III: an increase in the number of competitors causes firms to raise their prices. Unlike other oligopoly models where competition leads to price reductions, Chen and Riordan [2007] achieves this result under the assumption of complete information and pure strategies. This region is characterized by a demand that is more elastic for the captive segment compared to the competitive one. This occurs because as the firm decreases its price, the marginal consumer in the monopoly segment always obtains zero surplus from choosing an alternative option (with infinite elasticity), while the marginal consumer in the competitive segment becomes increasingly enticed by competing brands (with finite elasticity). As the number of firms increases, the captive segment contracts while the competitive segment expands, reducing the overall average demand elasticity and ultimately resulting in a higher equilibrium price.

Expanding on this framework, the study conducted by Reggiani [2014] explores the

Spokes Model where firms strategically determine their locations to implement price discrimination based on location-specific pricing. Unlike the assumptions of Chen and Riordan [2007], consumers in this study consider all available alternative suppliers in their decision-making process. As a result, consumers located in areas where there are no nearby suppliers are not bound to any specific firm. The firms in this model bear the transportation costs and deliver the products directly to the consumers' addresses. Furthermore, since firms make personalized offers to individual consumers, they essentially engage in Bertrand competition, where each firm has different costs associated with serving various locations. Consistently with established findings in the literature, the equilibrium location pattern in this model aligns with minimizing social costs. In scenarios with a sufficiently high number of firms, a highly asymmetric location pattern may emerge (Reggiani [2014]). In such cases, one firm supplies a product that generally appeals to a wide range of consumers, while other firms concentrate on serving specific niche markets.

In contrast to the approach taken by Reggiani [2014], our setting introduces spatial price discrimination performed by firms situated at the origin of each spoke and does not include a location choice stage. We also retain the assumption of captive consumers, as it allows us to analyze the impact of data usage for discrimination purposes on demand elasticity. In essence, our research focuses on the interplay between oligopolistic and monopolistic demand when considering the competitive advantage derived from the sale of data. Moreover, we assume that, as in all standard spatial models transportation costs are a measure of consumers' disutility, and location is a product characteristic.

In the context of digital economics, the Spokes Model is used by Rhodes [2011] to study the widespread use of search-related advertising in online markets, where sponsored links are prominently displayed on webpages and consumers frequently click on them. Firms actively compete for these desirable positions at the top of the page. This raises the question of why such prominence holds value in these contexts, considering that visiting additional websites is nearly costless. In the framework presented, consumers know their product valuations but are unaware of which retailer sells each specific product. The Spokes model incorporates the search results presented by gatekeepers like Google or Bing, either in a random order or with a specific firm given prominence. The primary contribution of the study is to demonstrate that a retailer in a prominent position earns significantly higher profits compared to other firms, even when the cost of searching websites and comparing products is essentially negligible (Rhodes [2011]).

Loginova [2022], examines the dynamics of price competition between online retailers, where some operate their own branded websites while others utilize online platforms like Amazon Marketplace. The study adopts the Spokes model due to its flexibility in accommodating both types of firms, given its non-localized nature. The firms face a trade-off in their sales strategy. Selling through Amazon enables a firm to access a larger customer base since consumers often rely on Amazon to discover alternatives, reducing search and comparison costs. On the other hand, establishing an independent website can enhance the perceived value of the firm's product and develop a brand reputation. In the long run, each firm must decide between utilizing Amazon or maintaining its own website, while in the short run, this choice cannot be altered. Analyzing the comparative statics of the resulting equilibria yields interesting insights. For instance, an increase in competition may lead to a decrease in the number of firms choosing Amazon as their sales channel

(Loginova [2022]). Additionally, a pure-strategy Nash equilibrium does not always exist, resulting in price dispersion. Firms are more likely to employ mixed strategies in less concentrated markets and when the increase in perceived product value is relatively small.

2.2 Model Setup

We consider a market in which horizontally differentiated firms sell a product to a mass of consumers, whose preferences can be observed by a firm only if it purchases customers' specific data from a Data Broker (DB). For example, firms sell their products via e-commerce solutions and the possibility of identifying the consumer through data acquired from a DB allows the firm to make personalized offers.

The competition framework is the Spokes Model introduced by Chen and Riordan [2007] and described in Section 2.1. There are N spokes (possible varieties), denoted by l_j, where $j \in \{1, \ldots, N\}$, and n firms. The firms are located at the origin of each spoke. Specifically, we focus the analysis on $N > 3$ and $n \in [2, N-1]$ for the ease of exposition, without restricting findings on market outcomes.

All the assumptions about consumer preferences, described in Section 2.1, are valid in our setting. According to these assumptions, for a consumer located on l_i, brand i is her first preferred brand (or local brand), and each of the other $N-1$ brands is equally likely to be her second preferred brand. Each firm encounters two groups of consumers, and its demand function consists of both a monopolistic and a competitive component.

Moreover, differently form Loginova [2022], we assume that utility values v fall in the range $[\frac{3}{2}$,2], which is included in Region III of Remark 1 and implies that, in the no-information scenario ($d = 0$), all firms carve out a portion of consumers in monopolistic markets (Figure 2.2).

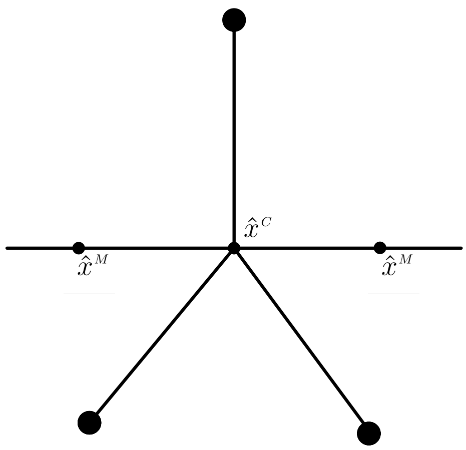


Figure 2.2. Spokes Model for $N = 5, n = 3, v \in [\frac{3}{2}$,2]

There is a Data Broker who possesses a dataset containing the geographical information of all consumers in the market. This Data Broker can sell a portion of this dataset to each firm that enters the market. To enhance the value of the data for the firms and increase their willingness to pay, the partition sold to each firm includes its own location, following the approach in Bounie et al. [2021]. Specifically, because the market exhibits symmetry, the partition sold starts from each firm's own location at the origin of each spoke.

We assume that the DB sells the same quantity of data (denoted by $d \in$ [0,1]) to all firms. Consumers on spoke l_j, where $j \in \{1, \dots, n\}$, belonging to the segment $d \in [0, \frac{1}{2})$, are identified by firm j: the firm knows their location and can perform on them first-degree price discrimination. If $d \in [\frac{1}{2}$,1] each firm j identifies consumers up to d on all spokes l_i, with $i \neq j$ and $i \in \{1, \dots, N\}$.

For any consumer, belonging to the segment d, who has both preferred variety i and j available, firm j and firm i provide a personalized price $p^C(x)$, where $x \in [0, d]$.

Consumers with one of the two preferences for an empty spoke and the other for variety j, where $j \in \{1, \dots, n\}$, receive a targeted personalized offer $p^M(x)$, where $x \in [0, d]$, specifically designed for the monopolistic market.

Unidentified consumers, with at least one choice available, will receive offers with basic price p_B.

The demand function described in Section 2.1 is redefined in the following Section taking into account the presence of data. Each informed firm offers its basic price exclusively to consumers about whom it possesses no information.

2.2.1 Selling mechanism and timing

We assume that the DB wants to sell a quantity d of data to all firms through a Take It Or Leave It (TIOLI) mechanism. The concept of "take it or leave it" in microeconomics and game theory refers to a situation where one player offers a specific deal or proposal to another player, and the receiving player can either accept the offer as it is or reject it outright. This type of interaction is often analyzed to understand strategic decision-making and negotiation dynamics in various economic settings. The "take it or leave it" approach assumes that the offering player (DB) has the power to set the terms unilaterally, and the other player's decision (firms) is limited to accepting or rejecting the offer without any further negotiation.

In this scenario, the DB offers an equal segment to each firm, and each firm individually and simultaneously accepts or declines the offer. In particular, the DB chooses the quantity of data d and offers to each firm i, where $i \in \{1, \dots, n\}$, a segment starting from the origin of spoke l_i. The DB determines the length d of the segment by maximizing the willingness of each firm to pay for the data: the DB extracts all surplus from competing firms and maximizes the differences between the profits of an informed firm and those of an uninformed firm as in Montes et al. [2019], Bounie et al. [2021] and Abrardi et al. [2022].

Letting $\pi^W(d)$ denote a firm's profit when it accepts the offer and all its competitors are informed (Inside Option), $\pi^L(d)$ denote its profits when it rejects the offer and faces all other firms with data (Outside Option), the price of each segment is:

$$w = \pi^W(d) - \pi^L(d)$$

All firms receive an equal quantity of data, resulting in a symmetric problem and an equal price for data across all firms. The length of the segment and the selling mechanism are common knowledge.

The DB's profits can be calculated by adding up the acceptance of offers from firms, which are equivalent to their respective willingness to pay for data:

$$\pi_{DB}(d) = n\left(\pi^W(d) - \pi^L(d)\right) \tag{2.7}$$

It is worth noting that if the DB aims to sell data to all firms, any selling mechanism that ensures the independence of the partition offered as an Inside Option to winning firms, as compared to the partition offered to competitors as an Outside Option, will produce identical outcomes (Bounie et al. [2022]). Auctions with negative externalities and reserve prices, as well as sequential bargaining, are examples of selling mechanisms within this class. When the DB chooses not to exclude any firms from data acquisition, it finds no advantage in using data as an exclusive tool to enhance the threat of the Outside Option. In fact, all firms face the same decision under any selling mechanism: if a firm accepts the DB's offer, all firms in the market become informed, whereas if a firm declines the offer, it becomes the only uninformed firm.

The timing of the model is as follows:

- Stage 1: The DB chooses the unique length d of each segment and the price of data w. Offers are non-renegotiable;
- Stage 2: Each firm individually and simultaneously accepts or rejects the offer and pay for data;
- Stage 3: Firms set basic prices p^B for the anonymous consumers;
- Stage 4: Firms set tailored prices $p^{\{C,M\}}(x)$ for the identified consumers if they have accepted the DB's offer. Consumers purchase the product and profits are made.

Stage 4 follows Stage 3 to ensure the existence of an equilibrium in pure strategies and is supported by managerial practices (Fudenberg and Villas-Boas [2007]).

2.2.2 All Firms informed

All firms hold a quantity $d \in [0,1]$ of information. In the range of $d \in [\frac{1}{2},1]$, personalized prices are implemented in all duopolistic markets, while the basic price is applied to captive consumers and influences competitors' tailored offer (Figure 2.3).

For any consumer located on l_j, with both preferred varieties j and $k \neq j$ are available (i.e. $j, k \in \{1, \ldots, n\}$), the distance of marginal consumer from firm j boils down to the equation (2.1) of the baseline framework. The competitive set of unidentified consumers, described in equation (2.2), can be redefined as:

$$q_j^{W,C} = \begin{cases} \dfrac{2}{N}\dfrac{1}{N-1}\displaystyle\sum_{k \neq j, k \in \{1,\ldots,n\}} \max\left[\min\left(\dfrac{1}{2} + \dfrac{p^W_{B,k} - p^W_{B,j}}{2} - d, 1-d\right), 0\right] & \text{if } d < \frac{1}{2} \\ 0 & \text{if } d \geq \frac{1}{2} \end{cases} \tag{2.8}$$

For the competitive fringe of consumers on the spoke l_j, where $x_j \in [0, d]$ and partitions are not overlapped, firm j matches the competitor basic price in utility level. In the case of overlapping partitions the personalized prices are driven down where both firms have data and Bertrand competition takes place:

$$\begin{cases} p_j^C(x) = p_{B,i}^W + 1 - 2x \quad \forall\, i \neq j, x \in [0, d] \quad \text{if } d < \dfrac{1}{2} \end{cases}$$

$$\begin{cases} p_j^C(x) = p_{B,i}^W + 1 - 2x & \forall\, i \neq j, x \in [0{,}1 - d] \\ p_j^C(x) = \min(1 - 2x{,}0) & x \in [1 - d{,}1] \end{cases} \text{if } d \geq \frac{1}{2} \tag{2.9}$$

For any consumer located on l_j whose second variety i with probability $\frac{1}{N-1}$ is not available (i.e. $i \notin \{1, \ldots, n\}$), if the data partition does not cover entirely l_j, the basic price set by firm j holds until $p_{B,j}^W + x_j \leq v$. Firm j's non-discriminatory demand function for this component of captive consumers described in (2.3), becomes:

$$\begin{cases} \frac{N-n}{N-1}\frac{2}{N} \min\left[\max\left(0, v - p_{B,j}^W - d\right), \frac{1}{2} - d\right] & \text{if } d < \frac{1}{2} \\ 0 & \text{if } d \geq \frac{1}{2} \end{cases} \tag{2.10}$$

Finally, consumers on l_i, with variety i not available and second choice j located on filled spokes (i.e. $j \in \{1, \ldots, n\}$), prefer purchasing from firm j to not purchasing if $p_{B,j}^W + (1 - x_i) \leq v$. If data partition goes beyond spoke l_j, personalized price is offered to a subset of these captive consumers and firm j's non-discriminatory demand, defined in (2.4), becomes:

$$\begin{cases} \dfrac{N-n}{N-1}\dfrac{2}{N} \min\left[\max\left(0, v - p_{B,j}^W - \dfrac{1}{2}\right), \dfrac{1}{2}\right] & \text{if } d < \frac{1}{2} \\ \dfrac{N-n}{N-1}\dfrac{2}{N} \min\left[\max\left(0, v - p_{B,j}^W - d\right), 1 - d\right] & \text{if } d \geq \frac{1}{2} \end{cases} \tag{2.11}$$

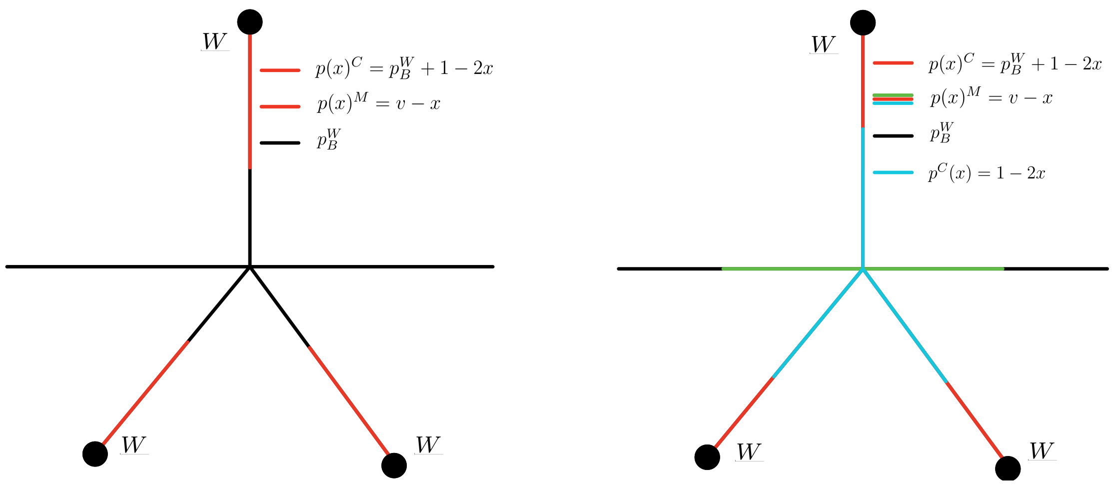
The monopolistic component of firm j's demand function can be rewritten by summing up (2.10) and (2.11):

$$q_j^{W,M} = \begin{cases} \dfrac{N-n}{N-1}\dfrac{2}{N}(v - p_{B,j}^W - d) & \text{if} \quad d \leq v - p_{B,j}^W \leq 1 \\ \dfrac{N-n}{N-1}\dfrac{2}{N}(1 - d) & \text{if} \quad v - p_{B,j}^W > 1 \end{cases} \tag{2.12}$$

As in the baseline framework (2.6), for $v - p_{B,j}^W < 1$ some non-discriminatory consumers who have firm j as their first or second choice and lack an available alternative prefer not buying to accept firm j's basic price. Moreover, when $v - p_{B,j}^W < d$, firm j will serve only the discriminatory component of monopolistic market.

For all the consumers on spoke l_i or l_j, with a first or second preference for variety j but only one alternative available (i.e. $i \notin \{1, \ldots, n\}$), firm j's personalized offers extract all the available surplus in the following way:

$$p_j^W(x) = v - x \quad x \in [0, d], \; \forall d \in [0{,}1] \tag{2.13}$$



(a) $N = 5,\ n = 3,\ d \in [0, \frac{1}{2})$

(b) $N = 5,\ n = 3,\ d \in [\frac{1}{2}, 1)$

Figure 2.3. Spatial pricing strategy when all firms (W) hold d data

By summing up (2.8) and (2.12), assuming that $|p^W_{B,k} - p^W_{B,j}| \leq 1$ and $p^W_{B,j} \leq v - d$, we can derive the set of consumers served by firm j's with its basic price:

$$q_j^W = \begin{cases} \dfrac{2}{N}\dfrac{1}{N-1} \displaystyle\sum_{k \neq j, k \in \{1,\ldots,n\}} \left(\dfrac{1}{2} + \dfrac{p^W_{B,k} - p^W_{B,j}}{2} - d \right) + q_j^{W,M} & \text{if } d < \frac{1}{2} \\ q_j^{W,M} & \text{if } d \geq \frac{1}{2} \end{cases} \tag{2.14}$$

In the scenario where there are overlapping partitions, each firm j serves only monopolistic markets at its basic price and captures its own spoke in duopolistic markets. Therefore, the profit function of firm j can be defined as a consequence, depending on d.

- If $d \in [0, \frac{1}{2})$:

$$\pi_j^W = q_j^{W,M} p^W_{B,j} + q_j^{W,C} p^W_{B,j} + \frac{n-1}{N-1}\frac{2}{N}\int_0^d p_j^C(x)\,dx + \frac{N-n}{N-1}\frac{2}{N}\int_0^d p_j^M(x)\,dx \tag{2.15}$$

- If $d \in [\frac{1}{2}, 1]$:

$$\pi_j^W = q_j^{W,M} p^W_{B,j} + \frac{n-1}{N-1}\frac{2}{N}\int_0^{\frac{1}{2}} p_j^C(x)\,dx + \frac{N-n}{N-1}\frac{2}{N}\int_0^d p_j^M(x)\,dx \tag{2.16}$$

Given the symmetry of the model, we focus on symmetric Bertrand-Nash (pure strategy) equilibria:

- Inside Option: When a generic firm i from the set of firms $\{1,\ldots,n\}$ accepts the TIOLI offer from the DB, it will faces $n-1$ competitors who have the same amount of information d. Each firm will set the base price p_B^{W*} at an equal level, and all firms will serve an equal number of non-discriminatory customers q^{W*}. The demand function, which is applicable where the base prices hold and is defined in equation (2.14), becomes:

$$q^{W*} = \begin{cases} \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2}-d\right)+\frac{N-n}{N-1}\frac{2}{N}\min(v-p_B^{W*}-d,1-d) & \text{if } d<\frac{1}{2} \\ \frac{N-n}{N-1}\frac{2}{N}\min(v-p_B^{W*}-d,1-d) & \text{if } d\geq\frac{1}{2} \end{cases} \tag{2.17}$$

2.2.3 One firm uninformed

When only one firm s is uninformed and all its $n-1$ competitors hold a quantity d of data, each informed firm j (Figure 2.4), with $j\neq s$, will face the following demand function:

$$q_j^W = \begin{cases} \frac{2}{N}\frac{1}{N-1}\left(\sum_{k\neq\{j,s\}}\left(\frac{1}{2}+\frac{p_{B,k}^W-p_{B,j}^W}{2}\right)+\frac{1}{2}+\frac{p_{B,s}^L-p_{B,j}^W}{2}-nd\right)+q_j^{W,M} & \text{if } d<\frac{1}{2} \\ \frac{1}{N-1}\frac{2}{N}\left(\frac{1}{2}+\frac{p_{B,s}^L-p_{B,j}^W}{2}-d\right)+q_j^{W,M} & \text{if } d\geq\frac{1}{2} \end{cases} \tag{2.18}$$

The first component of the second equation in (2.18) represents the duopolistic demand which arises from competition with the only uninformed firm s.

Moreover, the tailored prices specifically designed by firm j for consumers with two preferred varieties available and one of them corresponding to variety s, will be defined as follows:

$$p_j^C(x) = p_{B,s}^L + 1 - 2x \quad x\in[0,d] \tag{2.19}$$

Firm s's demand function boils down to equation (2.6) of the basline framework:

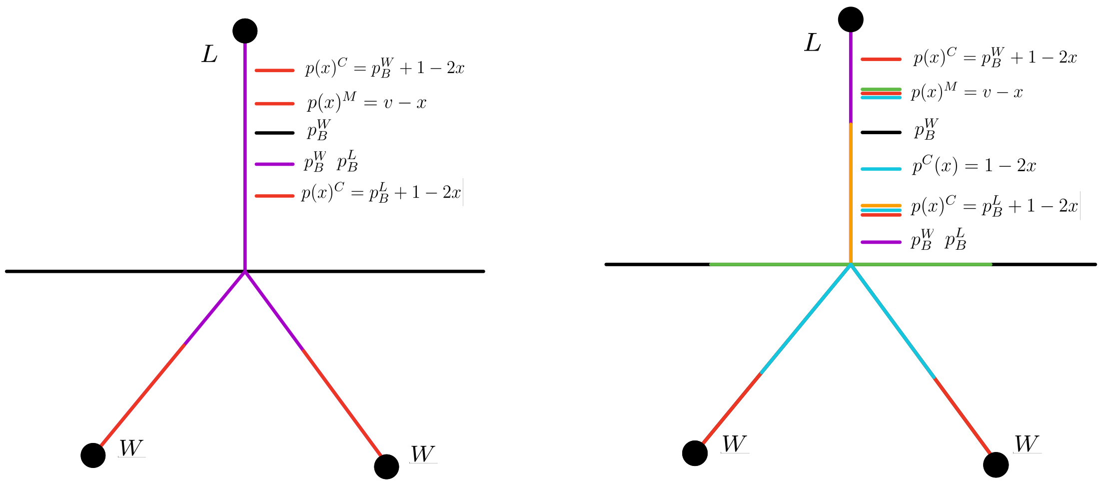
$$q_s^L = \frac{2}{N}\frac{1}{N-1}\sum_{k\neq s,k\in\{1,\ldots,n\}}\left(\frac{1}{2}+\frac{p_{B,k}^W-p_{B,s}^L}{2}\right)+\frac{N-n}{N-1}\frac{2}{N}\min(v-p_{B,s}^L,1) \tag{2.20}$$

Firm s's profit corresponds to $\pi_s^L = p_{B,s}^L q_s^L$.

