Regulating data sales: the role of data selling mechanisms

Telecommunications Policy 48 (2024) 102813

Contents lists available at ScienceDirect

Telecommunications Policy

journal homepage: www.elsevier.com/locate/telpol

Regulating data sales: The role of data selling mechanisms[☆]

Laura Abrardi, Carlo Cambini, Flavio Pino [*]

Politecnico di Torino, Department of Management, Corso Duca degli Abruzzi, 24, 10129 Turin, Italy

ARTICLE INFO

JEL classification:
L12
L41
L86

Keywords:
Data Broker
Price discrimination
Competition
Auctions
TIOLI
Entry

ABSTRACT

We analyze the effects of different data selling mechanisms of a monopolistic Data Broker (DB) who sells consumer data to firms in a downstream market with free entry, where data can be used for consumer price discrimination. We consider three possible data selling mechanisms, namely auctions with and without reserve prices, and Take-It-Or-Leave-It offers, which exhibit decreasing levels of DB's bargaining power towards firms. We highlight the emergence of an entry barrier effect in the downstream market, regardless of the data selling mechanism. Moreover, we show that the auction-based selling mechanisms, and particularly the auction with reserve prices, induce the DB to sell the lowest quantity of data, implying the lowest level of consumer surplus. Conversely, under TIOLI, the DB floods the market for data, selling to all firms data partitions that overlap over subsets of consumers. Imposing the sale of non-overlapping partitions to all firms would improve consumer surplus and welfare.

1. Introduction

The central role of consumer data, an essential input for firms due to the rising share of online retail sales (Cramer-Flood, 2023), has prompted the rise of the Data Broker (DB) industry. DBs harvest consumer data from multiple sources, combine them into ready-to-use information and sell them to downstream firms (FTC, 2014). The amount of data collected and traded is staggering. Oracle, one of the largest consumer DBs, aggregates 3 billion user profiles from 15 million different websites, 1 billion mobile users, billions of purchases from grocery chains and 1500 large retailers, credit reporting agencies including Visa and Mastercard, as well as 700 million messages from social media, blogs, and consumer review sites, every day (Christl, 2017). Information sold by DBs is used by downstream firms to improve their targeting of consumers and adopt price discrimination strategies.

Despite aggregating detailed and sensitive information on a large share of the digital population,[1] DBs do not acquire data from individuals directly, relying instead on various sources such as public records, third-party transactions, and by monitoring online activities, often without consumers' awareness (ACCC, 2023). In this respect, they run a completely different business than online platforms that directly interact with consumers, such as social media and search engines, with significant policy implications. In recent years, regulatory actions such as the European Digital Service Act, Digital Markets Act, and the Data Act have been adopted to shape the market for data and ensure competitiveness. However, such regulations are targeted toward firms that directly gather data from consumers and only tangentially affect DBs, which are neither digital gatekeepers nor very large online platforms (Krämer et al., 2023; Ruschemeier, 2022). DBs are thus outside of the current scope of digital policies, and information on their industry is lacking. To overcome this, the Australian Competition and Consumer Commission (ACCC) has recently started to investigate DBs'

☆ We thank Marc Borreau, Joan Calzada, Antoine Dubus, Jan Krämer, Luca Sandrini, Tommaso Valletti and Patrick Waelbroeck for useful comments on previous versions of this manuscript.

* Corresponding author.
E-mail address: flavio.pino@polito.it (F. Pino).

[1] Acxiom, another DB, declares owning data about 68% of the digital population. See https://shorturl.at/dqLQ1

https://doi.org/10.1016/j.telpol.2024.102813
Received 4 March