Given the symmetry of the model, we focus on symmetric Bertrand-Nash (pure strategy) equilibria:

- Outside Option: When a generic firm i from the set of firms $\{1,\ldots,n\}$ rejects the TIOLI offer from the DB, it will face $n-1$ competitors who have the same amount of information d. The uninformed firm i will establish a distinct base price p_B^{L*}, while all its competitors will set a common base price p_B^{WL*}. The demand function of the firm without data, defined in equation (2.20), becomes:

$$q^{L*} = \begin{cases} \dfrac{n-1}{N-1}\dfrac{2}{N}\left(\dfrac{1}{2}+\dfrac{p_B^{WL*}-p_B^{L*}}{2}\right)+\dfrac{N-n}{N-1}\dfrac{2}{N}(v-p_B^{L*}) & \text{if} \quad 0 \leq v-p_B^{L*} \leq 1 \\ \dfrac{n-1}{N-1}\dfrac{2}{N}\left(\dfrac{1}{2}+\dfrac{p_B^{WL*}-p_B^{L*}}{2}\right)+\dfrac{N-n}{N-1}\dfrac{2}{N} & \text{if} \quad v-p_B^{L*} > 1 \end{cases} \tag{2.21}$$



(a) $N = 5$, $n = 3$, $d \in [0, \frac{1}{2})$ (b) $N = 5$, $n = 3$, $d \in [\frac{1}{2}, 1)$

Figure 2.4. Spatial pricing strategy when only one firm (L) holds no information

2.3 Equilibrium Prices and DB's strategies

The analysis of equilibrium prices and DB's strategies will be categorized according to the values assumed by the number of firms n. When the number of firms n is sufficiently lower than the count of unoccupied spokes, represented as $N - n$, each individual firm will encounter a higher prevalence of consumers who are captive to their offerings. Conversely, when the number of firms becomes larger, the collective sum of consumers in duopolistic markets exceeds the aggregate count of captive consumers. We will designate the region where $n < n_2^I(v, N)$ as the domain characterized by a prevalence of monopolistic sub-markets, while the region where $n \geq n_2^I(v, N)$ will be referred to as the domain characterized by a prevalence of competitive sub-markets.

$$\begin{cases} 2 \leq n < n_2^I = \lceil \frac{N+v-1}{v} \rceil^+ & \text{Monopolistic Sub-Markets Prevalence (M.S.P.)} \\ \lceil \frac{N+v-1}{v} \rceil^+ = n_2^I \leq n \leq N & \text{Competitive Sub-Markets Prevalence (C.S.P.)} \end{cases}$$

The range of n values designed above emerges from the equilibrium analysis conducted in the scenario where the Data Broker sells non-overlapping segments and enables a comprehensive treatment of market outcomes. In the case of overlapping segments, we will explore the impact of a generic competitive or monopolistic sub-markets prevalence on prices and profits.

When $d = 0$, assuming that $v \in [\frac{3}{2},2]$, equilibrium prices and profits fall in the region III of Remark 1 (Chen and Riordan [2007]): all firms find convenient to carve out a portion of the monopolistic market.

As d gets higher, for all values of (n, v), the two effects on firm's profit described in Section 1.1 take place (Thisse and Vives [1988]):

- Surplus Extraction Effect (S.E.E.): when d increases, data allow firms to identify consumers and charge them with a tailored price, which exactly matches their willingness to pay for the product (the last two terms of (2.16) and (2.15)). In particular, firms extract all the surplus available from captive consumers, after deducting only transportation costs (2.13).

- Intensified Competition Effect (I.C.E.): when d increases, on each duopolistic market, the distance between the origin of each firm's spoke and the average location of non-discriminatory consumers also increases. This distance poses a greater challenge in attracting these consumers. Consequently, each firm lowers its base price in an attempt to lure these consumers away from competitors. This price reduction affects also the rent extraction from captive consumers and leads to a decline in the amount of profit extracted from all the non-discriminatory segments (the first two terms of (2.15)). Furthermore, for $d \in [\frac{1}{2},1]$, the base price only applies to captive consumers (first term of (2.16)), and the I.C.E. on firm's profit occurs only through Bertrand competition on overlapping duopolistic segments (2.9).

When the number of firms is sufficiently lower than the number of empty spokes $N - n$ (M.S.P.), for all values of the pair (d, v), basic prices will be greater or equal than $v - 1$ (absence of basic prices war). In this situation, the relevance of monopolistic demand outweighs that of competitive demand, resulting in the market being either incompletely covered or being fully covered with the marginal captive consumer surplus that approaches zero. At the same time, the rent extraction from discriminated captive consumers outweights the intensified competition effect that drives down non-discriminatory profits.

When the number of firms significantly exceeds the number of empty spokes $N - n$ (C.S.P.) and the quantity of information d approaches $\frac{1}{2}$ from the left, the group of unidentified captive consumers holds less significance compared to the consumer base in competitive sub-markets. Simultaneously, the area comprising non-discriminatory consumers in each duopolistic market that can be reached by basic prices diminishes and moves farther away from the origin of each spoke. As a result, the competitive pressure from rivals forces each firm to reduce its base price below $v-1$, leading to an intensive decline in rent extraction from the non-discriminatory monopolistic market and to a decrease in overall profits. In this region, differently from the region of M.S.P., there is an upward discontinuity in $d = \frac{1}{2}$, due to the I.C.E. that drives basic prices down when the duopolistic markets are not fully covered by data. Indeed, for $d \in [\frac{1}{2},1]$, all consumers in each duopolistic market are offered a tailored price and firms set basic prices only for captive consumers: as in the case of monopolistic sub-markets prevalence, the monopolistic market will be fully covered with basic price equal to $v-1$. However, in the case of overlapping segments, as n increases, the intensified competition effect will affect profits only through competitive markets and, differently from the non-overlapping segments scenario, will not drive down the rent extraction from captive consumers.

2.3.1 Monopolistic Sub-Markets Prevalence

In the region of Monopolistic Sub-Markets Prevalence, every firm, whether operating in the Inside or Outside Option, will focus its efforts on extracting profits from the group of captive consumers.

Assuming $v \in [\frac{3}{2}, 2]$, if $d = 0$ equilibrium prices boil down to the Region III of Remark 1. In this region of v, for values of $n \in [2, N-1]$, basic prices will be maintained at a level above $v-1$ (elastic monopolistic demand). Introducing data, as the value of d increases, the non-discriminatory market diminishes. On one hand, firms no longer find it advantageous to carve out a portion of the monopolistic market. On the other hand, if there is a prevalence of captive consumers, they are not compelled by competitive forces to excessively lower their baseline prices.

In the Inside Option, all firms hold a quantity d of data and set an equal price p_B^{W*}, capturing the set of consumers described in (2.17). The system of reaction functions for all firms enables the determination of equilibrium prices and profits, as outlined in the following propositions:

Proposition 1. [All Firms Informed | Equilibrium Prices for $\mathbf{d} \in [\mathbf{0},\mathbf{1}]$] *Assuming $v \in [\frac{3}{2}, 2]$, when all firms buy a quantity $d \in [0,1]$ of data, the equilibrium basic prices as a function of (d, n, v, N) will be:*

- *if $n \in [2, n_2^I(v, N))$ and $d \in [0, \frac{1}{2})$*

$$p^{W*}(d, n, v, N) = \begin{cases} p_1^{W*} \ || \ 0 \leq d < d_1^I & || \ p^{W*} > v-1 \\ p_2^{W*} \ || \ d_1^I \leq d < \frac{1}{2} & || \ p^{W*} = v-1 \end{cases} \tag{2.22}$$

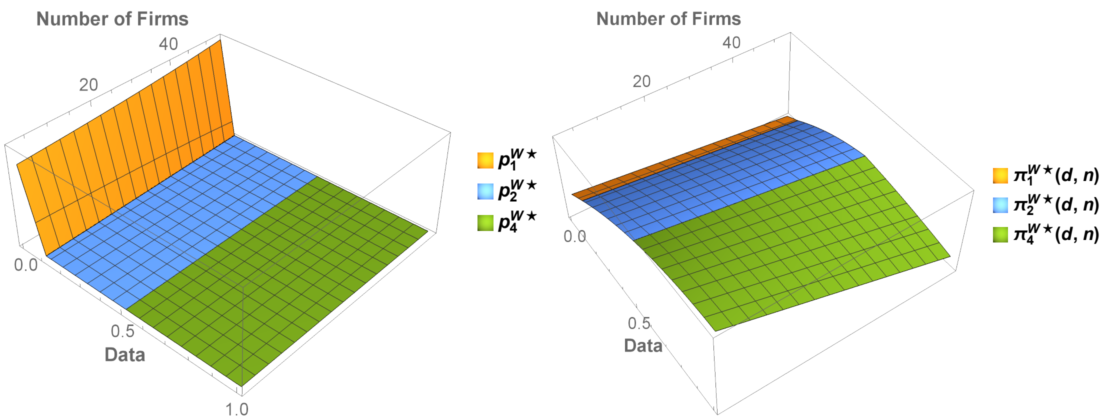
- *if $n \in [2, N-1]$ and $d \in [\frac{1}{2}, 1]$*

$$p_4^{W*} \ || \ \frac{1}{2} \leq d \leq 1 \ || \ p^{W*} = v-1 \tag{2.23}$$

$$p_1^{W*} = \frac{2v(N-n) - 2d(N-1) + (n-1)}{4N - 3n - 1}$$

$$d_1^I(n, v, N) = \frac{(2-v)(2N-n-1)}{2(N-1)} \in [0, \frac{1}{2}) \ \forall \ (n, v, N)$$

As d increases, non-discriminatory consumers in the duopolistic sub-markets, on average, are located further away from each firm's position. To attract these consumers away from their competitors, informed firms reduce their basic prices (2.22). However, Proposition 1 demonstrates that there is a threshold value lower than $\frac{1}{2}$, denoted as $d_1^I(n, v, N)$, beyond which firms no longer find it advantageous to decrease prices as this implies an unsustainable reduction in the rent extraction from captive consumers (Figure 2.5,a).



(a) Inside Option Basic Prices (b) Inside Firm Profits

Figure 2.5. Inside Option profits and basic prices in the region of M.S.P., for $d \in [0,1]$, $n \in [2, n_2^I(v) = 53]$, $v = \frac{19}{10}$ and $N = 100$

For $d \in [\frac{1}{2},1]$, each firm maximizes its non-discriminatory profit exclusively among unidentified captive consumers (2.23). Consequently, the firm's base price can no longer be utilized to exert competitive pressure in duopolistic sub-markets. As a result, the firm's uniform price p_4^{W*} and its range of definition are independent of the number of filled spokes n (Figure 2.5,a). In the Hotelling framework, when monopolistic sub-markets are absent, as d approaches $\frac{1}{2}$, basic prices tend to approach zero and Bertrand competition takes place on the entire unit line (Bounie et al. [2021]). This occurs because there are no more unidentified consumers left to be served in the market. Differently from Hotelling, in the Spokes Model, each firm establishes its basic price to serve the remaining part of non-discriminatory captive consumers and employs it as a benchmark to determine customized prices (2.9) for competitive consumers for whom it possesses exclusive information.

Lemma 1.1. *For prices and boundaries defined in* ***Proposition*** *1 the following statements hold:*

1) p^{W*} *is continuous in* d*;*

2) p_1^{W*} *is decreasing in* d*,* $\forall\ d \in [0, d_1^I)$*, and increasing in* n*,* $\forall\ d \in [0, \frac{2-v}{3}]$*.*

3) $d_1^I(n, v)$ *is decreasing in* n *and in* v*.*

As in Chen and Riordan [2007], when $d \in [0, d_1^I(n, v))$, point 2) of Lemma 1.1 states that there exists a specific region within this interval where an increase in entry leads to higher basic prices: the property is described in Section 2.1 and is due to the fact that, as a firm lowers its price, the marginal consumer in the monopoly segment always has zero surplus (infinite elasticity). As d gets higher ($d \geq \frac{2-v}{3}$), the non-discriminatory market shrinks, and each firm tends to carve out fewer captive consumers. This re-establishes the familiar influence of new entries in oligopolistic competition by restoring the balance between the two components of demand elasticity, leading to a reduction in prices as n increases.

Moreover, as stated in point 3), when both n and v increase, firms are motivated to decrease their basic prices for lower values of $d_1^I(n, v, N)$. This incentive arises due to the following reasons: when v increases, reducing the base prices to $v-1$ results in a relatively smaller decline in profits, and as n increases, competitive pressure compels firms to reduce their prices (Figure 2.5).

In the following Lemma 1.2 we highlight the distinctive features of Inside Option profits in the case of Monopolistic Sub-Markets Prevalence.

Lemma 1.2. [All Firms Informed | Equilibrium Profits for d ∈ [0,1]] *For the prices and boundaries defined in* ***Proposition*** *1, firms' profits are the following:*

$$\pi^{W*}(d, n, v, N) = \begin{cases} \pi_1^{W*} \mid\mid 0 \leq d < d_1^I & \mid\mid p^{W*} > v - 1 \\ \pi_2^{W*} \mid\mid d_1^I \leq d < \frac{1}{2} & \mid\mid p^{W*} = v - 1 \\ \pi_4^{W*} \mid\mid \frac{1}{2} \leq d \leq 1 & \mid\mid p^{W*} = v - 1 \end{cases}$$

$$\pi_2^{W*} = \frac{-d^2(n + N - 2) + 2d(N - 1) + (v - 1)(2N - n - 1)}{(N - 1)N} \tag{2.24}$$

$$\pi_4^{W*} = \frac{2d^2(n - N) + 4d(-nv + N + v - 1) + n + 4(N - 1)v - 4N + 3}{2(N - 1)N} \tag{2.25}$$

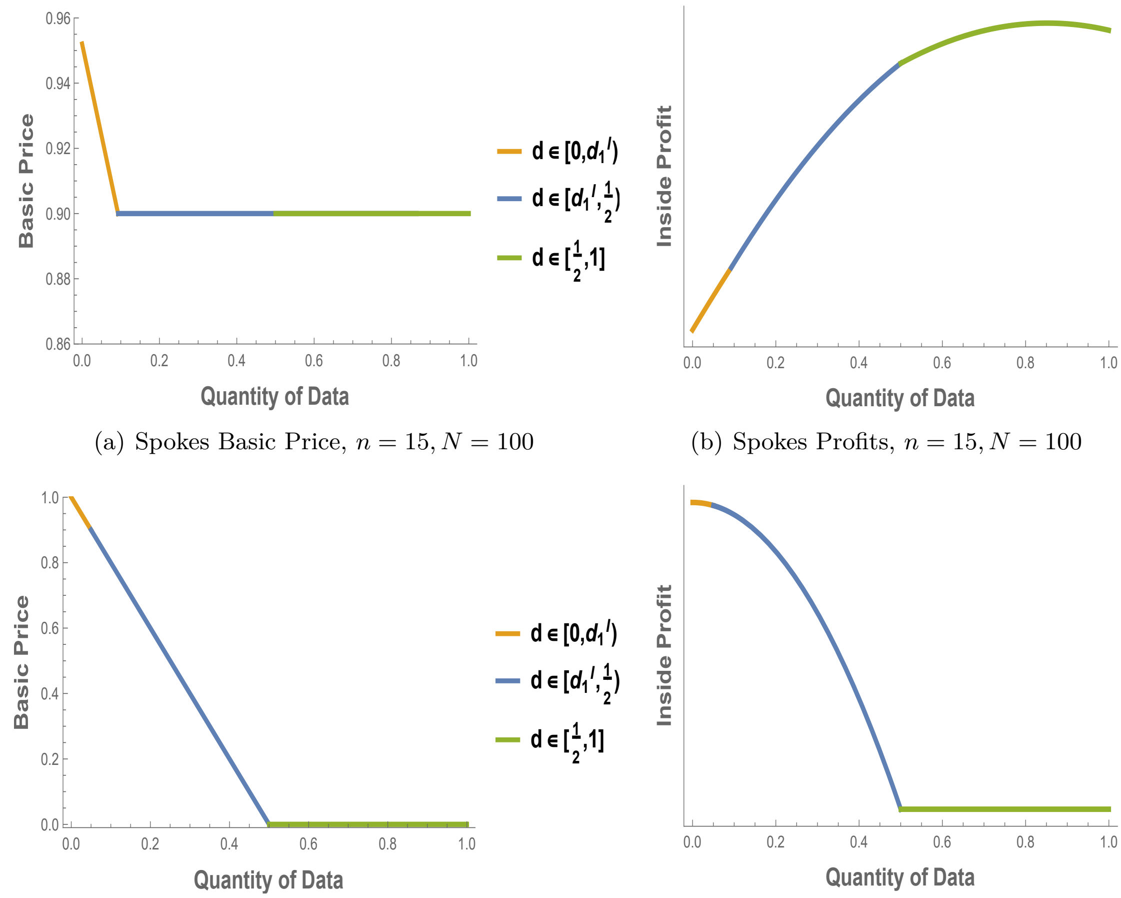
- π_1^{W*} *is increasing in* d, $\forall\ d \in [0, d_1^I)$;
- π_2^{W*} *is increasing in* d, $\forall\ d \in [d_1^I, \frac{1}{2})$;
- π_4^{W*} *is increasing in* d, $\forall\ (d, n) \in [\frac{1}{2}, d^{\pi_4^{W*}}(n, v, N)) \times [2, n^{\pi_4, I}(v, N))$ *and always decreasing otherwise;*

$$d^{\pi_4^{W*}}(n, v, N) = \frac{N - 1 - v(n - 1)}{N - n} \tag{2.26}$$

See the Appendix for the analytic expressions of π_1^{W*}.

Within the domain of monopolistic sub-markets dominance, for $d \in [0, \frac{1}{2})$, the S.E.E. resulting from the identification of captive consumers (last term in (2.15)) surpasses the decrease in rent extraction from non-discriminatory consumers caused by the declining trend of basic prices (first and second term in (2.15)). Moreover, for $d \in [d_1^I, \frac{1}{2})$, firms target a specific kink of the non-discriminatory demand function ($p_2^{W*} = v-1$) and profits (π_2^{W*}) are exclusively moved and driven up by price discrimination (Figure 2.6,b).

Since the analytical expression of prices and profits in the case of overlapping segments does not depend on $n_2^I(v, N)$, in this sub-chapter, we will focus the analysis on the impact



(a) Spokes Basic Price, $n = 15, N = 100$

(b) Spokes Profits, $n = 15, N = 100$

(c) Hotelling Basic Price, $n = 2, N = 2$

(d) Hotelling Profits, $n = 2, N = 2$

Figure 2.6. Prices and profits comparison between Hotelling and Spokes Model when all firms have the same amount of information

of a general prevalence of monopolistic sub-markets on profit outcomes. In particular, for $d \in [\frac{1}{2},1]$, each firm's profits are driven exclusively by the second and third component of (2.16): respectively the rent extraction effect that arises from prices tailored to competitive and monopolistic markets. When n is sufficiently small compared to $N - n$, the S.E.E. originated from identified captive consumers (2.13) outweighs the Bertrand competition effect that occurs in overlapping duopolistic segments (2.9). As a consequence, π_4^{W*} is increasing until values of $d^{\pi_4^{W*}}(n, v, N)$ very close to 1 (Figure 2.5,b).

In the Hotelling framework, when both firms possess data, as is the case in all models of non-localized spatial competition without elastic monopolistic demand, profits decrease for $d \in [0, \frac{1}{2})$. However, once the competitive market is fully covered by data, profits remain constant. This pattern emerges because, as d increases, both firms adopt an aggressive approach in setting their basic prices to expand their market share. However, once firms identify all consumers in their market segment ($d = \frac{1}{2}$), additional data only serve to identify consumers that cannot be attracted away, thereby exerting no impact on either the basic pricing strategy or profits. Differently from Hotelling, in the Spokes Model, when there are no more non-discriminatory competitive consumers to be poached, the existence of a monopolistic market leads to ongoing changes in the profit function for $d \in [\frac{1}{2}, 1]$, influenced by the last two terms of (2.16).

Recalling that the Spokes Model with $N = 2$ and $n = 2$ boils down to the Hotelling framework, we give a representation of the profits in Figure 2.6, highlighting the distinctive feature of these two settings.

In the Outside Option scenario, one of the n firms rejects the DB's offer and faces $n-1$ competitor holding d data. The uninformed firm set a uniform price p_B^{L*} captures the set of consumers described in (2.21). Moreover, in the case of overlapping segments each informed firm will use its basic price p_B^{WL*} not only to serve captive consumers but also to compete with the uninformed firm on the shared non-discriminatory duopolistic market. Based on the condition that d is greater or lower than $\frac{1}{2}$, we will describe firms' profits and equilibrium prices in the following Propositions 2 and 3.

Proposition 2. [One Firm Uninformed | Equilibrium Prices for $\mathbf{d} \in [\mathbf{0}, \frac{\mathbf{1}}{\mathbf{2}})$] *Assuming that the DB offers a quantity $d \in [0, \frac{1}{2})$ of data, when $n \in [2, n_2^O(v, N))$ and one firm rejects the DB's offer, whereas all the other firms accept, the equilibrium basic prices and differences between prices as a function of (d, n, v, N) will be:*

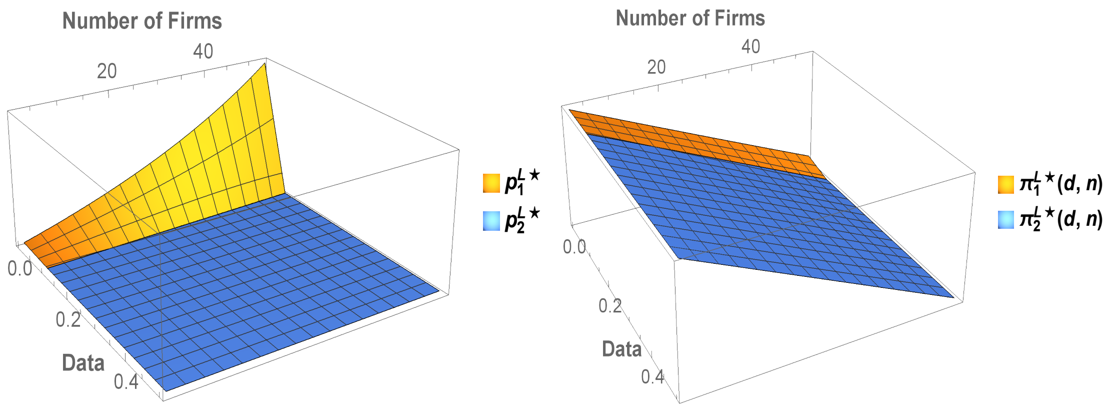
$$p^{L*}(d, n, v) = \begin{cases} p_1^{L*} || \ \left(p^{L*} > v-1, p^{WL*} > v-1\right) \ || \ 0 \leq d < d_1^O \\ p_2^{L*} || \ \left(p^{L*} > v-1, p^{WL*} = v-1\right) \ || \ d_1^O \leq d < \frac{1}{2} \end{cases} \tag{2.27}$$

$$p^{L*}(d, n, v) = \begin{cases} p_1^{L*} = \dfrac{2d(n + N - 1 - nN) + (4N - 2n - 1)(2Nv - n(2v-1) - 1)}{(4N - 2n - 1)(4N - 3n - 1)} \\ p_2^{L*} = \dfrac{v}{2} \end{cases}$$

$$p^{L*}(d, n, v) - p^{WL*}(d, n, v) = \begin{cases} D_1^{O*} = \dfrac{2d(N-1)}{4N - 2n - 1} \\ D_2^{O*} = 1 - \dfrac{v}{2} \end{cases}$$

$$d_1^O(n, v, N) = \frac{(2-v)(4N - 2n - 1)}{4(N-1)} \in [0, \frac{1}{2}) \ \forall \ (n, v, N)$$

The presence of information asymmetry puts the only uninformed firm in the market at a disadvantage compared to its rivals, for two primary reasons. Firstly, it lacks the ability to charge customized prices to extract surplus from consumers. Secondly, it is limited to using the basic price for surplus extraction, while its competitors can utilize both the basic price and tailored prices. Consequently, the uninformed firm ends up setting a higher basic price, leading to a loss in market share. Moreover, within the range of $v \in [\frac{3}{2}, 2]$, any firm without data would prefer to maintain its uniform price above $v - 1$ in order to capitalize on the monopolistic market.



(a) Uninformed Firm's Basic Price (b) Uninformed Firm Profits

Figure 2.7. Outside Option profits and prices in the region of M.S.P. when the DB offers non-overlapping segments ($d \in [0, \frac{1}{2})$) for $n \in [2, n_2^I(v) = 53]$, $v = \frac{19}{10}$ and $N = 100$

As in the Inside Option case, Proposition 2 shows that also when one firm is uninformed and n is in the region of M.S.P., basic prices are greater or equal to $v - 1$ (2.27). Each informed firm has no incentive to lower its uniform price (p_B^{WL}) under $v - 1$ and focuses on extracting profit from the monopolistic sub-markets as they have greater relevance than the competitive duopolistic segments (Figure 2.9,a). For small values of d, firms with data find it convenient to carve out a portion of captive consumers and as d increases they grant full market coverage with marginal captive consumer's surplus at zero, poaching more market share to the firm without data. At the same time, the firm without data chooses to maintain a consistently higher price, capitalizing on its monopolistic demand (Figure 2.7,a). Indeed, for $d \in [d_1^O, \frac{1}{2})$ the uninformed firm anticipates that all its informed competitors will target a specific kink of the non-discriminatory demand function ($p_2^{WL*} = v - 1$), and maximizes fully ($p_2^{L*} = \frac{v}{2}$) the rent extraction from captive consumers (second term of (2.21)).

Lemma 2.1. *For the prices and boundaries defined in* ***Proposition*** *2, the followings are valid:*

1) *p_1^{L*} is decreasing in d, $\forall\ d \in [0, d_1^O)$;*

2) *p_1^{L*} is increasing in n, $\forall\ d \in [0, d_{p_1}^{L*}) \subset [0, d_1^O)$. Moreover, $d_{p_1}^{L*}(n, v, N)$ is decreasing in n, $\forall\ n \in [2, N-1]$;*

3) *D_1^{O*} is increasing in n and d, $\forall\ d \in [0, d_1^O)$ and $\forall\ n \in [2, N-1]$;*

4) *Uninformed firm's profit are decreasing in d, $\forall\ d \in [0, d_1^O)$ and constant in d, $\forall\ d \in [d_1^O, \frac{1}{2})$*

See the Appendix for the analytic expressions of $d_{p_1}^{L*}(n, v, N)$.

As in the Inside Option case, for both firms, there exists a region of d, where an increase in n, leads to higher uniform price due to similar factors that were discussed in the context of demand elasticity balancing, as explained for Lemma 1.1. In particular, it can be easily shown that firms with data increase their basic price in response to new entries until values of d lower than $d^{L*}_{p_1}(n, v, N)$.

Independently from the fact that prices decrease or increase with respect to n, point 3) of Lemma 2.1 states that, as new firms enter the market and as d increases, the difference between prices increases and firms with data conquer more consumers on the duopolistic spokes shared with the uninformed firm. These phenomena arise for values of d where both prices are greater than $v-1$ and can be explained by considering that firms with access to information are less affected by a uniform price reduction, as they serve a smaller portion of unidentified consumers. Moreover, as the number of firms (n) increases, the significance of monopolistic sub-markets diminishes but, when the uninformed firm decreases (or increases) its price, the resulting reduction (or gain) in rent extraction from captive consumers is more pronounced compared to the informed firms. As a result, until the uninformed firm finds it convenient to keep its monopolistic demand elastic, it will respond less aggressively than its informed competitors to market entries.

In the outside option scenario, when $d \in [\frac{1}{2}, 1]$, the analytic expression of firm's equilibrium prices and profits do not depend on $n^I_1(v, N)$. We will use the following proposition also in Section 2.3.2 and we describe in Lemma 3.1 the effect of n on equilibrium outcomes.

Proposition 3. [One Firm Uninformed | Equilibrium Prices for $\mathbf{d} \in [\frac{1}{2}, 1]$] *Assuming that the DB offers a quantity $d \in [\frac{1}{2}, 1]$ of data, when $n \in [2, N-1]$ and one firm rejects the DB's offer, whereas all the other firms accept, the equilibrium basic prices and differences between prices as a function of (d, n, v, N) will be:*

$$p^{L*}(d,n,v) = \begin{cases} p^{L*}_6 || & \left(p^{L*} > v-1, p^{WL*} = v-1\right) || \ \frac{1}{2} \le d < d^O_5 \\ p^{L*}_7 || & \left(p^{L*} > v-1, p^{WL*} = v-1\right) || \ d^O_5 \le d < d^O_6 \\ p^{L*}_8 || & \left(p^{L*} > v-1, p^{WL*} = v-1\right) || \ d^O_6 \le d < d^O_7 \\ p^{L*}_9 || & \left(p^{L*} > v-1, p^{WL*} = v-1\right) || \ d^O_7 \le d \le 1 \end{cases} \tag{2.28}$$

$$p^{L*}(d,n,v) = \begin{cases} p^{L*}_6 = \dfrac{v}{2} \\ p^{L*}_7 = 2d + v - 2 \\ p^{L*}_8 = \dfrac{1}{2}\left(\dfrac{(1-d)(n-1)}{N-n} + v\right) \\ p^{L*}_9 = p^{L*}_8(d = d^O_7) = \dfrac{2v(N-n)+n-1}{4N-3n-1} \end{cases}$$

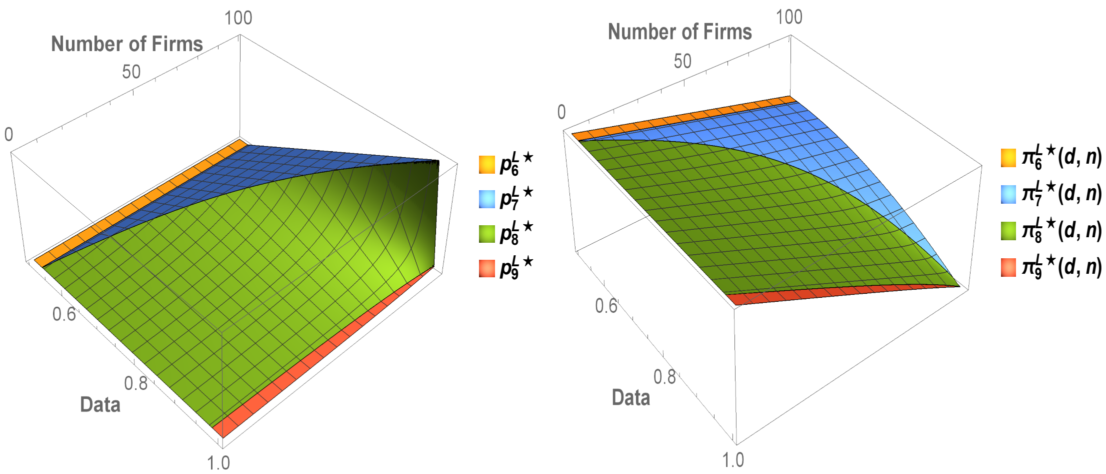
$$p^{L*}(d,n,v) - p^{WL*}(d,n,v) = \begin{cases} D^{O*}_5 = 1 - \dfrac{v}{2} \\ D^{O*}_6 = 2d - 1 \\ D^{O*}_7 = \dfrac{1}{2}\left(\dfrac{(1-d)(n-1)}{N-n} - v + 2\right) \\ D^{O*}_8 = \dfrac{(2-v)(2N-n-1)}{4N-3n-1} \end{cases}$$

$$d_5^O(v) = \frac{4-v}{4}$$

$$d_6^O(n,v,N) = \frac{v(n-N)}{4N-3n-1} + 1$$

$$d_7^O(n,v,N) = \frac{N(v+2)-n(v+1)-1}{4N-3n-1}$$

When $d \in [\frac{1}{2},1]$, there exists a threshold value $d_5^O(v)$ beyond which every firm with data should lower its basic price (p_B^{WL}) under $v-1$ to fully exploit the competitive advantage offered by information but instead keeps its uniform price equal to $v-1$ (Figure 2.9,a). Consequently, for $d \geq d_5^O(v)$ each informed firm fails to attract more non-discriminatory consumers in the duopolistic market shared with its uninformed competitor, as this would imply an unsustainable reduction in the rent extraction from captive consumers. It is worth noting that d_5^O remains independent of the value n. Therefore, regardless of whether the market is predominantly competitive or monopolistic, the significance of poaching additional consumers from the only uninformed firm is lower in comparison to extracting profits from captive consumers.



(a) Uninformed Firm's Basic Price

(b) Uninformed Firm Profits

Figure 2.8. Outside Option profits and prices when the DB offers overlapping segments ($d \in [\frac{1}{2},1]$) for $n \in [2,99]$, $N = 100$ and $v = \frac{19}{10}$

Anticipating that the informed firms' basic prices are fixed at $v-1$, the uninformed competitor recognizes that it can raise its basic price (p_B^L) until the indifferent consumer falls within the range where the informed firm practices price discrimination, i.e. $p_B^L \leq p_7^{L*}$. More specifically, the uninformed firm will face a demand function comprising both an inelastic competitive component and an elastic monopolistic component (refer to the Appendix for precise definitions). As long as $d \in [d_5^O, d_6^O)$, the uninformed firm increases its price up to p_7^{L*} in order to extract the maximum profit from the relatively inelastic competitive demand (Figure, 2.8,a). However, for $d \in [d_6^O, d_7^O)$, fully leveraging the inelasticity of the duopolistic demand would result in the loss of a substantial number of consumers in the monopolistic sub-markets. Consequently, the uninformed firm lowers its basic price at p_8^{L*} (below p_7^{L*}) in order to retain a larger customer base.

Furthermore, the uninformed firm is aware that its $n-1$ informed competitors will capture identified consumers until the tailored price for the shared duopolistic markets exceeds zero, i.e. (2.19) ≥ 0. When $d \geq d_7^O$, the firm lacking data ceases to decrease its base price, anticipating that further competitive consumers will not be poached ($p_9^{L*} = p_8^{L*}(d = d_7^O)$).

Lemma 3.1. *For the prices and boundaries defined in* ***Proposition*** *3, the followings are valid:*

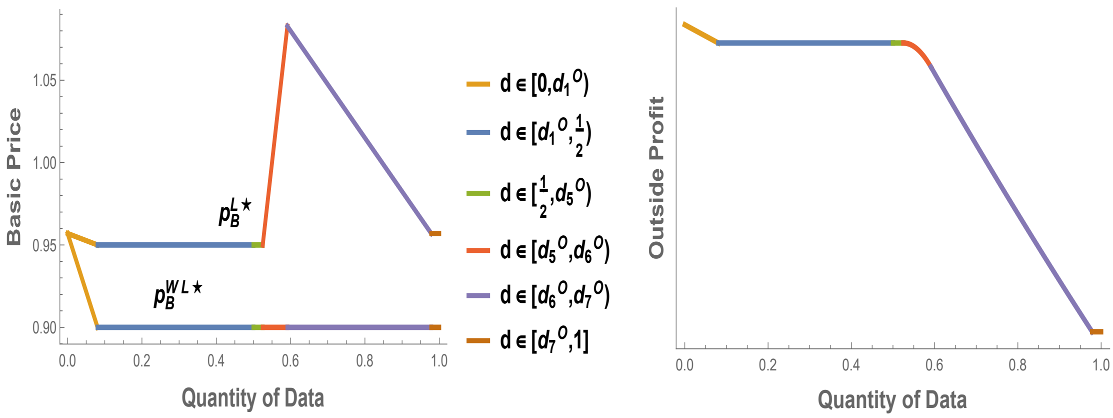
1) p_8^{L*} is increasing in n, $\forall\ n \in [2, N-1]$ and decreasing in d, $\forall\ d \in [d_6^O, d_7^O)$;

2) d_6^O is increasing in n, $\forall\ n \in [2, N-1]$;

3) d_7^O is increasing in n, $\forall\ n \in [2, N-1]$. In particular, $d_7^O = 1$ for $n = N$;

4) Uninformed firm's profit is decreasing in d, $\forall\ d \in [d_5^O, d_7^O)$ and constant in d, $\forall\ d \in [\frac{1}{2}, d_5^O)$ and $\forall\ d \in [d_7^O, 1]$.

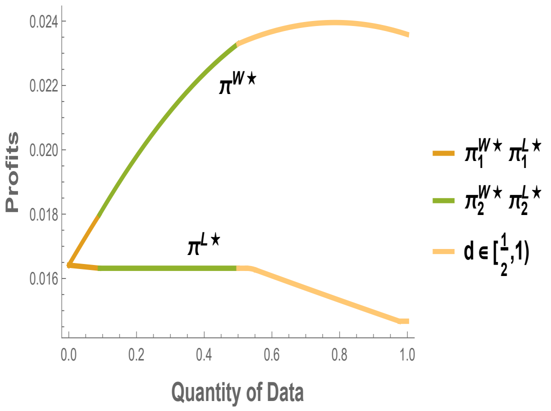
As n increases, points 1) and 2) of Lemma 3.1, shows on one hand that an uninformed firm has the incentive to raise price in the region $[d_6^O, d_7^O)$ and on the other hand that greater relevance acquired by competitive sub-markets (where the demand is relatively inelastic) makes convenient to raise uniform price at its maximum level (p_7^{L*}) for greater values of d_6^O.



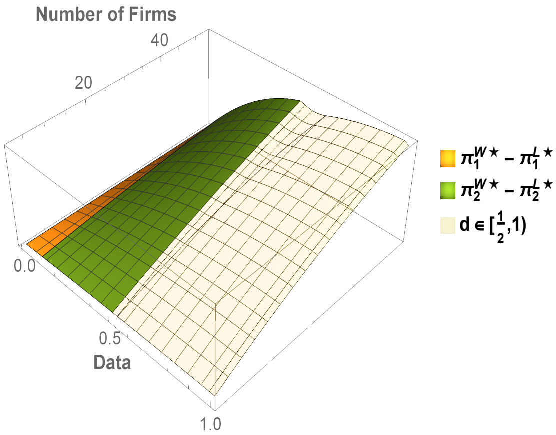
(a) Outside Option Basic Prices, $n < n_2^I(v)$ (b) Uninformed Firm Profits, $n < n_2^I(v)$

Figure 2.9. Outside Option profits and basic prices in the region of M.S.P., for $d \in [0,1]$, $v = \frac{19}{10}$ and $N = 100$

It is worth noticing in Figure 2.9 (a) that, if monopolistic sub-markets dominate, the Outside Option basic prices remain continuous as the value of d approaches $\frac{1}{2}$ from the left. In this scenario, all informed firms unanimously strive to maintain rent extraction from non-discriminatory captive consumers, while avoiding a price war. This holds even as each duopolistic segment not covered by information diminishes and the conquest of competitive unidentified consumers becomes increasingly challenging.



(a) Firm's Profits, $n < n_2^I(v)$



(b) DB's Profits, $n < n_2^I(v)$

Figure 2.10. DB's Profits in the region of Monopolistic Sub-Markets Prevalence for $d \in [0,1]$, $n \in [2, n_2^I(v) = 53]$, $v = \frac{19}{10}$ and $N = 100$

In the upcoming observation, we present a distinctive feature of the DB's strategy that will prove beneficial in determining the optimal quantity of data to be sold. Specifically, we use the notation d^{DB*} to denote the optimal quantity of data to be sold when the complete data set is available for offering, i.e. $d \in [0,1]$.

Observation 1. *For $n < n_2^I(v, N)$, DB's profits are increasing for $d \in [0, \frac{1}{2})$ and consequently:*

- *If the DB can offer a quantity of data $d \in [0,1]$, its optimum strategy d^{DB*} falls in $[\frac{1}{2},1]$ (non-exclusive information).*

For $d \in [0, \frac{1}{2})$, the Inside Option profits are always increasing in d (Lemma 1.2) and uninformed firm's profits in the Outside Option scenario exhibits always a constant or decreasing behavior (Lemma 2.1). Moreover, the continuity of prices and profits is ensured for $d = \frac{1}{2}$. Recalling that the DB aims to maximize (2.7), it becomes evident that the DB's profits exhibit an increasing behavior within the range of $d \in [0, \frac{1}{2})$.

2.3.2 Competitive Sub-Markets Prevalence

When n is greater than $n_2^I(v, N)$, the portion of the market not covered by discriminatory pricing, where competition occurs based on uniform prices, becomes more significant than the non-discriminated monopolistic market. Moreover, capturing consumers located far from the origin of each spoke (where firms are placed) becomes more challenging, as the majority of them have a preference for two specific brands. Consequently, the prevalence of competitive sub-markets leads to a reduction of all basic prices under $v-1$ for sufficiently high values of d.

For what concerns the equilibrium outcomes in the case of overlapping partitions, all the results exposed in Section 2.3.1 hold for $n \in [2, N-1]$. Therefore, for $d \in [\frac{1}{2}, 1]$, in this Section we will limit to highlight the effect of a generic competitive sub-markets prevalence on equilibrium prices and profits.

On the other hand, regarding the region of d which falls in $[0, \frac{1}{2})$, we will extend Proposition 1 and 2, to outline the structure of equilibrium prices and profits for values of n that exceed $n_2^I(v, N)$. Furthermore, to facilitate clarity and computational simplicity, we present and prove Proposition 2 while considering a fixed value of $N = 100$.

Proposition 4. [Extension of Proposition 1] *Assuming $v \in [\frac{3}{2}, 2]$, when all firms buy a quantity $d \in [0,1]$ of data, the equilibrium basic prices will be:*

- *if $n \in [n_2^I(v, N), N-1]$ and $d \in [0, \frac{1}{2})$*

$$p^{W*}(d, n, v, N) = \begin{cases} p_1^{W*} \mid\mid 0 \leq d < d_1^I & \mid\mid p^{W*} > v-1 \\ p_2^{W*} \mid\mid d_1^I \leq d < d_2^I & \mid\mid p^{W*} = v-1 \\ p_3^{W*} \mid\mid d_2^I \leq d < d_{NE}^I & \mid\mid p^{W*} < v-1 \\ NE \mid\mid d_{NE}^I \leq d < \frac{1}{2} & \mid\mid \textit{No Equilibria} \end{cases} \tag{2.29}$$

- *if $n \in [2, N-1]$ and $d \in [\frac{1}{2}, 1]$*

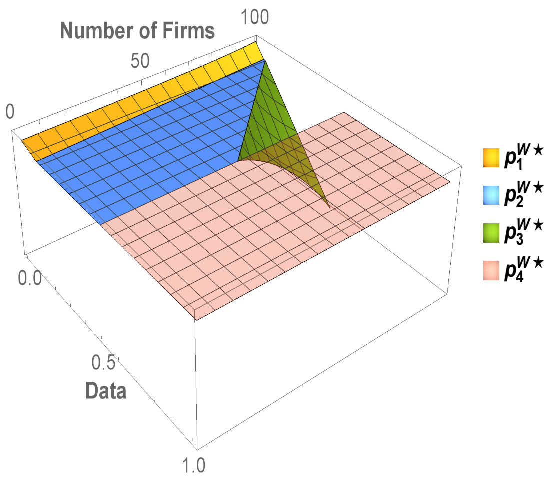
$$p_4^{W*} \mid\mid \frac{1}{2} \leq d \leq 1 \mid\mid p^{W*} = v-1 \tag{2.30}$$

$$p_3^{W*} = \frac{2N - 2d(N-1) - n - 1}{n-1}$$

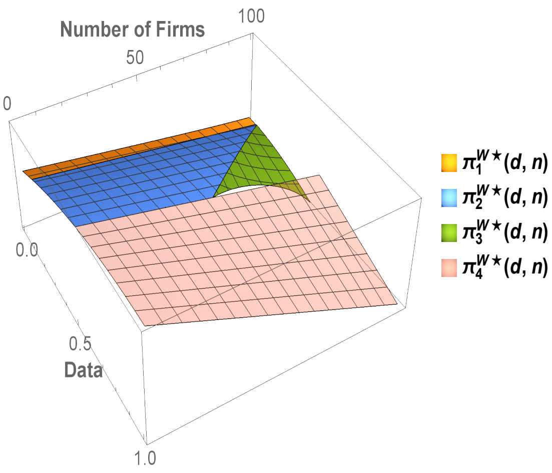
$$d_2^I(n, v, N) = 1 - \frac{(n-1)v}{2(N-1)}$$

See Proposition 1 and Appendix for analytic expression of p_1^{W*}, p_2^{W*}, d_1^I and d_{NE}^I.

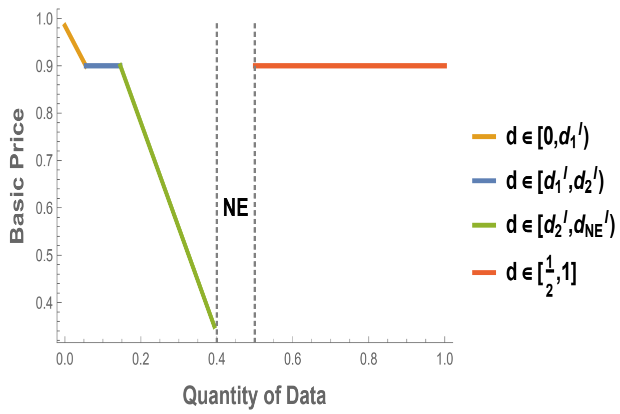
Proposition 4 highlights one distinctive characteristic of the Spokes Model with price discrimination: the absence of pure strategy Nash equilibria resulting from intensified competition for non-discriminatory consumers in duopolistic markets (2.29). In the region of competitive sub-markets prevalence, when d and v are sufficiently high, to protect their non-discriminatory consumers and attempt to attract an additional part of them, firms must lower basic prices (p_B^W) to such an extent that it becomes advantageous to raise



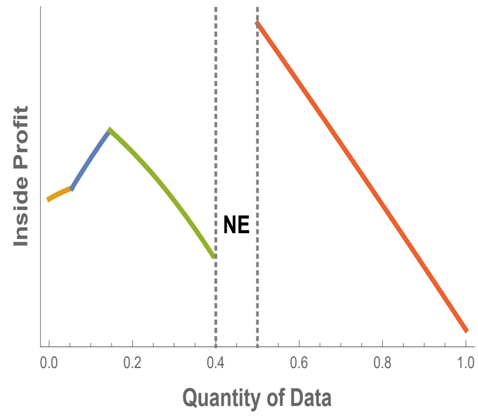
(a) Inside Option Basic Price, $2 \leq n \leq N-1$



(b) Informed Firm Profits, $2 \leq n \leq N-1$



(c) Inside Option Basic Price for $n \geq n_2^I$



(d) Informed Firm Profits for $n \geq n_2^I$

Figure 2.11. Inside Option prices and profits in the region of Competitive Sub-Markets Prevalence for $d \in [0,1]$, $N = 100$ and $v = \frac{19}{10}$

them to $v-1$, focusing solely on the monopolistic component. Hence, there is a specific range of values for d, represented by $[d_{NE}^I, \frac{1}{2})$, within which a profitable deviation from the candidate equilibrium price exists (Figure 2.11,a & c).

It is easy to prove that d_2^I is decreasing in n. New entries in the market expand the region of d where basic prices (p_3^{W*}) are lower than $v-1$ and intensify competition for the unidentified consumers in duopolistic sub-markets (2.29). Consequently, as shown below, in the region of competitive prevalence, an increase in n or in d, diminishes the rent extraction from non-discriminatory consumers and neutralizes the effect of total surplus extraction from the discriminatory monopolistic sub-markets.

When data segments are overlapped, as already shown in Proposition 1, each firm serves at its basic price only non-discriminatory captive consumers and finds it convenient to ensure full market coverage with a basic price (p_4^{W*}) equal to $v-1$ (2.30). Therefore, as d approaches $\frac{1}{2}$ from the left, the basic price decreases below $v-1$ and then abruptly jumps up to $v-1$ when $d=\frac{1}{2}$, resulting in a discontinuity (Figure 2.11,a & c).

Lemma 4.1. [Extension of Lemma 1.2] *For the prices and boundaries defined in* ***Proposition*** *4, firms' profits are the following:*

$$\pi^{W*}(d,n,v,N)=\begin{cases}\pi_1^{W*} \ || \ 0\leq d<d_1^I & || \ p^{W*}>v-1\\ \pi_2^{W*} \ || \ d_1^I\leq d<d_2^I & || \ p^{W*}=v-1\\ \pi_3^{W*} \ || \ d_2^I\leq d<d_{NE}^I & || \ p^{W*}<v-1\\ \pi_4^{W*} \ || \ \frac{1}{2}\leq d\leq 1 & || \ p^{W*}=v-1\end{cases}$$

- π_1^{W*} *is both increasing and decreasing in d,* $\forall\ (n,d)\in[n^{\pi_1,I}(v,N),N-1]\times[0,d_1^I)$*;*
- π_2^{W*} *is increasing in d,* $\forall\ d\in[d_1^I,d_2^I)$*;*
- π_3^{W*} *is decreasing in d,* $\forall\ d\in[d_2^I,d_{NE}^I)$*;*
- π_4^{W*} *is decreasing in d,* $\forall\ d\in[\frac{1}{2}$,1]*.*

See Lemma 1.2 and Appendix for the analytic expressions of π_1^{W*},π_2^{W*}, π_3^{W*}, π_4^{W*} *and* $n^{\pi_1,I}(v,N)$*.*

As illustrated in Lemma 1.2, profits are mainly driven up by the discrimination of captive consumers, as in each duopolistic sub-market the presence of all firms equipped with the same quantity of data intensifies competition and reduces the rent extraction from unidentified consumers.

When $d\in[0,\frac{1}{2})$, each firm would prefer to maintain its basic price equal to or greater than $v-1$ to extract a substantial rent from non-discriminatory consumers. However, in the region where competitive sub-markets dominate, the majority of consumers belong to duopolistic sub-markets, and the need to protect them from competitors compels firms to lower their basic prices below $v-1$, resulting in decreased profits from non-discriminatory consumers (Figure 2.11,b & d). Therefore, if d is sufficiently high, the reduction in the first two terms of (2.15) outweighs the S.E.E. arising from the last term of (2.15), leading to a decrease in profit π_3^{W*}.

When $d\in[\frac{1}{2}$,1], profits (π_4^{W*}) decrease in d. This is because, when competitive sub-markets dominate, the impact of Bertrand competition intensifies in competitive segments where data overlap, becoming more significant in relation to extracting rent from identified captive consumers.

The discontinuity in prices as indicated by Proposition 4 leads to a corresponding discontinuity in profits when $d=\frac{1}{2}$. Specifically, as d approaches $\frac{1}{2}$ from the left, π_3^{W*} experiences a decline, followed by a sudden increase to π_4^{W*} at $d=\frac{1}{2}$, where the price war disappears completely (Figure 2.11,b & d).

Proposition 5. [Extension of Proposition 2] *Fixing* $N=100$ *and assuming that the DB offers a quantity* $d\in[0,\frac{1}{2})$ *of data, when* $n\in[n_2^O(v)$,99] *and one firm rejects the DB's offer, whereas all the other firms accept, the equilibrium basic prices and differences between prices as a function of* (d,n,v) *will be:*

- *if* $v \in (\frac{11}{7}, 2]$
  - *if* $n \in [n_2^O(v), n_3^O(v))$

$$p^{L*}(d,n,v) = \begin{cases} \{p_1^{L*}, p_2^{L*}\} & \|\ 0 \leq d < d_2^O \\ p_3^{L*} \|\ \ (p^{L*} > v-1, p^{WL*} < v-1) & \|\ d_2^O \leq d < d_{NE}^I \end{cases} \tag{2.31}$$

  - *if* $n \in [n_3^O(v), n_4^O(v))$

$$p^{L*}(d,n,v) = \begin{cases} \{p_1^{L*}, p_2^{L*}, p_3^{L*}\} & \|\ 0 \leq d < d_3^O \\ p_4^{L*} \|\ \ (p^{L*} = v-1, p^{WL*} < v-1) & \|\ d_3^O \leq d < d_{NE}^I \end{cases} \tag{2.32}$$

  - *if* $n \in [n_4^O(v), N-1]$

$$p^{L*}(d,n,v) = \begin{cases} \{p_1^{L*}, p_2^{L*}, p_3^{L*}, p_4^{L*}\} & \|\ 0 \leq d < d_4^O \\ p_5^{L*} \|\ \ (p^{L*} < v-1, p^{WL*} < v-1) & \|\ d_4^O \leq d < d_{NE}^I \end{cases} \tag{2.33}$$

$$p^{L*}(d,n,v) - p^{WL*}(d,n,v) = \begin{cases} D_4^{O*} = \dfrac{198(d-1)+(n-1)v}{n} \\ D_5^{O*} = \dfrac{198d}{2n-1} \end{cases}$$

See Proposition 2 and Appendix for the missing analytic expressions and for a complete outline of equilibrium prices in the case of $v \in [\frac{3}{2}, \frac{11}{7}]$

As demonstrated in Proposition 2, the uninformed firm, facing a competitive disadvantage, sets a higher uniform price (p_B^L) compared to its informed competitors. At the same time, if a firm rejects the DB offer, it will cover with its uniform price the entire monopolistic market thereby preferring to keep the price above $v-1$. Proposition 5 reveals that, when $d \geq d_2^O$ and $n \geq n_2^O$, the non-discriminatory consumer segment shrinks and the greater relevance acquired by each duopolistic sub-market shared with informed competitors forces informed firms' basic price (p_B^{WL}) below $v-1$ (2.31). In line with what occurs in the region of M.S.P. for each informed firm, there exists a region of d where the uninformed firm targets a specific kink of the demand function ($p_4^{L*} = v-1$), as excessively lowering the base price would result in an unsustainable decrease in rent extraction from captive consumers (2.32). Moreover, if the informed firms engage in a highly intense price war and the competitive demand gains significant importance ($d \geq d_4^O$ and $n \geq n_4^O$), the uninformed firm responds by lowering its uniform price below $v-1$ in order to mitigate a substantial loss of consumers within its competitive segment (2.33).

Lemma 5.1. *For the prices and boundaries defined in* ***Proposition*** *5, the followings are valid:*

1) p_3^{L*}, p_4^{L*} *and* p_5^{L*} *are decreasing in* n *and* d*, in their respective region of definition;*

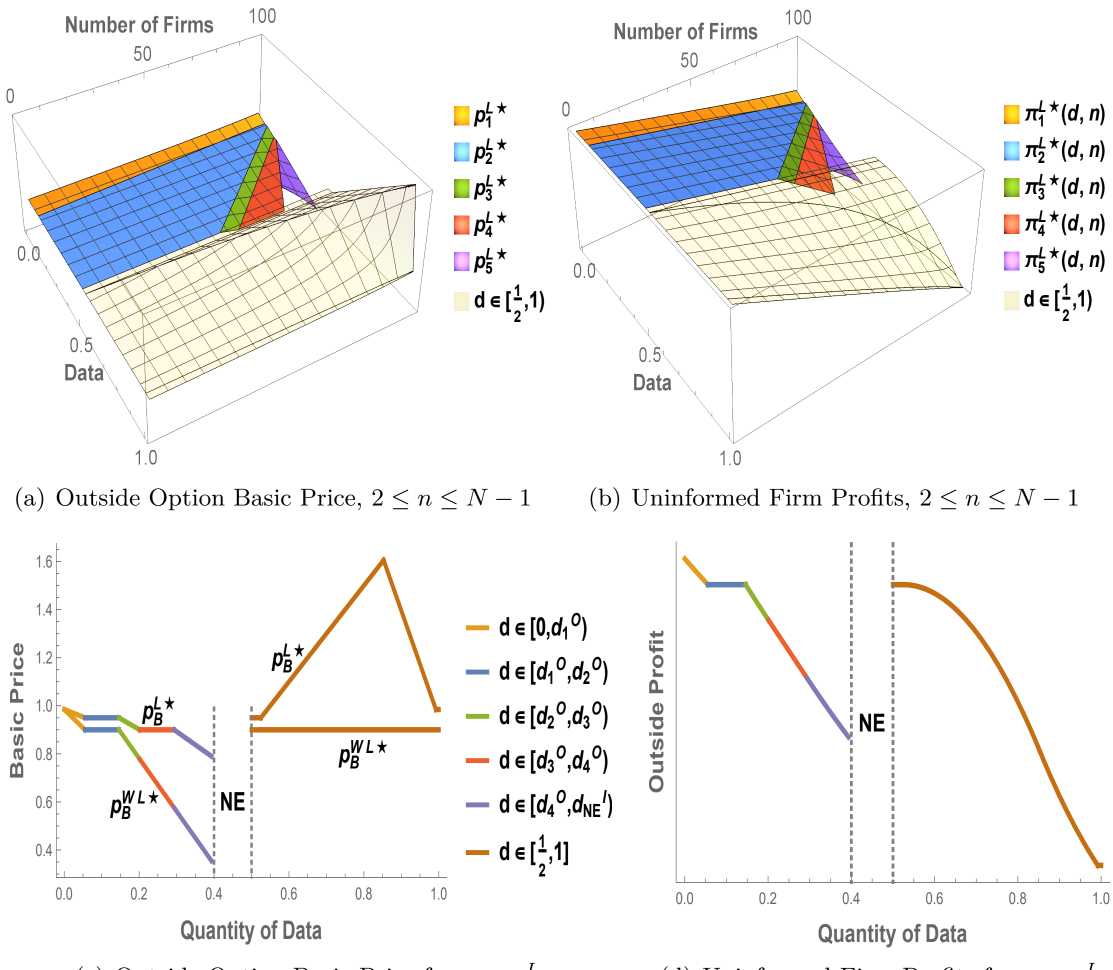
2) D_3^{O*} *and* D_4^{O*} *are increasing in* n *and* d*, in their respective region of definition;*

3) D_5^{O*} *is decreasing in* n *and increasing in* d*, in its respective region of definition;*

4) Uninformed firm's profits are decreasing or constant in d*,* $\forall\ d \in [0, d_{NE}^I)$*.*

Consistently with the findings of Lemma 2.1, the difference between p_B^{L*} and p_B^{WL*} remains increasing in n until the monopolistic demand of the uninformed firm is elastic.

Indeed, for $d \in [d_4^O, d_{NE}^I)$, the firm without data grants full market coverage (2.33) with marginal captive consumer surplus greater than zero. As n increases within this range of d, the associated difference between basic prices D_5^{O*} decreases, indicating that the uninformed firm becomes more assertive in countering the attempts of informed firms to lure away consumers in duopolistic sub-markets.



(a) Outside Option Basic Price, $2 \leq n \leq N-1$ (b) Uninformed Firm Profits, $2 \leq n \leq N-1$

(c) Outside Option Basic Price for $n > n_2^I$ (d) Uninformed Firm Profits for $n > n_2^I$

Figure 2.12. Outside Option profits and prices in the region of Competitive Sub-Markets Prevalence for $d \in [0,1]$, $N = 100$, $v = \frac{19}{10}$

Furthermore, by examining Figure 2.12, we can observe a similar discontinuity occurring in both basic prices and profits, as seen in the Inside Option scenario. As the value of d approaches d_{NE}^I, informed firms initiate a price war, prompting their uninformed competitor to respond more aggressively in an attempt to retain its customer base in duopolistic markets. Consequently, within the region of C.S.P., the uninformed firm's uniform price declines when duopolistic sub-markets shared by informed competitors are not fully covered by data and rise abruptly to p_6^{L*} when $d = \frac{1}{2}$ (2.28). This discontinuity in Outside Option prices leads to a corresponding disruption in the uninformed firm's profits. Specifically, within the range of $d \in [d_2^O, d_{NE}^I)$, the uninformed firm's profits undergo a significant decline, followed by a sudden surge at $d = \frac{1}{2}$ when the basic price war among informed firms completely disappears.

For what concerns the Outside Option scenario, in the case of overlapping segments, Proposition 3 holds true for $n \in [2, N-1]$. Specifically, within a general region of C.S.P., when $d \geq d_5^O$, the inelastic competitive demand encountered by the uninformed firm becomes more significant in relation to the captive consumer base. As the value of n increases, two outcomes arise: firstly, the firm without data will raise its uniform price to the maximum permissible price p_7^{L*} until higher values of d_6^O are reached, and secondly, the uniform price (p_8^{L*}) for $d \in [d_6^O, d_7^O)$ will consistently exceed the prices observed when monopolistic markets dominate.

We expand on Observation 1 by elucidating the distinctive characteristics of the optimal strategy employed by the DB in the region dominated by competitive sub-markets. Specifically, we use the notation $d^{DB,E*}$ to represent the optimal strategy of the DB when selling non-overlapping (exclusive) segments, i.e. $d \in [0, \frac{1}{2})$. The presence of the profits discontinuity at $d = \frac{1}{2}$ adds intrigue to examining the scenario of exclusive information sales, as it leads to fundamental insights regarding the non-exclusive information sale.

Observation 2. [Extension of Observation 1] *Fixing $N = 100$, for $n \in [n_2^I(v), N-1]$, DB's profits are both increasing and decreasing for $d \in [0, \frac{1}{2})$ and consequently:*

1) If the DB is forced to sell a quantity of data $d \in [0, \frac{1}{2})$ (exclusive information) and $v \in [\frac{42}{25}, \frac{9}{5}]$, its strategy $d^{DB,E*}$ can be defined as follows:

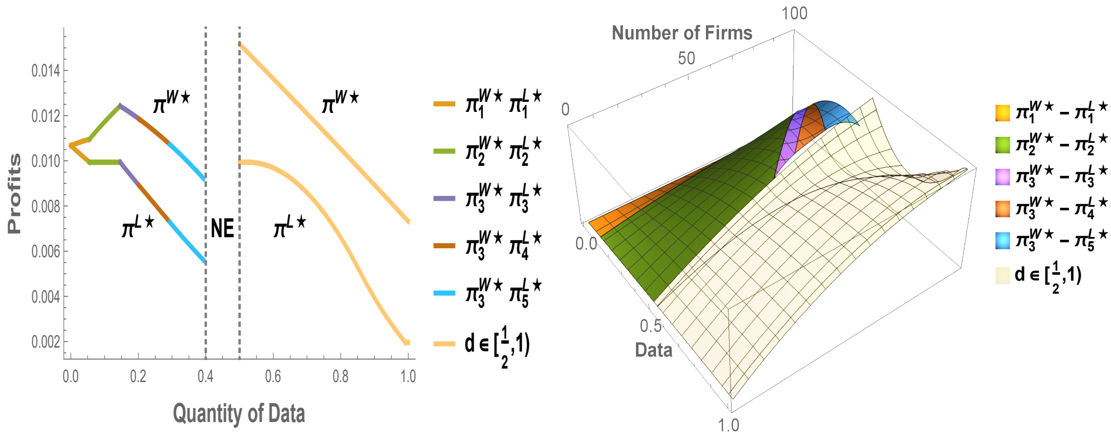
$$d^{DB,E*} = \begin{cases} d_2^I(n,v) \mid\mid n \in [n_2^I(v), n_5^{DB,E*}(v)) & \mid\mid \\ d_5^{DB,E}(n,v) \mid\mid n \in [n_5^{DB,E*}(v), n_5^{DB,E,U*}(v)) & \mid\mid d_5^{DB,E} \in (d_2^O, d_3^O] \\ d_3^O(n,v) \mid\mid n \in [n_5^{DB,E,U*}(v), n_6^{DB,E,L*}(v)) & \mid\mid \\ d_6^{DB,E}(n,v) \mid\mid n \in [n_6^{DB,E,L*}(v), n_7^{DB,E*}(v)) & \mid\mid d_6^{DB,E} \in (d_3^O, d_4^O) \\ d_7^{DB,E}(n,v) \mid\mid n \in [n_7^{DB,E*}, 99] & \mid\mid d_7^{DB,E} \in (d_4^O, \frac{1}{2}) \end{cases}$$

Moreover:

- *$d_5^{DB,E}(n,v)$ is almost always increasing in n;*
- *$d_6^{DB,E}(n,v)$ is almost always increasing in n;*
- *$d_7^{DB,E}(n,v)$ is almost always decreasing in n.*

2) If the DB can offer a quantity of data $d \in [0,1]$ (non-exclusive information), its optimum strategy d^{DB*} falls in $[\frac{1}{2},1]$.

Within the context of C.S.P., as d approaches $\frac{1}{2}^-$, the Data Broker encounters the intensified competition effect present in non-discriminatory duopolistic markets. This effect exerts a profound influence, leading to a dual impact. Firstly, it significantly affects the declining behavior of Inside Option rents. Secondly, it causes an upward discontinuity in the profits of both informed and uninformed firms. When the partitions do not overlap, the competition for non-discriminatory consumers in the competitive sub-markets also lowers the extraction of rents from unidentified captive consumers (Figure 2.11,d). At the same time, if a firm rejects the offer from the Data Broker, all of its informed competitors enter into a price war, enabling them to entice a considerable share of the uninformed firm's consumers in the duopolistic market. This, in turn, intensifies the threat posed by the Outside Option (Figure 2.12,d). Therefore, the DB's optimal strategy for selling exclusive information is influenced by two opposing factors (Figure 2.13,a). On the one hand, the threat of the uninformed firm's outside option pushes the DB's strategy upwards. On the other hand, the I.C.E. that occurs for non-discriminatory consumers in each duopolistic market drives the Inside Option profits and the DB's strategy downwards.



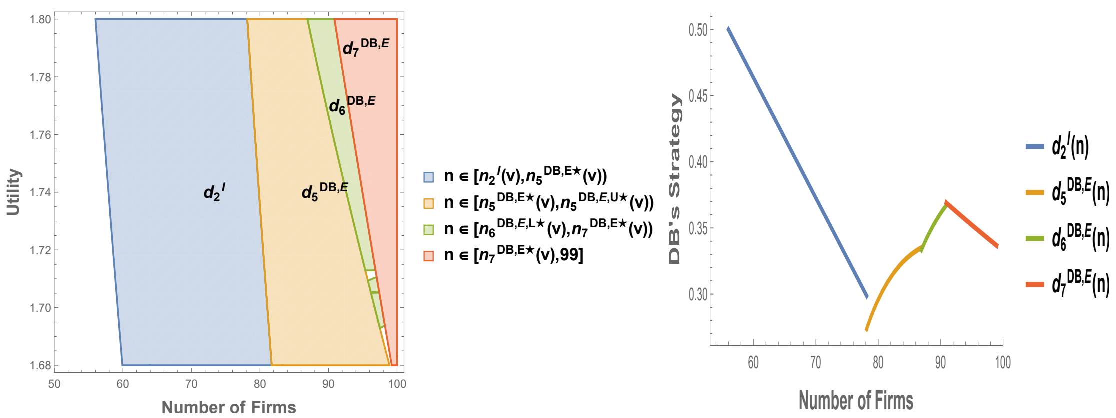
(a) Firm's Profits for $n > n_2^I$ (b) DB's Profits, $2 \leq n \leq N-1$

Figure 2.13. DB's Profits in the region of Competitive Sub-Markets Prevalence for $d \in [0,1]$, $v = \frac{19}{10}$ and $N = 100$

In cases where the value of n is relatively low ($n \in [n_2^I(v), n_5^{DB,E*}(v))$, the DB's profits global optimum aligns with a corner solution (d_2^I). This solution represents the maximum threshold within the d region where basic prices are equal to or greater than $v-1$ (2.29). Going beyond d_2^I by selling more data would lead to a significant decline in rent extraction from captive consumers, without sufficient compensation from the reduction in Outside Option rents. As a result, the intensified competition effect outweighs the threat of the Outside Option, and the optimal strategy demonstrates a decreasing trend as n increases (Figure 2.14,blue).

As n increases, competitive sub-markets become more relevant, intensifying the price war among informed competitors. This poses a substantial threat to the uninformed firm, risking the loss of more consumers in duopolistic markets. Specifically, when $n \in [n_5^{DB,E*}(v), n_7^{DB,E*})$, the optimal strategy lies in a d region where the uninformed firm maintains its base price at or above $v-1$, enabling it to continue extracting rent from captive consumers while also carving out a portion of them ((2.31) and (2.32)). Simultaneously, in order to maintain its monopolistic demand with sufficient elasticity, the uninformed firm becomes more susceptible to the aggressive pricing strategy implemented by its informed competitors. Consequently, within this range of n values, the presence of the Outside Option threat drives the upward movement of the optimal quantity of data to be sold ($d_5^{DB,E}$ and $d_6^{DB,E}$), resulting in a generally increasing trend as n increases (Figure 2.14,yellow & green).

When the number of firms is sufficiently close to the number of available spokes ($n \in [n_7^{DB,E*}(v), N-1]$), the optimal quantity of data to be sold falls within a d region where the uninformed firm is compelled to lower its price below $v-1$, thereby reducing rent extraction from captive consumers (2.33). Indeed, it reacts with equal intensity to the aggressive approach pursued by its informed competitors. Consequently, the threat of rejecting the DB's offer loses significance once again in comparison to the declining behavior of Inside Option rents, causing the DB's equilibrium offer ($d_7^{DB,E}$) to decrease with n (Figure 2.14,red).



(a) DB's Strategy Definition

(b) DB's Strategy for $v = \frac{9}{5}$

Figure 2.14. DB's Optimal Strategy in the Region of Competitive Sub-Markets Prevalence and in the exclusive information scenario for $N = 100$ and $v \in [\frac{42}{25}, \frac{9}{5}]$

When $d = \frac{1}{2}$, however, the I.C.E. caused by the unidentified consumers in duopolistic sub-markets suddenly ceases to exist. Indeed, each informed firm, regardless of the presence of unidentified consumers shared with the uninformed competitor, independently opts to raise its basic price solely to maximize non-discriminatory monopolistic profit, increasing its Inside Option rents (Figure 2.11,c & d). Concurrently, the uninformed firm foresees that its informed competitors will swiftly elevate their basic price and responds by adjusting its strategy accordingly, resulting in an increase in its Outside Option profits (Figure 2.12,c & d). Point 2) of Observation 2 shows that when $d = \frac{1}{2}$, the gain obtained in the Inside Option scenario for extracting additional profit from captive unidentified consumers far exceeds the gain obtained by an uninformed firm in the outside option due to the end of price wars among informed firms (Figure 2.13,a). This happens mainly because when d reaches $\frac{1}{2}$, the intensified competition effect does not affect any more monopolistic rents. As a result, Observation 2 proves that, regardless of the number of filled Spokes and the utility value, the DB will mitigate competition by selling a quantity of data greater or equal than $\frac{1}{2}$.

2.3.3 DB's Optimum Strategy

Based on Observations 1 and 2 discussed in the previous sections, it is evident that the DB's preference is to sell non-exclusive information, regardless of whether the market is predominantly monopolistic or competitive. By employing this approach, the DB can steer clear of the price war that ensues among informed competitors in cases where the duopolistic markets are not fully covered by data. Moreover, this strategy ensures that rent extraction from monopolistic sub-markets remains unneutralized. The subsequent proposition presents the equilibrium strategy of the DB, encompassing all values of $n \in [2, N-1]$ and $v \in [\frac{3}{2},2]$. As the Outside Option scenario and the DB's strategies are discussed and demonstrated in Section 2.3.2 while assuming $N = 100$, we will apply the same assumption to the final results presented in this Section.

Proposition 6. [Extension of Observations 1 and 2] *Fixing $N = 100$, if the DB can offer the entire data set ($d^{DB} \in [0,1]$), independently from the prevalence of monopolistic or competitive markets, its optimal strategy consists of selling almost all data. Specifically, $d^{DB*}(n, v) \in [\frac{17}{20},1]$ and:*

$$d^{DB*} = \begin{cases} d^{\pi_4^{W*}}(2.26) \;||\; n \in [2, n_4^{DB*}(v)) & ||\; d^{\pi_4^{W*}} \in (d_8^O,1) \\ d_8^O(n,v) \;||\; n \in [n_4^{DB*}(v), N-1] \;|| & \end{cases}$$

Moreover:

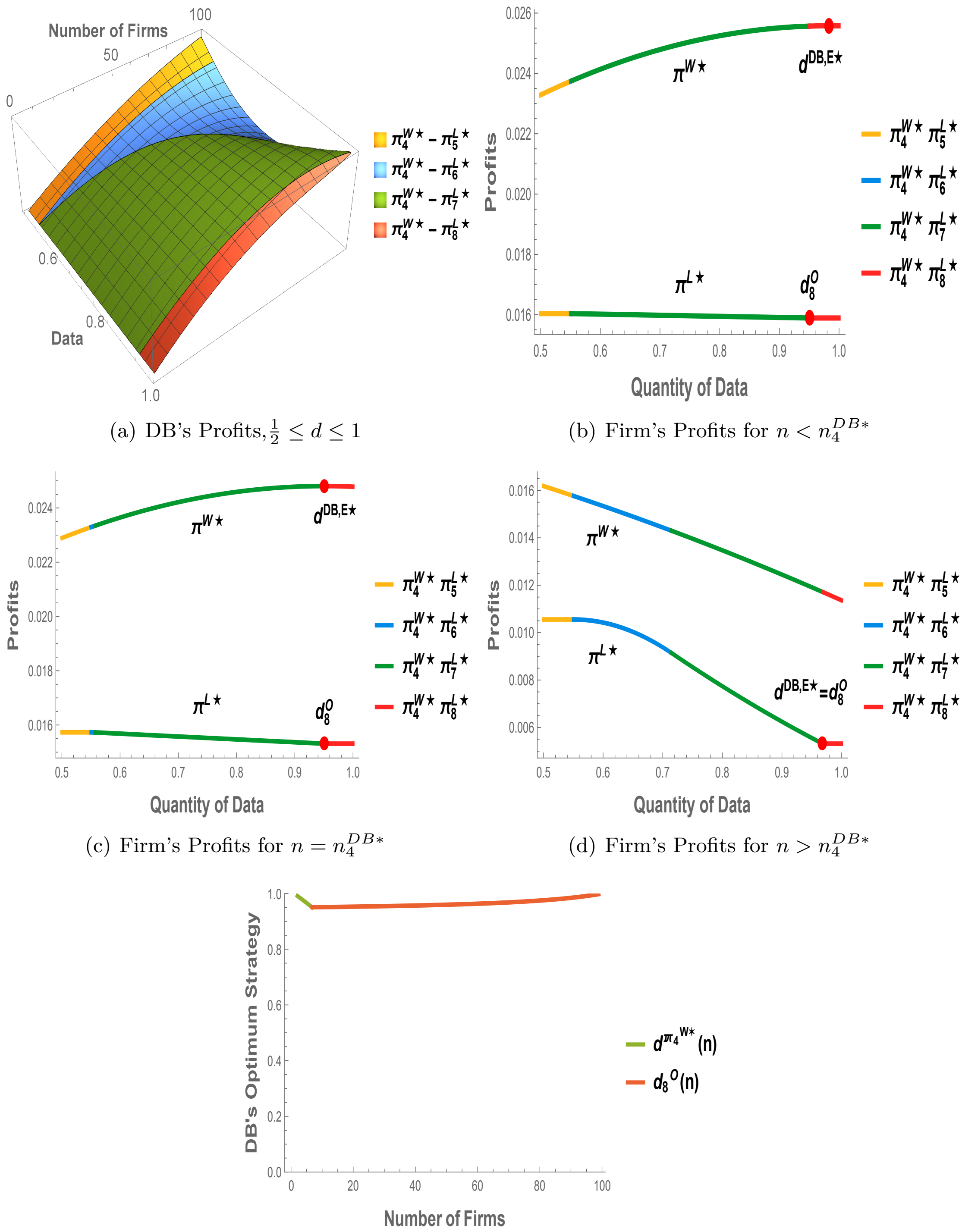
1) *The DB's profits evaluated at $d = d_8^O(n, v)$ are greater than the DB's profits evaluated at $d = \frac{1}{2}$, $\forall\ (n, v) \in [2,99] \times [\frac{3}{2},2]$;*

2) *(2.26) is always decreasing in n;*

3) *$d_8^O(n, v)$ is always increasing in n.*

In the previous section, we demonstrated that the optimal amount of data to be sold lies within the range of $d^{DB*} \in [\frac{1}{2},1]$. Therefore, our attention will now be directed toward examining the profits achieved in the context of non-exclusive information. Taking into account the insights provided by Lemma 1.2 and 4.1, it becomes evident that in situations where segments overlap, the profits of the Inside Option, referred to as π_4^{W*}, experience an increase until d reaches $d^{\pi_4^{W*}}$ within the range of $n \in [2, n^{\pi_4,I}(v))$ and decreases for $d \in [\frac{1}{2},1]$ otherwise. Moreover, it is important to note that $n_2^I(v) > n^{\pi_4,I}(v)$. Consequently, it can be concluded that, in the Inside Option scenario, for the majority of values of n, the Bertrand Competition Effect prevails over the Surplus Extraction Effect arising from monopolistic markets. Furthermore, in the case of non-exclusive information, the DB's strategy is primarily influenced by the perceived threat posed by the Outside Option (Figure 2.15,a). Indeed, point 1) of Proposition 6 highlights that the DB consistently seeks to minimize the profits of the uninformed firm in the Outside Option, irrespective of the values assumed by (n, v). Indeed, taking into account Proposition 3, $d_8^O(n, v)$ represents a threshold value beyond which additional data offer no competitive advantage to the informed firms in the Outside Option scenario, resulting in constant rents for the uninformed firm.

Specifically, for values of n within the range $[2, n_4^{DB*})$, where $n_4^{DB*} < n^{\pi_4, I}(v)$, the DB successfully minimizes the rents of the uninformed firm while maximizing the profits of the Inside Option (Figure 2.15,b). This is achieved by setting d up to $d^{\pi_4^{W*}}$, where π_4^{W*} is maximized thanks to the Surplus Extraction Effect from discriminated consumers in monopolistic sub-markets. When $n = n_4^{DB*}$, the value of d that maximizes the Inside Option profits ($d^{\pi_4^{W*}}$) coincides with $d_8^O(n, v)$ (Figure 2.15,c). As n exceeds n_4^{DB*}, the DB finds it advantageous to fully leverage the threat of the Outside Option, even though profits from the Inside Option generally decrease with increasing d (Figure 2.15,d). Moreover, as n increases, the competitive sub-markets and the corresponding threat of being uninformed gain more relevance for the firm rejecting the DB's offer. Consequently, the DB's optimal strategy approaches $d^{DB*} \approx 1$ as n approaches $N - 1$ (Figure 2.15,e). Hence, the DB strategically leaves unidentified only a small portion of the market segment, where firms are not compelled to base their pricing on marginal cost, but it always finds it advantageous to sell nearly all of its data to all firms.



(a) DB's Profits, $\frac{1}{2} \leq d \leq 1$

(b) Firm's Profits for $n < n_4^{DB*}$

(c) Firm's Profits for $n = n_4^{DB*}$

(d) Firm's Profits for $n > n_4^{DB*}$

(e) DB's optimal strategy, $2 \leq n \leq N-1$

Figure 2.15. DB's Profits and Optimal Strategies for $N = 100$, $v = \frac{19}{10}$

2.4 Conclusions

With the steady growth of online services, DBs have become central players in the digital economy. Their ability to extract valuable information from consumers' data allows them to influence competition in retail markets, with important welfare implications. This study extends the growing strand of literature on the competitive effects of DBs on downstream markets by modeling an oligopoly market where all firms serve a subset of captive consumers. The marginal value of data in the Spokes Model is two-fold: on the one hand, more data enable firms to extract total surplus from a greater turf of captive consumers, while on the other hand, they foster intensified competition within duopolistic segments. Examining a market where products hold limited utility and some captive consumers choose not to buy, our findings indicate that the Data Broker reliably accomplishes total market coverage. This involves selling equally portioned segments, encompassing almost all data, to all companies involved.

In the Hotelling framework, as demonstrated by Bounie et al. [2021], the DB consistently sells only exclusive information, reserving an indiscriminate market segment to mitigate competition for unidentified consumers. Similarly, in a comparable scenario where the downstream market follows the Salop model and the number of firms is sufficiently high, Abrardi et al. [2022] discovers that the DB chooses to leave a territory of shared unidentified consumers. This strategy arises because, before the data becomes ineffective in influencing market outcomes, the decline in the Inside Option profits is not adequately compensated by a proportionate decrease in profits for the uninformed competitor. Moreover, the DB can temper competition only selling exclusive information to all firms.

However, in the Spokes Model, where monopolistic markets exist, even in the presence of low market concentration the DB mitigates competition by selling a low quantity of non-exclusive information. This effectively eliminates the price war that occurs for unidentified consumers, enabling each informed firm to fully exploit the surplus from non-discriminatory captive consumers. Moreover, in our framework, even a single empty spoke, where informed firms can fully exploit surplus, negatively impacts the uninformed firms to such an extent that the DB prefers to let the Bertrand Competition Effect affects the duopolistic component of Inside Option profits to maximize the threat of being uninformed.

While this work did not undertake a thorough analytical welfare analysis, it is readily apparent that when the count of competitive sub-markets is notably fewer than monopolistic sub-markets, selling data will considerably disadvantage consumers. Moreover, due to the fact that captive consumers, irrespective of their significance, are almost identified by all firms, they will be always negatively affected by the surplus extraction effect of data within monopolistic segments. Regarding policy formulation, this study proposes that the mere existence of a limited sub-market of captive consumers motivates the DB to sell nearly all available data. The DB's optimal strategy promotes heightened competition and guarantees comprehensive market reach. Consequently, if the prevalence of consumers within competitive sub-markets is significant enough, there might be no need for regulatory interventions to ensure the pro-competitive effect of data.

Appendix A

Proofs and Definitions

Proof of Proposition 1. For $d = 0$, assuming $v \in [\frac{3}{2},2]$, the prices described in Region III of Remark 1 hold. For each region of $d(n, v, N)$ described in Proposition 1, assuming that all firms have a quantity d of data, we construct a symmetric equilibrium where the equilibrium basic prices satisfy a unique property that can hold only in the assumed region of parameter values.

- $\mathbf{d} \in [\mathbf{0}, \mathbf{d_1^I}(\mathbf{n}, \mathbf{v}, \mathbf{N}))$

When $d \in [0, \frac{1}{2})$, if firm j accepts the DB offer, it will set the basic price p_{Bj}^W to face $n-1$ informed competitors with d data over each duopolistic market and to serve captive consumers with a preference for one of the $N-n$ varieties.

Suppose that the symmetric equilibrium prices satisfy:

$$\begin{cases} p_1^{W*} > v - 1 \\ p_1^{W*} < v - \frac{1}{2} \end{cases}$$

The non-discriminatory demand facing firm j, when p_{Bj}^W is in the neighborhood of p_B^{W*}, can be expressed as:

$$q_j^W = \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_1^{W*} - p_{Bj}^W}{2} - d\right) + \frac{N-n}{N-1}\frac{2}{N}(v - p_{Bj}^W - d)$$

Firm's j first-order condition for profit maximization is:

$$q_j^W - \frac{n-1}{N-1}\frac{1}{N}p_{Bj}^W - \frac{N-n}{N-1}\frac{2}{N}p_{Bj}^W = 0$$

Therefore, at a symmetric equilibrium:

$$p_1^{W*} = \frac{2v(N-n) - 2d(N-1) + (n-1)}{4N - 3n - 1}$$

The requirement that $p_1^{W*} > v - 1$ is satisfied if and only if:

$$d < \frac{(2-v)(2N-n-1)}{2(N-1)}$$

Denoting the above quantity as $d_1^I(n,v,N)$, it can be shown that:

$$d_1^I(n,v,N) \;\in\; [0,\frac{1}{2}),\; \forall\; (n,v,N) \in [2,N-1] \times [\frac{3}{2},2] \times [2,+\infty)$$

Moreover, it is easy to show that $d_1^I(n,v,N)$ is decreasing in n and v.

For $p_{Bj}^W \geq p_1^{W*} + 1 - 2d$ the demand function q_j^W undergoes the following change:

$$q_j^W = \frac{N-n}{N-1}\frac{2}{N}(v - p_{Bj}^W - d)$$

and the corresponding informed firm's non-discriminatory profits are increasing for $p_{Bj}^W \leq \frac{v-d}{2}$, when $v \leq 2-d$. It is worth noticing that $p_1^{W*} + 1 - 2d > \frac{v-d}{2}$, for $d \in [0, d_1^I)$, $n \in [2, N-1]$ and $v \in [\frac{3}{2},2]$. Hence, to verify that p_1^{W*} is also globally optimal, notice that, since $v - 1 < p_1^{W*} < v - \frac{1}{2}$, it suffices if any deviation to any $p_{Bj}^W < p_1^{W*}$ cannot be profitable. But since the second-order condition is satisfied for $p_{Bj}^W < p_1^{W*}$, no global deviation can be profitable. Thus p_1^{W*} is a symmetric equilibrium when $d \in [0, d_1^I(n,v,N))$ and only in this region.

- $\mathbf{d \in [d_1^I(n,v,N), \frac{1}{2})}$

When $d \in [d_1^I(n,v,N), \frac{1}{2})$, suppose that a symmetric equilibrium price satisfies:

$$p_2^{W*} = v - 1$$

The non-discriminatory demand facing firm j, when p_{Bj}^W is in the neighborhood of p_B^{W*}, can be expressed as:

$$q_j^W = \begin{cases} \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_2^{W*} - p_{Bj}^W}{2} - d\right) + \frac{N-n}{N-1}\frac{2}{N}(v - p_{Bj}^W - d) & \text{if } p_{Bj}^W > v-1 \\ \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_2^{W*} - p_{Bj}^W}{2} - d\right) + \frac{N-n}{N-1}\frac{2}{N}(1-d) & \text{if } p_{Bj}^W < v-1 \end{cases}$$

In order for $p_B^{W*} = v - 1$ to be an equilibrium, a slight increase of p_{Bj}^W at p_2^{W*}, should not increase profits:

$$q_j^W + p_{Bj}^W \frac{\partial q_j^W}{\partial p_{Bj}^W} = d(1-N) + \frac{1}{2}(2-v)(2N-n-1) \leq 0$$

which holds if $d \geq d_1^I(n,v,N)$. Also, a slight decrease of p_{Bj}^W at p_2^{W*} should not decrease profit:

$$q_j^W + p_{Bj}^W \frac{\partial q_j^W}{\partial p_{Bj}^W} = d(1-N) - \frac{1}{2}(n-1)v + N - 1 \geq 0$$

which holds if and only if:

$$d \leq d_2^I(n,v,N) = \frac{v(1-n) + 2N - 2}{2N-2}$$

By imposing that $d_2^I(n,v,N) > \frac{1}{2}$, as a function of n, we obtain:

$$n < n_2^I(v,N) = \lceil \frac{N+v-1}{v} \rceil^+$$

where $n_2^I(v,N)$ defines the boundary for the region of Monopolistic Sub-Markets Prevalence assumed in Proposition 1.

Recalling from the proof of Proposition 1 that for $p_{Bj}^W \geq p_2^{W*} + 1 - 2d$ the demand function q_j^W consists only its the monopolistic component, it is easy to show that $p_2^{W*} + 1 - 2d > \frac{v-d}{2}$, for $d \in [d_1^I, \frac{1}{2})$, $n \in [2, N-1]$ and $v \in [\frac{3}{2}, 2]$. Hence, p_2^{W*} is also globally optimal, if firm j cannot benefit from any deviation to $p_{Bj}^W < v-1$. But since the second-order condition is satisfied for $p_{Bj}^W < v-1$ (any kink of the profit function makes it more concave), no global deviation can be profitable. In other words, when $d \in [0, \frac{1}{2})$ and $n < n_2^I(v,N)$, firms do not find it convenient to lower prices under $v-1$: extracting rent from non-discriminatory monopolistic market is more relevant than trying to poach competitors' consumers.

- $\mathbf{d} \in [\mathbf{\frac{1}{2}}, \mathbf{1}]$

When $d \in [\frac{1}{2}, 1]$, if firm j accepts the DB offer, it will set the basic price p_{Bj}^W to serve only captive consumers with a preference for one of the $N-n$ varieties.

Firm j's non-discriminatory demand,in this region of d, can be expressed as:

$$\begin{cases} v - p_{Bj}^W - d & \text{if } v - p_{Bj}^W < 1 \\ 1 - d & \text{if } v - p_{Bj}^W > 1 \end{cases}$$

Assuming that $p_{Bj}^W > v-1$ and computing the first-order condition, we obtain:

$$p_{Bj}^W = \frac{v-d}{2}$$

Observing that:

$$\frac{v-d}{2} < v-1 \quad \forall\, (d,v) \in [\frac{1}{2}, 1] \times [\frac{3}{2}, 2]$$

the equilibrium prices in this region will be:

$$p_4^{W*} = v-1$$

As $p_4^{W*} = v-1$, the equilibrium prices and their region of definition remain unaffected by n. □

Proof of Lemma 1.1. Point 2) of the Lemma is easily demonstrated by observing the analytic expression of p_1^{W*}. For what concerns point 3), we compute the first-order derivative of $p_1^{W*}(d,n,v,N)$ with respect to n, obtaining the following expression:

$$\frac{\partial p_1^{W*}(d,n,v,N)}{\partial n} = \frac{2(N-1)(2-v-3d)}{(3n-4N+1)^2}$$

which is greater than 0 when:

$$d \leq \frac{2-v}{3}$$

and lower than 0 otherwise. □

Proof of Lemma 1.2. Recalling that the non-discriminatory demand function where basic prices hold is described in equation (2.17), firm's profits in the Inside Option scenario can be defined as:

$$\pi^{W*} = \begin{cases} q^{W*} p_B^{W*} + \int_0^d p_B^{W*} + 1 - 2x \; dx + \int_0^d v - x \; dx & \text{if } d < \frac{1}{2} \\ q^{W*} p_B^{W*} + \int_0^{1-d} p_B^{W*} + 1 - 2x \; dx + \int_{1-d}^{\frac{1}{2}} 1 - 2x \; dx + \int_0^d v - x \; dx & \text{if } d \geq \frac{1}{2} \end{cases}$$

$\boldsymbol{\pi_1^{W*}}$ **is increasing in d**: the informed firms' profit function π_1^{W*} is concave in d within its range of definition ($d \in [0, d_1^I)$). As a consequence, it is sufficient to show that the first-order derivative of the profit is greater or equal to zero when d reaches the upper bound of its valid range. By evaluating the derivative of the profit function at $d = d_1^I(n, v, N)$, we obtain:

$$\frac{\partial \pi_1^{W*}}{\partial d}(d = d_1^I) = \frac{a+b}{(N-1)^2 N(3n - 4N + 1)}$$

$$a = v\left(-3n^3 + 3(3(n-2)n - 5)N + 9n - 8N^3 + 24N^2 + 2\right)$$

$$b = (n - 2N + 1)\left(3n^2 - 3n(N+1) + N(7 - 2N) - 2\right)$$

By imposing that $\frac{\partial \pi_1^{W*}}{\partial d} \geq 0$ as a function of v, we obtain the following upper bound:

$$v \leq \frac{-18n^2N + 6n^3 + 8nN^2 + 20nN - 10n + 8N^3 - 32N^2 + 22N - 4}{-9n^2N + 3n^3 + 18nN - 9n + 8N^3 - 24N^2 + 15N - 2} \tag{A.1}$$

which is always increasing in n for $2 \leq n \leq N$.

The lower bound on v beyond which the basic price p_1^{W*} holds, can be obtained by expressing $n_2^I(v, N)$ as a function of v:

$$v \geq \frac{N-1}{n-1} \tag{A.2}$$

which is always decreasing in n. By imposing that (A.2) is lower than $\frac{3}{2}$ as a function n, the following threshold value arise:

$$\frac{3}{2} = \frac{N-1}{n-1} \iff n = \frac{1}{3}(2N+1)$$

Equation (A.1) evaluated at $n = \frac{1}{3}(2N+1)$, gives $\frac{16}{11}$, which is lower than $\frac{3}{2}$. This means that π_1^{W*} can be decreasing in d only when $n \geq n_2^I(v, N)$.

$\boldsymbol{\pi_2^{W*}}$ **is increasing in d**: when $d \in [d_1^I(n, v, N), \frac{1}{2})$, each firm sets the basic price equal to $v - 1$ granting full market coverage. The analytic expression of profits in this region

(π_2^{W*}) is defined in equation (2.24). It is easy to show that π_2^{W*} is concave in d. Ensuring that the first-order derivative is greater or equal than zero, we obtain:

$$\frac{\partial \pi_2^{W*}}{\partial d} \geq 0 \iff d \leq \frac{N-1}{n+N-2}$$

By imposing that the above condition on d is greater or equal than $\frac{1}{2}$, follows that:

$$\frac{N-1}{n+N-2} \geq \frac{1}{2} \iff 2-N < n \leq N$$

$\pi_4^{\mathbf{W}*}$ **is both increasing and decreasing in d**: when $d \geq \frac{1}{2}$, the equilibrium prices do not depend on (d, n) and firm's profit functions are moved by personalized prices: on one hand there is the S.E.E. in the monopolistic markets, and on the other there is the I.C.E. in duopolistic markets where Bertrand competition takes place. The analytic expression of profits in this region (π_4^{W*}) is defined in equation (2.25). It is easy to show that π_4^{W*} is concave in d. Ensuring that the first-order derivative is equal to zero, we obtain:

$$\frac{\partial \pi_4^{W*}}{\partial d} = 0 \iff d^{\pi_4^{W*}} = \frac{nv-N-v+1}{n-N}$$

Moreover, by by imposing $d^{\pi_4^{W*}} \leq \frac{1}{2}$, as a function of n, we obtain:

$$n \geq n_{\pi 4}^{I} = \frac{N+v-1}{v}$$

It easy to show that $n_{\pi 4}^{I}(v, N) < n_2^{I}(v, N)$, $\forall\, v \in [\frac{3}{2}$,2]. Therefore, π_4^{W*} is increasing in d when $(d, n) \in [\frac{1}{2}, d^{\pi_4^{W*}}] \times [2, n_{\pi 4}^{I})$.

$\square$

Proof of Proposition 2. For $d = 0$, assuming $v \in [\frac{3}{2}$,2], the prices described in Region III of Remark 1 hold. For each region of $d(n, v, N)$ described in Proposition 2, assuming that all firms, except one, have a quantity d of data, we construct a symmetric equilibrium where the equilibrium basic prices satisfy a unique property that can hold only in the assumed region of parameter values.

- $\mathbf{d} \in [\mathbf{0}, \mathbf{d_1^O}(\mathbf{n}, \mathbf{v}, \mathbf{N}))$

When $d \in [0, \frac{1}{2})$, if firm i rejects the offer and all other firms $j \neq i$ hold a quantity d of data, it will face $n-1$ informed competitors and will serve its monopolistic market at a basic price p_{Bi}^{L}. Each firm j will face $n-2$ informed competitors, the uninformed firm i, and will serve captive consumers at a basic price p_{Bj}^{WL}.

Suppose that a symmetric equilibrium prices satisfies:

$$\begin{cases} p_1^{WL*} > v-1 \\ p_1^{WL*} < p_1^{L*} \\ p_1^{L*} > v-1 \\ p_1^{L*} < v-\frac{1}{2} \end{cases}$$

The non-discriminatory demand facing each firm j, when p_{Bj}^{WL} is in the neighborhood of p_1^{WL*} and p_1^{L*}, can be expressed as:

$$\begin{aligned} q_j^{WL} =& \frac{n-2}{N-1}\frac{2}{N}\left(\frac{1}{2}+\frac{p_1^{WL*}-p_{Bj}^{WL}}{2}-d\right)+\frac{1}{N-1}\frac{2}{N}\left(\frac{1}{2}+\frac{p_1^{L*}-p_{Bj}^{WL}}{2}-d\right)+ \\ &+\frac{N-n}{N-1}\frac{2}{N}(v-p_{Bj}^{WL}-d) \end{aligned}$$

The uniformed firm i's demand function, when p_{Bi}^{L} is in the neighborhood of p_1^{WL*} , can be expressed as:

$$q_i^L = \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2}+\frac{p_1^{WL*}-p_{Bi}^{L}}{2}\right)+\frac{N-n}{N-1}\frac{2}{N}(v-p_{Bi}^{L})$$

By imposing first-order conditions on both firms' profit functions, we obtain the following systems of equations:

$$\begin{cases} q_j^{WL}-\dfrac{n-1}{N-1}\dfrac{1}{N}p_{Bj}^{WL}-\dfrac{N-n}{N-1}\dfrac{2}{N}p_{Bj}^{WL}=0 \\ q_i^{L}-\dfrac{n-1}{N-1}\dfrac{1}{N}p_{Bi}^{L}-\dfrac{N-n}{N-1}\dfrac{2}{N}p_{Bi}^{L}=0 \end{cases}$$

Therefore, at a symmetric equilibrium:

$$\begin{cases} p_1^{L*}=\dfrac{2d(n(-N)+n+N-1)+(2n-4N+1)(n(2v-1)-2Nv+1)}{(2n-4N+1)(3n-4N+1)} \\ p_1^{WL*}=\dfrac{4d(N-1)(n-2N+1)+(2n-4N+1)(n(2v-1)-2Nv+1)}{(2n-4N+1)(3n-4N+1)} \end{cases}$$

The difference between p_1^{L*} and p_1^{WL*} is:

$$p_1^{L*}-p_1^{WL*}=\frac{2d(N-1)}{4N-2n-1}$$

which is bounded above by 1 and below by 0 in its proper region of definition. Therefore, each informed firm j establishes a base price lower than the price set by firm i, thereby capturing a portion of the duopolistic demand from the uninformed firm on spoke l_i.

Moreover, both basic prices are greater than $v-1$ if and only if p_1^{WL*} satisfies this criterion. By imposing the condition $p_1^{WL*}>v-1$, as a function of d, we obtain:

$$d<\frac{(2-v)(4N-2n-1)}{4(N-1)}$$

Denoting the above quantity as $d_1^O(n,v,N)$, it can be shown that:

$$d_1^O(n,v,N)\ \in\ [0,\frac{1}{2}),\ \forall\ (n,v,N)\in[2,N-1]\times[\frac{3}{2},2]\times[2,+\infty)$$

Moreover, it is easy to show that $d_1^O(n,v,N)$ is decreasing in n and v.

To verify that p_1^{L*} is also globally optimal, notice that, since $v-1 < p_1^{L*} < v - \frac{1}{2}$, it suffices if any deviation to any $p_{Bi}^{L} < p_1^{WL*}$ cannot be profitable. But since the second-order condition is satisfied for $p_{Bi}^{L} < p_1^{WL*}$, no global downward deviation can be profitable.

The same holds true for p_1^{WL*} but, in the case of informed firms, we have to check for any incentive to deviate upwards. Indeed, for $p_{Bj}^{WL} \geq p_1^{WL*} + 1 - 2d$, the demand function q_j^{WL} undergoes the following change:

$$q_j^{WL} = \frac{1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_1^{L*} - p_{Bj}^{WL}}{2} - d\right) + \frac{N-n}{N-1}\frac{2}{N}(v - p_{Bj}^{WL} - d) \tag{A.3}$$

and the corresponding profits' first-order derivative evaluated at $p_{Bj}^{WL} = p_1^{WL*} + 1 - 2d$ is lower than 0 for $d \in [0, d_1^O)$. Moreover, it is worth noticing that $p_1^{WL*} + 1 - 2d > \frac{v-d}{2}$, for $d \in [0, d_1^O)$, $n \in [2, N-1]$ and $v \in [\frac{3}{2}, 2]$. Thus $\{p_1^{L*}, p_1^{WL*}\}$ is a symmetric equilibrium when $d \in [0, d_1^O(n, v, N))$ and only in this region.

- $\mathbf{d \in [d_1^O(n, v, N), \frac{1}{2})}$

When $d \in [d_1^O(n, v, N), \frac{1}{2})$, suppose that a symmetric equilibrium prices satisfies:

$$\begin{cases} p_2^{L*} = \frac{v}{2} \\ p_2^{WL*} = v - 1 \end{cases}$$

For a generic uniformed firm i, $p_2^{L*} = \frac{v}{2}$ is the best response to $p_2^{WL*} = v-1$. Furthermore, $p_2^{L*} > v-1$, since $\frac{v}{2} > v-1$ for $v \in [\frac{3}{2}, 2]$.

The non-discriminatory demand facing firm j, when p_{Bj}^{WL} is in the neighborhood of p_2^{L*} and p_2^{WL*}, can be expressed as:

$$q_j^{WL} = \begin{cases} \frac{n-2}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_2^{WL*} - p_{Bj}^{WL}}{2} - d\right) + \frac{1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_2^{L*} - p_{Bj}^{WL}}{2} - d\right) + \frac{N-n}{N-1}\frac{2}{N}(v - p_{Bj}^{WL} - d) \\ \text{if } p_{Bj}^{WL} > v-1 \\ \frac{n-2}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_2^{WL*} - p_{Bj}^{WL}}{2} - d\right) + \frac{1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_2^{L*} - p_{Bj}^{WL}}{2} - d\right) + \frac{N-n}{N-1}\frac{2}{N}(1-d) \\ \text{if } p_{Bj}^{WL} < v-1 \end{cases}$$

In order for $p_2^{WL*} = v-1$ to be an equilibrium, a slight increase of p_{Bj}^{WL} at p_2^{WL*}, should not increase profits:

$$q_j^{WL} + p_{Bj}^{WL}\frac{\partial q_j^{WL}}{\partial p_{Bj}^{WL}} = -Nd + d + \frac{1}{4}(v-2)(2n - 4N + 1) \leq 0$$

which holds if $d \geq d_1^O(n, v, N)$. Also, a slight decrease of p_{Bj}^{WL} at p_2^{WL*} should not decrease profit:

$$q_j^{WL} + p_{Bj}^{W}\frac{\partial q_j^{WL}}{\partial p_{Bj}^{W}} = \frac{1}{4}(-4d(n-1) - 2nv + 4N + v - 2) \geq 0$$

which holds if and only if:

$$d \leq d_2^O(n, v, N) = \frac{-2nv + 4N + v - 2}{4N - 4}$$

By imposing that $d_2^O(n,v,N) > \frac{1}{2}$ as a function of n, we obtain:

$$n < n_2^O(v,N) = \lceil \frac{4N+v}{2v+2} \rceil^+$$

where $n_2^I(v,N)$ defines the boundary for the region of Monopolistic Sub-Markets Prevalence assumed in Proposition 1.

Recalling that for $p_{Bj}^{WL} \geq p_2^{WL*} + 1 - 2d$ the informed firm's demand function q_j^{WL} undergoes the change defined in (A.3), it is easy to show that the corresponding first-order derivative of informed firm's profits evaluated at $p_{Bj}^{WL} = p_2^{WL*} + 1 - 2d$ is lower than 0, for $d \in [d_2^I, \frac{1}{2})$, $n \in [2, N-1]$ and $v \in [\frac{3}{2}, 2]$. Hence, p_2^{WL*} is also globally optimal, if firm j cannot benefit from any deviation to $p_{Bj}^{WL} < v - 1$. But since the second-order condition is satisfied for $p_{Bj}^{WL} < v - 1$ (any kink of the profit function makes it more concave), no global deviation can be profitable. In other words, when $d \in [0, \frac{1}{2})$ and $n < n_2^O(v,N)$, firms do not find it convenient to lower prices under $v - 1$: extracting rent from the non-discriminatory monopolistic market is more relevant than trying to poach competitors' consumers.

□

Proof of Lemma 2.1. In order to prove point 1) of this Lemma, we compute the first-order derivative of p_1^{L*} with respect to d:

$$\frac{\partial p_1^{L*}}{\partial d} = \frac{2(n(-N) + n + N - 1)}{(2n - 4N + 1)(3n - 4N + 1)}$$

which is lower than 0, $\forall\ n \in [2, N-1]$.

By differentiating p_1^{L*} with respect to n, we calculate the following first-order derivative:

$$\frac{\partial p_1^{L*}}{\partial n} = \frac{2(N-1)\left(2d(3(n-2)n + 2N(7-4N) - 3) - (v-2)(2n-4N+1)^2\right)}{(2n-4N+1)^2(3n-4N+1)^2}$$

which is lower than 0 when:

$$d \leq d_{p_1}^{L*} = \frac{(v-2)(2n-4N+1)^2}{6(n-2)n + 4N(7-4N) - 6}$$

Moreover, $d_{p_1}^{L*}(n,v,N)$ is decreasing in n, $\forall\ (n,v) \in [2, N-1] \times [\frac{3}{2}, 2]$, and is lower than $d_1^O(n,v,N)$, $\forall\ (n,v) \in [2, N-1] \times [\frac{3}{2}, 2]$.

Recalling that the uninformed firm's demand function is described in equation (2.21), firm's profits in the outside option scenario can be defined as:

$$\pi^{L*} = q^{L*} p_B^{L*}$$

For $d \in [0, d_1^O)$, basic price of the uninformed firm boils down to p_1^{L*} and the corresponding expression for profits is:

$$\pi_1^{L*} = \frac{(2N-n-1)(2d(+n+N-1-nN) + (4N-2n-1)(n(2v-1) - 2Nv + 1))^2}{(N-1)N(4N-2n-1)^2(4N-3n-1)^2}$$

which is decreasing in d in its proper region of definition.

For $d \in [d_1^O, \frac{1}{2})$, basic price of the uninformed firm boils down to p_2^{L*} and the corresponding expression for profits is:

$$\pi_2^{L*} = \frac{v^2(2N - n - 1)}{4(N-1)N}$$

which is independent of d. □

Proof of Proposition 3. When $d \in [\frac{1}{2}$,1] and a generic firm i rejects the DB's offer, it will faces $n-1$ informed competitor at a basic price p_{Bi}^L. Each firm $j \neq i$, at the same time, will set its basic price p_{Bj}^{WL} to serve only its non-discriminatory captive consumers. We assume that each informed firm sets its basic price p_B^{WL*} equal to $v-1$ and we compute the uninformed firm's best response for $d \in [\frac{1}{2}$,1]. The proof is concluded showing that all informed firms have no incentive to deviate from $v-1$.

- $\mathbf{d \in [\frac{1}{2}, d_5^O(n, v, N))}$

Suppose that a symmetric equilibrium prices satisfies:

$$\begin{cases} p_6^{L*} = \frac{v}{2} \\ p_B^{WL*} = v - 1 \end{cases}$$

Recalling that, as shown in the proof of Proposition 2, for a generic uniformed firm i, $p_6^{L*} = \frac{v}{2}$ is the best response to $p_B^{WL*} = v-1$ and $p_6^{L*} > v-1$ for $v \in [\frac{3}{2}$,2], it is easy to show that this equilibrium holds as long as the indifferent consumer remains in the non-discriminatory market. In other words, for every v, there exists a value of d beyond which, given the prices described above, the informed firm can no longer attract unidentified consumers away from the data-less firm. This boundary can be obtained by imposing:

$$\frac{1}{2} + \frac{p_6^{L*} - p_B^{WL*}}{2} - d \geq 0$$

which gives:

$$d \leq d_5^O(v, n, N) = \frac{4 - v}{4}$$

- $\mathbf{d \in [d_5^O(n, v, N)), d_6^O(n, v, N))]}$

When $d \in [d_5^O(v, n, N)$,1], assuming that each informed firm has no incentive to deviate from $p_B^{WL*} = v-1$, the uninformed firm i will face the following demand function:

$$q_i^L = \frac{n-1}{N-1}\frac{2}{N}(1-d) + \frac{N-n}{N-1}\frac{2}{N}(v - p_{Bi}^L) \tag{A.4}$$

The first component of (A.4) represent firm's i (inelastic) competitive demand function which is equal to $\frac{n-1}{N-1}\frac{2}{N}(1-d)$ when:

$$\frac{1}{2} + \frac{p_B^{WL*} - p_{Bi}^L}{2} \geq 1 - d$$

By imposing the above condition on p^L_{Bi}, we obtain:

$$p^L_{Bi} \leq 2d + v - 1 \tag{A.5}$$

Firm's i first-order condition for profit maximization when the demand function is described by (A.4) is the following:

$$q^L_i - \frac{N-n}{N-1}\frac{2}{N}p^L_{Bi} = 0$$

which gives:

$$p^L_{Bi} = \frac{1}{2}\left(\frac{(1-d)(n-1)}{N-n} + v\right) \tag{A.6}$$

When d is such that (A.5) is not satisfied we will have a corner solution and the equilibrium price will be $p^{L*}_7 = 2d + v - 1$. Indeed, by imposing (A.5) as a function of d we obtain:

$$d \leq d^O_6(n, v, N) = 1 + \frac{v(n-N)}{4N - 3n - 1}$$

- $\mathbf{d} \in [\mathbf{d^O_6}(\mathbf{n}, \mathbf{v}, \mathbf{N})), \mathbf{d^O_7}(\mathbf{n}, \mathbf{v}, \mathbf{N}))]$

Hence, if $d > d^O_6(n, v, N)$, firm i would not find it advantageous to increase the uniform price to $2d + v - 1$, as doing so would result in an excessive loss of captive consumers: in this region of d, the equilibrium price will be $p^{L*}_8 =$ (A.6).

Furthermore, when $d > d^O_6(n, v, N)$, each firm j customizes prices for consumers within the duopolistic market it shares with the uninformed firm, employing the following approach:

$$p^C_j(x) = p^L_{Bi} + 1 - 2x$$

Given that the above expression is decreasing in x, to ensure that $p^C_j(x) \geq 0$ it is sufficient to show that:

$$p^L_{Bi} + 1 - 2d \geq 0$$

The above condition is satisfied when:

$$d \leq d^O_7(n, v, N) = \frac{N(v+2) - n(v+1) - 1}{4N - 3n - 1}$$

Therefore, for $d \geq d^O_7(n, v, N)$, data do not allow informed firms to conquer more consumers on spoke l_i and p^L_{Bi} becomes constant and equal to (A.6) evaluated at $d = d^O_7(n, v, N)$.

The uninformed firm's symmetric equilibrium prices can be defined as:

$$\begin{cases} p^{L*}_6 = \frac{v}{2} & \text{if } d \in [\frac{1}{2}, d^O_5(n, v, N)) \\ p^{L*}_7 = 2d + v - 1 & \text{if } d \in [d^O_5(n, v, N), d^O_6(n, v, N)) \\ p^{L*}_8 = \text{(A.6)} & \text{if } d \in [d^O_6(n, v, N), d^O_7(n, v, N)) \\ p^{L*}_9 = \frac{2v(N-n)+n-1}{4N-3n-1} & \text{if } d \in [d^O_7(n, v, N), \frac{1}{2}] \end{cases}$$

Now we need to prove that every informed firm has no incentive to deviate from $p_B^{WL*} = v-1$ when $p_B^{L*} = \{p_7^{L*}, p_8^{L*}\}$.

The non-discriminatory demand facing firm j, when p_{Bj}^{WL} is in the neighborhood of p_B^{L*}, can be expressed as:

$$q_j^{WL} = \begin{cases} \frac{1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_B^{L*}-p_{Bj}^{WL}}{2} - d\right) + \frac{N-n}{N-1}\frac{2}{N}(v - p_{Bj}^{WL} - d) & \text{if } p_{Bj}^{WL} > v-1 \\ \frac{1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_B^{L*}-p_{Bj}^{WL}}{2} - d\right) + \frac{N-n}{N-1}\frac{2}{N}(1-d) & \text{if } p_{Bj}^{WL} < v-1 \end{cases}$$

If $d \in [d_5^O, d_6^O)$, the uninformed firm will set its uniform price to p_7^{L*}. The second order condition for profit maximization of the informed firm is satisfied for $p_{Bj}^{WL} \leq p_7^{L*} + 1 - 2d = v-1$. In order for $p_B^{WL*} = v-1$ to be an equilibrium, a slight increase of p_{Bj}^{WL} at p_B^{WL*}, should not increase profits:

$$q_j^{WL} + p_{Bj}^{WL}\frac{\partial q_j^{WL}}{\partial p_{Bj}^{WL}} = n(d+v-2) - N(d+v-2) - \frac{v}{2} + \frac{1}{2} \leq 0$$

which holds if and only if:

$$v \geq \frac{-2dn + 2dN + 4n - 4N - 1}{2n - 2N - 1} \tag{A.7}$$

It is easy to show that (A.7) is lower than $\frac{3}{2}$, $\forall\ (n,d) \in [2, N-1] \times [\frac{1}{2},1]$.

Also, a slight decrease of p_{Bj}^{WL} at p_B^{WL*} should not decrease profit:

$$q_j^{WL} + p_{Bj}^{W}\frac{\partial q_j^{WL}}{\partial p_{Bj}^{W}} = -n + N - \frac{v}{2} + \frac{1}{2} \geq 0$$

which holds if and only if:

$$v \leq 2N - 2n + 1 \tag{A.8}$$

It is easy to show that (A.8) is greater than 3, $\forall\ (n,d) \in [2, N-1] \times [\frac{1}{2},1]$.

For $d \in [d_7^O,1]$, the uninformed firm consistently sets a price $p_B^{L*} = \{p_8^{L*}, p_9^{L*}\}$ that is always lower than p_7^{L*}. Firm j's second order condition for profit maximization is satisfied for $p_{Bj}^{WL} \leq p_B^{L*} + 1 - 2d < v-1$. Therefore, also in this situation, each informed firm has the incentive to establish $p_B^{WL*} = v-1$ when (A.7) and (A.8) hold. Given that, these two conditions are satisfied for all values of $d \in [\frac{1}{2},1]$, we can conclude that all firms with data are in equilibrium with a basic price $p_B^{WL*} = v-1$.

□

Proof of Lemma 3.1. To establish the validity of point 1) in the Lemma, we calculate the first-order derivative of $p_8^{L*}(d,n,v,N)$ with respect to n, resulting in the subsequent expression:

$$\frac{\partial p_8^{L*}(d,n,v,N)}{\partial n} = \frac{(1-d)(N-1)}{2(n-N)^2}$$

which is always negative in its proper region of definition.

In order to prove point 2) and 3) it is sufficient to observe carefully the analytic expression of $d_6^O(d, n, v, N)$ and $d_7^O(d, n, v, N)$.

Recalling that the uninformed firm's demand function is described in equation (2.21), firm's profits in the outside option scenario can be defined as:

$$\pi^{L*} = q^{L*} p_B^{L*}$$

For $d \in [\frac{1}{2}, d_5^O)$, the basic price of the uninformed firm boils down to p_6^{L*} and the corresponding expression for profits is:

$$\pi_6^{L*} = \frac{v^2(2N - n - 1)}{4(N - 1)N}$$

which is independent of d.

For $d \in [d_5^O, d_6^O)$, the basic price of the uninformed firm boils down to p_7^{L*} and the corresponding expression for profits is:

$$\pi_7^{L*} = \frac{2(d - 1)(2d + v - 2)(n - 2N + 1)}{(N - 1)N}$$

which is decreasing in d in its proper region of definition.

For $d \in [d_6^O, d_7^O)$, the basic price of the uninformed firm boils down to p_8^{L*} and the corresponding expression for profits is:

$$\pi_8^{L*} = \frac{(d(-n) + d - nv + n + Nv - 1)^2}{2(N - 1)N(N - n)}$$

which is decreasing in d in its proper region of definition.

For $d \in [d_7^O$,1], the basic price of the uninformed firm boils down to $p_8^{L*}(d, n, v, N)$ evaluated at $d = d_7^O(n, v, N)$ and consequently the expression for profits is:

$$\pi_9^{L*} = \pi_8^{L*}(d = d_7^O) = \frac{2(N - n)(-2nv + n + 2Nv - 1)^2}{(N - 1)N(3n - 4N + 1)^2}$$

which is independent of d. □

Proof of Proposition 4. We extend the proof of Proposition 1, providing the definition of the symmetric equilibrium basic prices for $n \geq n_2^I(v, N)$ and $d \geq d_2^I(n, v, N)$. Indeed, in the proof of Proposition 1, we show that the upper boundary $d_2^I(n, v, N)$, which bounds from above the region of d where the symmetric equilibrium price p_2^{W*} holds, is lower than $\frac{1}{2}$ when $n \geq n_2^I(v, N)$. Finally, the proof is concluded by examining the potential incentive to deviate the basic price towards $v - 1$ in order to focus solely on the monopolistic market.

- $\mathbf{d \in [d_2^I(n, v, N)), \frac{1}{2})}$

We assume that for values of d greater than $d_2^I(n, v, N)$, the symmetric equilibrium price satisfies:

$$p_3^{W*} < v - 1$$

The non-discriminatory demand facing each informed firm j, when p^W_{Bj} is in the neighborhood of p^{W*}_3, can be expressed as:

$$q^W_j = \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2}+\frac{p^{W*}_3 - p^W_{Bj}}{2} - d\right) + \frac{N-n}{N-1}\frac{2}{N}(1-d)$$

Firm's j first-order condition for profit maximization is:

$$q^W_j - \frac{n-1}{N-1}\frac{1}{N}p^W_{Bj} + \frac{N-n}{N-1}\frac{2}{N} = 0$$

Therefore, at a symmetric equilibrium:

$$p^{W*}_3 = \frac{2N - 2d(N-1) - n - 1}{n-1}$$

To confirm the results obtained in the proof of Proposition 1 in terms of boundaries definition, we ensure that the requirement $p^{W*}_3 < v-1$ is satisfied if and only if:

$$d \geq d^I_2(n, v, N) = \frac{v(1-n) + 2N - 2}{2N-2}$$

Finally, it is necessary to verify that each informed firm has no incentive to deviate globally. At the candidate equilibrium, the second-order condition is satisfied for $p^W_{Bj} \leq p^{W*}_3 + 1 - 2d$. Furthermore, for $v - 1 \geq p^W_{BJ} > p^{W*}_3 + 1 - 2d$, demand is perfectly inelastic, profit is increasing in p^W_{Bj} and, for $p^W_{Bj} \geq v-1$ firm j 's profits are declining if $v \geq 2-d$. Therefore, if $v \geq 2-d$, the possibly most profitable deviation is $p^W_{Bj} = v-1$. Recalling the expression for the Inside Option profits in the non-overlapping segments scenario (2.15), the deviation is not profitable if $\pi^W(p^W_{Bj} = v-1) \leq \pi^W(p^W_{Bj} = p^{W*}_3)$, i.e.:

$$\frac{(d-1)(v-1)(n-N)}{(N-1)N} \leq \frac{(2d(N-1) + n - 2N + 1)^2}{(n-1)(N-1)N}$$

which holds if and only if:

$$v \leq \frac{\frac{(1-2d)^2(n-1)^2}{(d-1)(n-N)} - 4dn - 4(d-1)N + 8d + 2n - 6}{2(n-1)} \tag{A.9}$$

or equivalently:

$$d \leq d^I_{NE,1}(n, v, N) = \frac{1}{4}\left(\frac{(n-1)^2(v-1)}{(N-1)^2} - \frac{(n-1)(v+1)}{N-1} - a + 4\right) \tag{A.10}$$

with:

$$a = \sqrt{\frac{(n-1)^2(v-1)(n-N)(n(v-1) - N(v+3) + 4)}{(N-1)^4}}$$

When $v \leq 2-d$, each firm maximize its monopolistic profit for $p^W_{Bj} = \frac{v-d}{2}$. Consequently, the deviation is not profitable if $\pi^W(p^W_{Bj} = \frac{v-d}{2}) \leq \pi^W(p^W_{Bj} = p_3^{W*})$, which holds only if:

$$v \leq d + \sqrt{2}\sqrt{\frac{b}{(n-1)(N-n)}} \tag{A.11}$$

with:

$$b = 4d^2N^2 - 8d^2N + 4d^2 + 4dnN - 4dn - 8dN^2 + 12dN - 4d + n^2 - 4nN + 2n + 4N^2 - 4N + 1$$

We use the notation $d^I_{NE,2}(n, v, N)$ to represent the reformulation of equation (A.11) in terms of the variable d. However, we choose not to provide its explicit analytical form.

To ensure that the above condition implies a profitable deviation, we impose:

$$(\text{A.11}) \leq 2-d$$

which holds if and only if:

$$\begin{cases} \frac{1}{3}(2N+1) < n < N \\ \dfrac{n^2-2nN+2N^2-2N+1}{n^2-nN-n+2N^2-3N+2} - \dfrac{\sqrt{\frac{n^3N-3n^2N-n^4+3n^3-3n^2+3nN+n-N}{(n^2-nN-n+2N^2-3N+2)^2}}}{\sqrt{2}} \leq d \leq \dfrac{1}{2} \end{cases} \tag{A.12}$$

We can easily demonstrate that the partial derivative of equations (A.9) and (A.11) with respect to variable d is negative. Simultaneously, when we evaluate equations (A.9) and (A.11) at the lower bound of the second condition in (A.12), they are equal. As a result, we can define the value of d that determines the boundary of existence of pure-strategy Nash Equilibria as follows:

- $n \in [n^I_2(v, N), \frac{1}{3}(2N+1)]$

$$d^I_{NE}(n, v, N) = d^I_{NE,1}(n, v, N)$$

- $n \in (\frac{1}{3}(2N+1), N-1]$

$$d^I_{NE}(n, v, N) = \begin{cases} d^I_{NE,1}(n, v, N) & \text{if } v > (\text{A.11}) \text{ evaluated at } d = (\text{A.12}) \\ d^I_{NE,2}(n, v, N) & \text{if } v \leq (\text{A.11}) \text{ evaluated at } d = (\text{A.12}) \end{cases}$$

□

Proof of Lemma 4.1. All the results exposed in the proof of Lemma 1.2 about the behavior of π_1^{W*}, π_2^{W*} and π_4^{W*} hold true in their proper region of definition even when $n \in [n^I_2(v, N), N-1]$. Therefore, it remains to study the Inside Option profits for $d \in [d^I_2, d^I_{NE})$. Specifically, in this region of d the informed firms' basic price corresponds to p_3^{W*} and profits can be expressed as:

$$\pi_3^{W*} = \frac{d^2\left(-\left(n^2+n(5N-7)+N(3-4N)+2\right)\right) - 2d(n-N)((n-1)v-4N+4)}{(n-1)(N-1)N}$$

which is concave in d, $\forall$ $(n,v) \in [2, N-1] \times [\frac{3}{2}, 2]$. Therefore, it is sufficient to show that its first-order derivative with respect to d, evaluated at the lower bound of its valid range of definition, is lower than 0. By evaluating the first order derivative of π_3^{W*} at $d = d_2^I(n,v,N)$, we obtain:

$$\frac{\partial \pi_3^{W*}}{\partial d}(d = d_2^I) = \frac{v\left(3nN + (n-5)n - 2N^2 + N + 2\right) - 2(N-1)(n+N-2)}{(N-1)^2 N}$$

which is always lower than 0 in its proper region of definition. □

Proof of Proposition 5. We extend the proof of Proposition 2, providing the definition of the symmetric equilibrium basic prices for $n \geq n_2^O(v,N)$ and $d \geq d_2^O(n,v,N)$. Indeed, in the proof of Proposition 2, we show that the upper boundary $d_2^O(n,v,N)$, which bounds from above the region of d where the symmetric equilibrium prices $\{p_2^{WL*}, p_2^{L*}\}$ hold, is lower than $\frac{1}{2}$ when $n \geq n_2^O(v,N)$. Moreover, the equilibrium prices are initially derived by assuming that the second-order condition for profit maximization is always satisfied and no firm has an incentive to deviate. Ultimately, the proof is completed by demonstrating that the upper limit for defining equilibrium prices in pure strategies derived in the case of the Inside Option, guarantees the validity of equilibrium prices in the context of the Outside Option.

Fixing $N = 100$, the boundaries $n_2^O(v,N)$ and $d_2^O(n,v,N)$ can be expressed as:

$$n \geq n_2^O(v) = \frac{v + 200}{2v}$$

$$d \geq d_2^O(n,v) = \frac{1}{396}(-2nv + v + 398)$$

Moreover, $n_2^O(v)$ can be re-formulated as a function of v:

$$v \geq \frac{200}{2n-1} \tag{A.13}$$

It is worth noticing that (A.13) is lower than 2 when $n \in [51, 99]$. Consequently, the couple of basic prices $\{p_2^{WL*}, p_2^{L*}\}$ holds for $d \in [d_1^O(n,v), \frac{1}{2})$ if and only if $n < 51$.

- $\mathbf{d} \in [\mathbf{d_2^O}(\mathbf{n},\mathbf{v}), \mathbf{d_3^O}(\mathbf{n},\mathbf{v}))$

We assume that, when $n \geq n_2^O(v)$ and $d \geq d_2^O(n,v)$, the symmetric equilibrium prices satisfy:

$$\begin{cases} p_3^{WL*} < v - 1 \\ p_3^{L*} > v - 1 \\ p_3^{L*} < v - \frac{1}{2} \end{cases}$$

The non-discriminatory demand facing each firm j, when p_{Bj}^{WL} is in the neighborhood of p_3^{WL*} and p_3^{L*}, can be expressed as:

$$\begin{aligned} q_j^{WL} = & \frac{n-2}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_3^{WL*} - p_{Bj}^{WL}}{2} - d\right) + \frac{1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_3^{L*} - p_{Bj}^{WL}}{2} - d\right) + \\ & + \frac{N-n}{N-1}\frac{2}{N}(1-d) \end{aligned} \tag{A.14}$$

The uniformed firm i's demand function, when p^L_{Bi} is in the neighborhood of p_3^{WL*} , can be expressed as:

$$q_i^L = \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_3^{WL*} - p^L_{Bi}}{2}\right) + \frac{N-n}{N-1}\frac{2}{N}(v - p^L_{Bi})$$

By imposing first-order conditions on both firms' profit functions, the symmetric equilibrium prices, fixing $N = 100$, are the following:

$$\begin{cases} p_3^{WL*} = \dfrac{(198d-199)(n-1)+2(n-100)nv}{n(2n-397)-1} \\ p_3^{L*} = \dfrac{396d(n-199)+n(2n-2v-795)+200v+79201}{n(397-2n)+1} \end{cases}$$

The difference between p_3^{L*} and p_3^{WL*} is:

$$p_3^{L*} - p_3^{WL*} = \frac{594d(n-133)+2(n-100)((n-1)v+n-397)}{n(2n-397)-1} \tag{A.15}$$

The requirement that $p_3^{L*} > v-1$ is satisfied if:

$$d \le d_3^O(n,v) = \frac{n(596-2n-197v)-v-198}{198(n-1)}$$

By imposing that $d_3^O(n,v) > \frac{1}{2}$, as a function of v, we obtain:

$$v < \frac{-2n^2+497n-99}{197n+1} \tag{A.16}$$

It is worth noticing that (A.16) is lower than 2 when $n \in$ [51,99]. Consequently, the couple of prices $\{p_3^{WL*}, p_3^{L*}\}$ holds for $d \in [d_2^O(n,v), d_3^O(n,v))$ if and only if $n \ge 51$.

The requirement that $p_3^{WL*} < v-1$ is satisfied if:

$$d \ge d_2^O(n,v)$$

which gives us further confirmation of the results obtained in the proof of Proposition 2. As shown in the proof of Proposition 2, we know that $d_2^O(n,v)$ is lower than $\frac{1}{2}$ when condition (A.13) on v holds.

Moreover, it is easy to show that the difference between prices in (A.15), is bounded below by 0 and above by 1, in its region of definition (i.e., $d \ge d_2^O(n,v)$ and $n \ge n_2^O(v)$).

- $\mathbf{d \in [d_3^O(n,v), d_4^O(n,v))}$

We assume that when (A.16) holds conversely and $d \ge d_3^O(n,v)$, the symmetric equilibrium prices satisfy:

$$\begin{cases} p_4^{WL*} < v-1 \\ p_4^{L*} = v-1 \end{cases}$$

The best response of each informed firm j to a uniform price $p_4^{L*} = v - 1$ set by the uninformed firm i, is equal to:

$$p_4^{WL*} = \frac{-198d - n + v + 198}{n}$$

It is easy to show that $p_4^{L*} < v - 1, \forall\, d \geq d_3^O(n, v)$.

The non-discriminatory demand facing firm i, when p_{Bi}^L is in the neighborhood p_4^{WL*}, can be expressed as:

$$q_i^L = \begin{cases} \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_4^{WL*} - p_{Bi}^L}{2}\right) + \frac{N-n}{N-1}\frac{2}{N}(v - p_{Bj}^L) & \text{if } p_{Bi}^L > v - 1 \\ \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2} + \frac{p_4^{WL*} - p_{Bi}^L}{2}\right) + \frac{N-n}{N-1}\frac{2}{N} & \text{if } p_{Bi}^L < v - 1 \end{cases}$$

In order for p_4^{L*} to be an equilibrium, a slight increase of p_{Bi}^L at p_4^{L*} do not increase profits when:

$$d \geq d_3^O(n, v) = \frac{n(596 - 2n - 197v) - v - 198}{198(n - 1)}$$

which is lower than $\frac{1}{2}$ when (A.16) holds conversely and gives us further confirmation of the equilibrium prices obtained for $d \in [d_2^O(n, v), d_3^O(n, v))$.

At the same time, a slight decrease of p_{Bi}^L at p_4^{L*} do not decrease profits when:

$$d \leq d_4^O(n, v) = \frac{1}{2}(2n - 1)\left(\frac{2}{n - 1} - \frac{v}{99}\right)$$

By imposing that $d_4^O(n, v) > \frac{1}{2}$, as a function of v, we obtain:

$$v < 99\left(\frac{1}{1 - 2n} + \frac{2}{n - 1}\right) \tag{A.17}$$

It is worth noticing that (A.17) is lower than 2 when $n \in$ [76,99]. Consequently, the couple of prices $\{p_4^{WL*}, p_4^{L*}\}$ holds for $d \in [d_3^O(n, v), \frac{1}{2})$ if and only if $n < 76$.

The difference between p_4^{L*} and p_4^{WL*} is:

$$p_4^{L*} - p_4^{WL*} = \frac{198(d - 1) + (n - 1)v}{n}$$

Moreover, it is easy to show that the above difference between prices , is bounded below by 0 and above by 1, in its region of definition (i.e., $d \geq d_3^O(n, v)$ and (A.16) holding conversely).

- $\mathbf{d} \in [\mathbf{d_4^O(n, v)}, \mathbf{\frac{1}{2}})$

We assume that when (A.17) holds conversely and $d \geq d_4^O(n, v)$, the symmetric equilibrium prices satisfy:

$$\begin{cases} p_5^{WL*} < v - 1 \\ p_5^{L*} < v - 1 \end{cases}$$

The non-discriminatory demand facing firm j, when p_{Bj}^{WL} is in the neighborhood of p_5^{WL*} and p_5^{L*}, is equal to (A.14).

At the same time, the non-discriminatory demand facing firm i, when p^L_{Bi} is in the neighborhood of p_5^{WL*} is equal to:

$$q_i^L = \frac{n-1}{N-1}\frac{2}{N}\left(\frac{1}{2}+\frac{p_5^{WL*}-p^L_{Bi}}{2}\right)+\frac{N-n}{N-1}\frac{2}{N}$$

By imposing first-order conditions on both firms' profit functions, the symmetric equilibrium prices, fixing $N=100$, are the following:

$$\begin{cases} p_5^{WL*} = \dfrac{198}{n-1}-\dfrac{396d}{2n-1}-1 \\ p_5^{L*} = \dfrac{198}{n-1}-\dfrac{198d}{2n-1}-1 \end{cases}$$

The difference between p_5^{L*} and p_5^{WL*} is:

$$p_5^{L*}-p_5^{WL*} = \frac{198d}{2n-1}$$

The requirement that $p_5^{L*} < v-1$ is satisfied if and only if:

$$d \geq d_4^O(n,v) = \frac{1}{2}(2n-1)\left(\frac{2}{n-1}-\frac{v}{99}\right)$$

which is lower than $\frac{1}{2}$ when (A.17) holds conversely and gives us further confirmation of the equilibrium prices obtained for $d \in [d_3^O(n,v), d_4^O(n,v))$.

- **No Incentive to Deviations**

All the equilibrium prices described in this proof hold until the second-order condition for profit maximization is satisfied. Indeed, if $p_B^{WL*} < v-1$, the second-order condition for firm j's profits is satisfied until $p^{WL}_{Bj} < p_B^{WL*}+1-2d$. At the same time, if $p_B^{L*} < v-1$, the second order condition for firm i's profit is satisfied until $p^L_{Bi} < p_B^{L*}+1$.

Moreover, as in all spatial competition frameworks, all informed firms, when one or more of its competitors are uninformed, set a greater uniform price with respect to the scenario where all firms hold the same quantity d of data. It is easy to show that $n_2^I(v) < n_2^O(v)$, $\forall v \in [\frac{3}{2},2]$, and that $d_2^I(n,v) < d_2^O(n,v)$, $\forall (v,n) \in [\frac{3}{2},2]\times[2,N-1]$. Indeed, in their respective region of definition and for values of $d \in [d_2^I(n,v),\frac{1}{2})$, the basic price set in the Inside Option scenario is lower or equal than the basic price set in the corresponding Outside option:

$$p_B^{W*}(d,n,v) = \{p_3^{W*}\} \leq p_B^{WL*}(d,n,v) = \begin{cases} p_3^{WL*} & \\ p_4^{WL*} & \forall\ (d,n,v)\in \quad [d_2^I(n,v),\frac{1}{2})\times \\ p_5^{WL*} & \qquad\qquad\quad [n_2^I(v),N-1]\times[\frac{3}{2},2] \end{cases}$$

As a result, when $d \geq d_2^I(n,v)$, if a generic firm j with data in the Inside Option scenario has no motivation to deviate, the same will hold true in the Outside Option case where a generic firm $i \neq j$ has no information whatsoever. We define the range of values for d in which the existence of pure strategy equilibria is guaranteed in the Inside Option scenario, thereby outlining the region where equilibrium prices are defined also in the Outside Option scenario.

Our focus lies in establishing a total ordering among the boundary that determines the existence of pure-strategy Nash Equilibria in the Inside Option scenario and all the boundaries that serve as lower bounds for the regions of d where the Outside Option basic prices fall below $v-1$.

For $v \geq 2-d$ the most profitable deviation for each informed firm is $p_B^W = v-1$ and its corresponding boundary is $d^I_{NE,1}(n,v)$.

By imposing that $d_2^O(n,v) \leq d^I_{NE,1}(n,v)$ as a function of v, we obtain:

$$v \geq v_2^O(n) = \frac{2n^3+195n^2-99n+100}{4n^3-10n^2+206n-101} + \sqrt{\frac{a}{(4n^3-10n^2+206n-101)^2}} \tag{A.18}$$

with:

$$a = 4n^6 - 820n^5 + 43245n^4 - 126250n^3 + 176425n^2 - 143004n + 50400$$

which is greater than (A.13), $\forall\ n \in [51{,}99]$. Moreover, recalling that $d_2^O(n,v) \leq \frac{1}{2}$ when (A.13) holds conversely and that (A.13) is lower than 2 when $n \geq 51$, it is ensured that the inequality holds true in the proper region of equilibrium prices definition.

By imposing that $d_3^O(n,v) \leq d^I_{NE,1}(n,v)$ as a function of v, we obtain:

$$v \geq v_3^O(n) \tag{A.19}$$

which is greater than $v_2^O(n)$ and (A.16), $\forall\ n \in [51{,}99]$. Moreover, recalling that $d_3^O(n,v) \leq \frac{1}{2}$ when (A.16) holds conversely and that (A.16) is lower than 2 when $n \geq 51$, it is ensured that the inequality holds true in the proper region of definition.

By imposing that $d_4^O(n,v) \leq d^I_{NE,1}(n,v)$ as a function of v, we obtain:

$$v \geq v_4^O(n) \tag{A.20}$$

which is greater than $v_2^O(n)$,$v_3^O(n)$ and (A.17), $\forall\ n \in [2{,}99]$.

If all the above mentioned boundaries $v_2^O(n)$, $v_3^O(n)$ and $v_4^O(n)$, are lower than 2 for a given value of n, they exhibit strict decreasing behavior as n increases. Consequently, inverse functions $v_2^O(n)^{-1} = n_2^{O,NE}(v)$, $v_3^O(n)^{-1} = n_3^O(v)$ and $v_4^O(n)^{-1} = n_4^O(v)$ determine the boundaries for defining equilibrium prices in terms of n.

Moreover, in the proof of Proposition 4, we show that for v sufficiently low, the analytical expression of $d^I_{NE}(n,v)$ changes and becomes $d^I_{NE,2}$ when $2-v \geq d^I_{NE,2}(n,v)$. Fixing $N=100$, this condition can be expressed as a function of v in the following way:

$$v \leq \frac{n^2-2n+19603}{n^2-101n+19702} + \frac{\sqrt{\frac{-n^4+103n^3-303n^2+301n-100}{(n^2-101n+19702)^2}}}{\sqrt{2}} \tag{A.21}$$

which is greater than $\frac{3}{2}$ for $n \in [67{,}99]$ and gives us further confirmation of the results obtained in (A.12) as $\frac{1}{3}(2N+1) = 67$ for $N=100$. Moreover, (A.21) is increasing for

$n \leq 89$, where it reaches its maximum ($v \leq \frac{11}{7}$), and decreasing otherwise. Therefore, for each value of $v \in [\frac{3}{2}, \frac{11}{7})$, there exist $n_1^{NE,2}(v)$ and $n_2^{NE,2}(v)$ which define the boundary of equilibrium prices definition as a function of v.

For $v \in (\frac{11}{7}, 2]$, the equilibrium prices and their respective boundaries are defined in Proposition 5.

For $v \in [\frac{3}{2}, \frac{11}{7}]$, the equilibrium prices are defined as follows:

- if $n \in [n_2^O(v), n_1^{NE,2}(v))$

$$p^{L*}(d,n,v) = \begin{cases} \{p_1^{L*}, p_2^{L*}\} & || \; 0 \leq d < d_2^O \\ p_3^{L*} || \; (p^{L*} > v-1, p^{WL*} < v-1) & || \; d_2^O \leq d < d_{NE,1}^I \end{cases}$$

- if $n \in [n_1^{NE,2}(v), n_2^{NE,2}(v))$

$$p^{L*}(d,n,v) = \begin{cases} \{p_1^{L*}, p_2^{L*}\} & || \; 0 \leq d < d_2^O \\ p_3^{L*} || \; (p^{L*} > v-1, p^{WL*} < v-1) & || \; d_2^O \leq d < d_{NE,2}^I \end{cases}$$

- if $n \in [n_3^O(v), n_4^O(v))$

$$p^{L*}(d,n,v) = \begin{cases} \{p_1^{L*}, p_2^{L*}, p_3^{L*}\} & || \; 0 \leq d < d_3^O \\ p_4^{L*} || \; (p^{L*} = v-1, p^{WL*} < v-1) & || \; d_3^O \leq d < d_{NE,1}^I \end{cases}$$

- if $n \in [n_4^O(v), N-1)$

$$p^{L*}(d,n,v) = \begin{cases} \{p_1^{L*}, p_2^{L*}, p_3^{L*}, p_4^{L*}\} & || \; 0 \leq d < d_4^O \\ p_5^{L*} || \; (p^{L*} < v-1, p^{WL*} < v-1) & || \; d_4^O \leq d < d_{NE,1}^I \end{cases}$$

□

Proof of Lemma 5.1. Firstly, recall that we are considering $N = 100$. In order to prove points 1) and 3) of this Lemma, it is sufficient to observe carefully the analytic expressions of basic prices in the proof of Proposition 5. For what concerns the difference between basic prices $D_3^{O*} = p_3^{L*} - p_3^{WL*}$, it is easy to show that the first order-derivative of D_3^{O*} with respect to n, is decreasing in d. Hence, evaluating $\frac{\partial D_3^{O*}}{\partial n}$ at the upper bound of its valid range of definition ($d_3^O(n,v)$) and imposing that it is greater than 0, the following condition on v arises:

$$v \geq \frac{-n^2 + 266n - 26401}{66n - 13134}$$

which is lower than (A.13) in its proper region of definition.

Consequently, we will limit to outline the definition and the distinctive features of the uninformed firm's rents when $n \in [n_2^O(v), N-1]$. In particular, the behavior of π_1^{L*} and π_2^{L*}, as described in the proof of Lemma 2.1, holds true also in the region of Competitive Sub-Markets Prevalence.

For $d \in [d_2^O, \min(d_3^O, d_{NE}^I))$, the basic price of the uninformed firm boils down to $p_3^{L*}(d, n, v)$ and the corresponding expression for profits is:

$$\pi_3^{L*} = \frac{(199 - n)((198d - 199)(n - 1) + 2(n - 100)nv)^2}{9900(n(397 - 2n) + 1)^2}$$

which is decreasing in d in its proper region of definition.

For $d \in [d_3^O, \min(d_4^O, d_{NE}^I))$, the basic price of the uninformed firm boils down to $p_4^{L*}(d, n, v)$ and the corresponding expression for profits is:

$$\pi_4^{L*} = \frac{(1 - v)(198d(n - 1) + n((n - 2)v + n - 397) + v + 198)}{9900n}$$

which is decreasing in d in its proper region of definition.

For $d \in [d_4^O, d_{NE}^I)$, the basic price of the uninformed firm boils down to $p_5^{L*}(d, n, v)$ and the corresponding expression for profits is:

$$\pi_5^{L*} = \frac{(198d(n - 1) + (n - 199)(2n - 1))^2}{9900(1 - 2n)^2(n - 1)}$$

which is decreasing in d in its proper region of definition. □

Proof of Observation 2. Recalling that $N = 100$ and the objective of the DB is to maximize (2.7), we provide the definition of the DB's profits expressing the boundaries of equilibrium prices definition as a function of v. In order to prove point 1) of this observation for the exclusive information scenario, it is necessary to establish a complete ordering among the boundaries that determine the basic prices of the Inside and Outside Option equilibrium in the region of Competitive Sub-Markets Prevalence ($n \geq n_2^I(v)$). Moreover, we compute the DB's optimum strategy as each informed had no incentive to deviate and the symmetric equilibria described in Propositions 4 and 5 held for all values of $d \in [0, \frac{1}{2})$. We conclude the proof of point 1) showing that when the DB maximizes its profit for $d \in [0, \frac{1}{2})$ and $v \in [\frac{11}{7}, \frac{19}{10}]$ it will sell a quantity of data which ensures the presence of equilibrium in pure-strategy. Next, in order to prove point 2), we will use the optimal strategies of the Data Broker derived in the exclusive information scenario to demonstrate that, regardless of the values assumed by (n, v), the DB always has an incentive to sell at least a quantity of data equal to $\frac{1}{2}$.

In the proof of Proposition 5 we show that each informed firm, when all its competitors are informed, has an incentive to lower basic prices for lower values of d with respect to the Outside Option scenario. In the Inside Option scenario, reformulating the boundary $n \geq n_2^I(v)$, as a function of v, we obtain $v \geq \frac{99}{n-1}$. In the Outside Option scenario, for $n \geq n_2^O(v)$, the boundaries above which the uninformed firm's uniform pricing strategy hold are defined by (A.13) for p_3^{L*}, by (A.16) for p_4^{L*} and by (A.17) for p_5^{L*}.

Therefore, for $d \in [0, \frac{1}{2})$, the DB's profits can be defined as follows:

- if $v \in [\frac{99}{n-1}, (\text{A.13}))$:

$$\pi^{DB,E}(d, n, v) = \begin{cases} \pi_1^{DB,E} = \pi_1^{W*} - \pi_1^{L*} & || \; 0 \leq d < d_1^I(n, v) \\ \pi_2^{DB,E} = \pi_2^{W*} - \pi_1^{L*} & || \; d_1^I(n, v) \leq d < d_1^O(n, v) \\ \pi_3^{DB,E} = \pi_2^{W*} - \pi_2^{L*} & || \; d_1^O(n, v) \leq d < d_2^I(n, v) \\ \pi_4^{DB,E} = \pi_3^{W*} - \pi_2^{L*} & || \; d_2^I(n, v) \leq d < \frac{1}{2} \end{cases} \tag{A.22}$$

- if $v \in [(A.13),(A.16))$:

$$\pi^{DB,E}(d,n,v) = \begin{cases} \{\pi_1^{DB,E}, \pi_2^{DB,E}, \pi_3^{DB,E}, \pi_4^{DB,E}\} & || \; 0 \leq d < d_2^O(n,v) \\ \pi_5^{DB,E} = \pi_3^{W*} - \pi_3^{L*} & || \; d_2^O(n,v) \leq d < \frac{1}{2} \end{cases} \tag{A.23}$$

- if $v \in [(A.16),(A.17))$:

$$\pi^{DB,E}(d,n,v) = \begin{cases} \{\pi_1^{DB,E}, \pi_2^{DB,E}, \pi_3^{DB,E}, \pi_4^{DB,E}, \pi_5^{DB,E}\} & || \; 0 \leq d < d_3^O(n,v) \\ \pi_6^{DB,E} = \pi_3^{W*} - \pi_4^{L*} & || \; d_3^O(n,v) \leq d < \frac{1}{2} \end{cases} \tag{A.24}$$

- if $v \in [(A.17),2]$:

$$\pi^{DB,E}(d,n,v) = \begin{cases} \{\pi_1^{DB,E}, \dots, \pi_5^{DB,E}, \pi_6^{DB,E}\} & || \; 0 \leq d < d_4^O(n,v) \\ \pi_7^{DB} = \pi_3^{W*} - \pi_5^{L*} & || \; d_4^O(n,v) \leq d < \frac{1}{2} \end{cases} \tag{A.25}$$

For $d \in [0, d_1^I(n,v))$, regardless of the values of (n,v), the analytical expression of $\pi^{DB,E}$ boils down to $\pi_1^{DB,E}$, which is concave in d, $\forall\ (n,v) \in [2{,}99] \times [\frac{3}{2}{,}2]$. By imposing the first order condition on $\pi_1^{DB,E}$, we determine the optimal value $d_1^{DB,E}(n,v)$. We find that $d_1^{DB,E}(n,v)$ satisfies $d_1^{DB,E}(n,v) \geq d_1^I(n,v)$, when the following condition holds:

$$v \geq v_1^{DB,E}(n)$$

The above expression is lower than $\frac{3}{2}$, $\forall\ n \in [2{,}99]$. Hence, π_1^{DB} is increasing in d, for $d \in [0, d_1^I)$.

The proofs of Lemma 1.2 and 2.1 show that, regardless of the values of (n,v), $\pi_2^{W*}(d,n,v)$ is increasing in d, $\forall\ d \in [d_1^I, d_2^I)$, and that uninformed firm's profits are always decreasing or constant in d. Hence, $\pi_2^{DB,E}(d,n,v)$ and $\pi_3^{DB,E}(d,n,v)$ are always increasing in d in their region of definition.

Additionally, Lemma 4.1 shows that, regardless of the values of (n,v), $\pi_3^{W*}(d,n,v)$ is decreasing in d, $\forall\ d \in [d_2^I, \frac{1}{2})$. Therefore, as $\pi_2^{L*}(n,v)$ does not depend on d (proof of Lemma 2.1), we can easily conclude that $\pi_4^{DB,E}(d,n,v)$ is always decreasing in d in its region of definition.

Therefore, when the DB's profits boil down to (A.22), they are increasing for $d \in [0, d_2^I)$ and decreasing otherwise. Consequently, the optimum quantity of data to be sold is $d^{DB,E*} = d_2^I(n,v)$.

Taking into account the analytical form of $\pi_5^{DB,E}$, we calculate its second-order derivative with respect to d and determine that the resulting expression is negative when $n \geq 64$. By recalling that (A.13) and (A.16) are less than 2 when $n \geq 51$, we establish that within its defined range, $\pi_5^{DB,E}$ exhibits convexity for $n \in [51{,}63]$ and concavity for $n \in [64{,}99]$. Moreover, deriving the optimum value $d_5^{DB,E}(n,v)$ we obtain the following expression:

$$d_5^{DB,E}(n,v) = \frac{f_5(n,v)}{g_5(n)}$$

where $f_5(n,v)$ is a sixth-degree polynomial in n, linear in v and $g_5(n)$ is a sixth-degree polynomial in n. Moreover, by imposing that $d_5^{DB,E}(n,v) \in [d_2^O(n,v), d_3^O(n,v)]$, we obtain that v must be included in the following boundaries:

$$v_5^{DB,E,L}(n) \leq v \leq v_5^{DB,E,U*}(n) \quad \forall\, n \in [74{,}99] \tag{A.26}$$

where both $v_5^{DB,E,L}$ and $v_5^{DB,E,U*}$ are rational expression with polynomials. Additionally, they are both decreasing in n and lower than 2, $\forall\ n \in [74{,}99]$. Moreover, $v_5^{DB,E,U*} >$ (A.16), $\forall\ n \in [74{,}99]$. When (A.26) holds, the local optimum falls within the range of definition of $\pi_5^{DB,E}$. Consequently, recalling that DB's profits are increasing for $d \in [0, d_2^I)$ and decreasing for $d \in [d_2^I, d_2^O)$, the global optimum can be computed by solving the following second-order inequality as a function of v:

$$\pi_3^{DB,E}(d = d_2^I(n,v), n, v) \geq \pi_5^{DB,E}(d = d_5^{DB,E}(n,v), n, v)$$

which gives:

$$v \leq v_5^{DB,E*}(n)$$

The above upper boundary $v_5^{DB,E*}(n)$ satisfies condition (A.26), $\forall\ n \in [74{,}99]$, and delimits from above the region of v where $d_2^I(n,v)$ is also a global optimum. In addition, we present a set of inequalities that establish a connection between the boundary responsible for the presence of local optimum values within the specified range of $\pi_5^{DB,E}$, and the upper boundary that ensures the validity of the DB's profit definition in (A.23):

$$\begin{cases} (\text{A.13}) < v_5^{DB,E,L}(n) < v_5^{DB,E*}(n) < v_5^{DB,E,U*}(n) & \forall\, n \in [74{,}99] \\ (\text{A.16}) < v_5^{DB,E*}(n) & \forall\, n \in [74{,}87] \\ (\text{A.16}) > v_5^{DB,E*}(n) & \forall\, n \in [88{,}99] \end{cases} \tag{A.27}$$

Therefore, when the DB's profits boil down to (A.23) the global optimum can be defined as follows:

$$\begin{aligned} d^{DB,E*} &= \Big\{ d_2^I(n,v) \quad \|(\text{A.13}) \leq v \leq (\text{A.16}) && \forall\, n \in [51{,}87] \\ d^{DB,E*} &= \begin{cases} d_2^I(n,v) & \|(\text{A.13}) \leq v < v_5^{DB,E*}(n) \\ d_5^{DB,E}(n,v) & \|v_5^{DB,E*}(n) \leq v \leq (\text{A.16}) \end{cases} && \forall\, n \in [88{,}99] \end{aligned} \tag{A.28}$$

Taking into account the analytical form of $\pi_6^{DB,E}$, we calculate its second-order derivative with respect to d and determine that the resulting expression is negative when $n \geq 71$. By recalling that (A.16) is less than 2 when $n \geq 51$ and (A.17) is less than 2 when $n \geq 76$, we establish that within its defined range, $\pi_6^{DB,E}$ exhibits convexity for $n \in [51{,}70]$ and concavity for $n \in [71{,}99]$. Moreover, deriving the optimum value $d_6^{DB,E}(n,v)$ we obtain the following expression:

$$d_6^{DB,E}(n,v) = \frac{99(n(3n-398)-1)-(n-1)((n-199)n+99)v}{n(n(n+493)-39698)}$$

By imposing that $d_6^{DB,E}(n,v) \in [d_3^O(n,v), d_4^O(n,v)]$, we deduce that v must lie within the following bounds:

$$v_6^{DB,E,L*}(n) \leq v \leq v_6^{DB,E,U}(n) \quad \forall\, n \in [74{,}99] \tag{A.29}$$

with:

$$v_6^{DB,E,L*}(n) = \frac{2n^5 + 390n^4 - 314220n^3 + 15897220n^2 - 78210n + 19602}{n^4 - 136920n^3 + 7918617n^2 - 38908n + 19602}$$

$$v_6^{DB,E,U}(n) = 198\left(\frac{1}{n-1} - \frac{99\,(n^2-199)}{n(n(n(2n+787)-40289)-19306)+19602}\right)$$

where $v_6^{DB,E,L*}$ is lower than 2 and decreasing in n, $\forall\, n \in [74{,}99]$ and $v_6^{DB,E,U}$ is lower than 2 and decreasing in n, $\forall\, n \in [78{,}99]$. In addition, we present a set of inequalities that establish a connection between the boundaries responsible for the presence of local optimum values within the specified range of $\pi_5^{DB,E}$ and $\pi_6^{DB,E}$ (A.29), and the boundaries that ensure the validity of the DB's profit definition in (A.24):

$$\begin{cases} (\text{A.16}) < v_5^{DB,E,U*}(n) < v_6^{DB,E,L*}(n) & \forall\, n \in [74{,}99] \\ v_6^{DB,E,L*}(n) < (\text{A.17}) & \forall\, n \in [74{,}79] \\ v_6^{DB,E,L*}(n) > (\text{A.17}) & \forall\, n \in [80{,}99] \\ v_5^{DB,E,U*}(n) < (\text{A.17}) & \forall\, n \in [74{,}79] \\ v_5^{DB,E,U*}(n) > (\text{A.17}) & \forall\, n \in [80{,}99] \end{cases} \tag{A.30}$$

Consequently, recalling that DB's profits are increasing for $d \in [0, d_2^I)$ and decreasing for $d \in [d_2^I, d_2^O)$, and considering the system of inequalities in (A.30) and in (A.27), when the DB's profits boil down to (A.24) the global optimum can be defined as follows:

$$\begin{aligned} d^{DB,E*} &= \begin{cases} d_2^I(n,v) & ||\ (\text{A.16}) \leq v < v_5^{DB,E*}(n) \\ d_5^{DB,E}(n,v) & ||\ v_5^{DB,E*}(n) \leq v < v_5^{DB,E,U*}(n) \\ d_3^O(n,v) & ||\ v_5^{DB,E,U*}(n) \leq v < v_6^{DB,E,L*}(n) \\ d_6^{DB,E}(n,v) & ||\ v_6^{DB,E,L*}(n) \leq v \leq \min(2, (\text{A.17})) \end{cases} && \forall\, n \in [74{,}79] \\ d^{DB,E*} &= \begin{cases} d_2^I(n,v) & ||\ (\text{A.16}) \leq v < v_5^{DB,E*}(n) \\ d_5^{DB,E}(n,v) & ||\ v_5^{DB,E*}(n) \leq v \leq (\text{A.17}) \end{cases} && \forall\, n \in [80{,}87] \\ d^{DB,E*} &= \begin{cases} d_5^{DB,E}(n,v) & ||\ (\text{A.16}) \leq v \leq (\text{A.17}) \end{cases} && \forall\, n \in [88{,}99] \end{aligned} \tag{A.31}$$

Taking into account the analytical form of $\pi_7^{DB,E}$, we calculate its second-order derivative with respect to d and determine that the resulting expression is negative when $n \geq 51$. By recalling that (A.17) is less than 2 when $n \geq 76$, we establish that within its defined range, $\pi_7^{DB,E}$ exhibits concavity for $n \in [76{,}99]$. Moreover, deriving the optimum value $d_7^{DB,E}(n,v)$ we obtain the following expression:

$$d_7^{DB,E}(n,v) = \frac{(2n-1)(198(n(3n-202)+1) - (n-100)(n-1)(2n-1)v)}{n(n(4n(n+492)-121559)+80877)-494}$$

Enforcing the inequality $d_7^{DB,E}(n,v) \geq d_4^O(n,v)$, the following condition on v arises:

$$v \geq v_7^{DB,E,L}(n) \quad \forall\, n \in [78{,}99] \tag{A.32}$$

with:

$$v_7^{DB,E,L}(n,v) = 198\left(\frac{1}{n-1} - \frac{198((n-1)n-99)}{n(n(4n(n+393)-81365)+21279)+19306}\right)$$

where $v_7^{DB,E,L}$ is lower than 2 and decreasing in n, $\forall\, n \in [78{,}99]$. Moreover, we establish the following inequalities among the boundaries that define the presence of local optimum values within the range of definition of $\pi_6^{DB,E}$ and $\pi_7^{DB,E}$, as well as the boundaries that delimit the range of definition of equilibrium prices:

$$(\text{A.17}) < v_7^{DB,E,L}(n) < v_6^{DB,E,U}(n) \quad \forall\, n \in [78{,}99] \tag{A.33}$$

Consequently, recalling that DB's profits are increasing for $d \in [0, d_2^I)$ and decreasing for $d \in [d_2^I, d_2^O)$ and considering the inequality (A.33), the global optimum can be computed by solving the following second-order inequality as a function of v:

$$\pi_6^{DB,E}(d = d_6^{DB,E}, n, v) \geq \pi_7^{DB,E}(d = d_7^{DB,E}(n,v), n, v)$$

which gives:

$$v \leq v_7^{DB,E*}(n)$$

The above upper boundary $v_7^{DB,E*}(n)$ satisfies condition (A.32),$\forall\, n \in [78{,}99]$ and delimits from above the region of v where $d_6^{DB,E}$ is also a global optimum. Therefore, considering the systems of inequalities in (A.30) and (A.33), when the DB's profits boil down to (A.25) the global optimum can be defined as follows:

$$\begin{aligned}
&d^{DB,E*} = \begin{cases} d_6^{DB,E}(n,v) & ||\ (\text{A.17}) \leq v \leq 2 \end{cases} && \forall\, n \in [76{,}77] \\
&d^{DB,E*} = \begin{cases} d_6^{DB,E}(n,v) & ||\ (\text{A.17}) \leq v < v_7^{DB,E*}(n) \\ d_7^{DB,E}(n,v) & ||\ v_7^{DB,E*}(n) \leq v \leq 2 \end{cases} && \forall\, n \in [78{,}79] \\
&d^{DB,E*} = \begin{cases} d_5^{DB,E}(n,v) & ||\ (\text{A.17}) \leq v < v_5^{DB,E,U*}(n) \\ d_3^O(n,v) & ||\, v_5^{DB,E,U*}(n) \leq v < v_6^{DB,E,L*}(n) \\ d_6^{DB,E}(n,v) & ||\ v_6^{DB,E,L*}(n) \leq v < v_7^{DB,E*}(n) \\ d_7^{DB,E}(n,v) & ||\ v_7^{DB,E*}(n) \leq v \leq 2 \end{cases} && \forall\, n \in [80{,}99]
\end{aligned} \tag{A.34}$$

To sum up, considering the definitions and boundaries of the optimal strategies outlined in equations (A.28), (A.31) and (A.34), while also ensuring that each of them remains above or equal to $\frac{3}{2}$, we can conclude that:

$$d^{DB,E*} = \begin{cases} d_2^I(n,v) & || \text{ (A.13)} \leq v \leq 2 \end{cases} \quad \forall\, n \in [51{,}73]$$

$$d^{DB,E*} = \begin{cases} d_2^I(n,v) & || \ \frac{3}{2} \leq v < v_5^{DB,E*}(n) \\ d_5^{DB,E}(n,v) & || \ v_5^{DB,E*}(n) \leq v < v_5^{DB,E,U*}(n) \\ d_3^O(n,v) & || v_5^{DB,E,U*}(n) \leq v < v_6^{DB,E,L*}(n) \\ d_6^{DB,E}(n,v) & || \ v_6^{DB,E,L*}(n) \leq v \leq 2 \end{cases} \quad \forall\, n \in [74{,}77]$$

$$d^{DB,E*} = \begin{cases} d_2^I(n,v) & || \ \frac{3}{2} \leq v < v_5^{DB,E*}(n) \\ d_5^{DB,E}(n,v) & || \ v_5^{DB,E*}(n) \leq v < v_5^{DB,E,U*}(n) \\ d_3^O(n,v) & || v_5^{DB,E,U*}(n) \leq v < v_6^{DB,E,L*}(n) \\ d_6^{DB,E}(n,v) & || \ v_6^{DB,E,L*}(n) \leq v < v_7^{DB,E*}(n) \\ d_7^{DB,E}(n,v) & || \ v_7^{DB,E*}(n) \leq v \leq 2 \end{cases} \quad \forall\, n \in [78{,}93]$$

$$d^{DB,E*} = \begin{cases} d_5^{DB,E}(n,v) & || \ \frac{3}{2} \leq v < v_5^{DB,E,U*}(n) \\ d_3^O(n,v) & || v_5^{DB,E,U*}(n) \leq v < v_6^{DB,E,L*}(n) \\ d_6^{DB,E}(n,v) & || \ v_6^{DB,E,L*}(n) \leq v < v_7^{DB,E*}(n) \\ d_7^{DB,E}(n,v) & || \ v_7^{DB,E*}(n) \leq v \leq 2 \end{cases} \quad \forall\, n \in [94{,}99]$$

To ensure that the above definition holds true, the DB's optimum strategy should fall where the existence of a symmetric equilibrium in pure strategies is guaranteed. Recalling from the proof of Proposition 5 that for $v \in (\frac{11}{7},2]$ the boundary which ensures the existence of equilibrium in pure strategies $d_{NE}^I(n,v)$ boils down to $d_{NE,1}^I(n,v)$ (A.10), and by imposing the condition that $d_6^{DB,E}(n,v)$ is greater than $d_{NE,1}^I(n,v)$, the following inequality constraint on v arises:

$$v \geq v_6^{DB,E,NE}(n)$$

where $v_6^{DB,E,NE}$ is bounded below by $\frac{9}{5}$ and is greater than $v_6^{DB,E,L*}(n)$, $\forall\ n \in [55{,}99]$. Additionally, by imposing the condition that $d_7^{DB,E}(n,v)$ is greater than $d_{NE,1}^I(n,v)$, we obtain the following inequality constraint on v:

$$v \geq v_7^{DB,E,NE}(n)$$

where $v_7^{DB,E,NE}$ is always increasing in n and is greater than $v_7^{DB,E*}(n)$, $\forall\ n \in [71{,}99]$. Moreover, noticing that $v_7^{DB,E,NE}(n = 71) > \frac{9}{5}$, we find that for $v \in (\frac{11}{7}, \frac{9}{5}]$ the DB's optimal strategy ensures pure strategies in equilibrium prices and profits. By taking into account that all the boundaries mentioned above are decreasing in their region of definition and considering that $v_5^{DB,E,U*}(n = 99) < \frac{42}{25}$, the definition of the optimal strategy mentioned above can be reformulated as a function of n, as stated in point 1) of Observation 2.

Considering the above results, in order to prove point 2), we need to define the DB's profits for $d = \frac{1}{2}$:

$$\pi^{DB}(d,n,v) = \begin{cases} \pi_1^{DB} = \pi_4^{W*} - \pi_6^{L*} \ || \ \frac{1}{2} \leq d < d_5^O(n,v) \end{cases}$$

Now we compare the DB's optimum strategy derived in the exclusive information scenario denoted by $\pi^{DB,E}(d = d^{DB,E*}, n, v)$ to the value assumed by π_1^{DB} for $d = \frac{1}{2}$.

When $\pi^{DB,E}(d, n, v)$ boils down to (A.22), the DB's profits are increasing for $d \in [0, d_2^I)$ and decreasing otherwise. Noticing that the difference between $\pi_1^{DB}(d = \frac{1}{2})$ and $\pi_3^{DB,E}(d = d_2^I)$ is increasing in v for $v \geq \frac{99}{n-1}$, it can be shown that:

$$\pi_1^{DB}(d = \frac{1}{2}, n, v) > \pi_3^{DB,E}(d = d_2^I, n, v) \quad \forall\, v \in [\frac{99}{n-1}, 2]$$

When $\pi^{DB,E}(d, n, v)$ boils down to (A.23) and $v \in [v_5^{DB,E*}(n, v), v_5^{DB,E,U*}(n, v))$, the DB's optimum strategy coincides with $d_5^{DB,E}$. Noticing that the difference between $\pi_1^{DB}(d = \frac{1}{2})$ and $\pi_5^{DB,E}(d = d_5^{DB,E})$ is increasing in v for $v \geq$ (A.13), it can be shown that:

$$\pi_1^{DB}(d = \frac{1}{2}, n, v) > \pi_5^{DB,E}(d = d_5^{DB,E}, n, v) \quad \forall\, v \in [v_5^{DB,E*}, 2]$$

When $\pi^{DB,E}(d, n, v)$ boils down to (A.24) and $v \in [v_6^{DB,E,L*}(n, v), v_7^{DB,E*}(n, v))$, the DB's optimum strategy coincides with $d_6^{DB,E}$. It can be shown that:

$$\pi_1^{DB}(d = \frac{1}{2}, n, v) > \pi_6^{DB,E}(d = d_6^{DB,E}, n, v) \quad \forall\, v \in [v_6^{DB,E,L*}, v_7^{DB,E*})$$

When $\pi^{DB,E}(d, n, v)$ boils down to (A.25) and $v \in [v_7^{DB,E*}(n, v), 2]$, the DB's optimum strategy coincides with $d_7^{DB,E}$. It can be shown that:

$$\pi_1^{DB}(d = \frac{1}{2}, n, v) > \pi_7^{DB,E}(d = d_7^{DB,E}, n, v) \quad \forall\, v \in [v_7^{DB,E*}, 2]$$

□

Bibliography

Laura Abrardi, Carlo Cambini, Raffaele Congiu, and Flavio Pino. User data and endogenous entry in online markets. *SSRN Electronic Journal*, 2022.

Paul Belleflamme, Wing Man Wynne Lam, and Wouter Vergote. Competitive imperfect price discrimination and market power. *Marketing Science*, 39:996–1015, 9 2020.

Dirk Bergemann and Alessandro Bonatti. Markets for information: An introduction. *Annual Review of Economics*, 11:85–107, 8 2019.

D. Bounie, A. Dubus, and P. Waelbroeck. *Market for Information and Selling Mechanisms*. Working papers of the Center of Economic Research at ETH Zurich. CER-ETH - Center of Economic Research at ETH Zurich, 2022.

David Bounie, Antoine Dubus, and Patrick Waelbroeck. Selling strategic information in digital competitive markets. *The RAND Journal of Economics*, 52:283–313, 6 2021.

Ramon Caminal and Adina Claici. Are loyalty-rewarding pricing schemes anticompetitive? *International Journal of Industrial Organization*, 25:657–674, 8 2007.

E Chamberlin. The theory of monopolistic competition. *Cambridge: Harvard University Press*, 1933.

Yongmin Chen and Michael H. Riordan. Price and variety in the spokes model. *The Economic Journal*, 117:897–921, 7 2007.

Zhijun Chen, Chongwoo Choe, and Noriaki Matsushima. Competitive personalized pricing. *Management Science*, 66:4003–4023, 9 2020.

J Crémer, Y Montjoye, H Schweitzer, European Commission, and Directorate-General for Competition. *Competition policy for the digital era*. Publications Office, 2019.

Emilie Dargaud and Carlo Reggiani. On the price effects of horizontal mergers: A theoretical interpretation. *Bulletin of Economic Research*, 67:236–255, 7 2015.

Drew Fudenberg and J. Miguel Villas-Boas. *Behavior-Based Price Discrimination and Customer Recognition*. Elsevier Science, Oxford, 2007.

Avi Goldfarb and Catherine Tucker. Digital economics. *Journal of Economic Literature*, 57:3–43, 3 2019.

Yiquan Gu, Leonardo Madio, and Carlo Reggiani. Exclusive Data, Price Manipulation and Market Leadership. CESifo Working Paper Series 7853, CESifo, 1 2019.

Harold Hotelling. Stability in competition. *The Economic Journal*, 39:41, 3 1929.

Philippe Jehiel and Benny Moldovanu. Auctions with downstream interaction among buyers. *The RAND Journal of Economics*, 31, 2000.

Qihong Liu and Konstantinos Serfes. Quality of information and oligopolistic price discrimination. *Journal of Economics and Management Strategy*, 13:671–702, 12 2004.

Oksana Loginova. Price competition online: Platforms versus branded websites. *Journal of Economics & Management Strategy*, 31:259–283, 4 2022.

Rodrigo Montes, Wilfried Sand-Zantman, and Tommaso Valletti. The value of personal information in online markets with endogenous privacy. *Management Science*, 65:1342–1362, 3 2019.

Flavio Pino. The microeconomics of data – a survey. *Journal of Industrial and Business Economics*, 49:635–665, 9 2022.

Andrea Prat and Tommaso Valletti. Attention oligopoly. *American Economic Journal: Microeconomics*, 14:530–557, 8 2022.

Carlo Reggiani. Spatial price discrimination in the spokes model. *Journal of Economics & Management Strategy*, 23:628–649, 9 2014.

Carlo Reggiani. Non-localised spatial competition: The “spokes model”. *Spatial Economics Volume I*, pages 129–167, 2020.

Andrew Rhodes. Can prominence matter even in an almost frictionless market? *The Economic Journal*, 121(556), 2011.

Steven C. Salop. Monopolistic competition with outside goods. *The Bell Journal of Economics*, 10:141, 1979.

Greg Shaffer and Z. John Zhang. Competitive one-to-one promotions. *Management Science*, 48:1143–1160, 9 2002.

Curtis Taylor and Liad Wagman. Consumer privacy in oligopolistic markets: Winners, losers, and welfare. *International Journal of Industrial Organization*, 34:80–84, 5 2014.

Jacques-Francois Thisse and Xavier Vives. On the strategic choice of spatial price policy. *The American Economic Review*, pages 122–137, 1988.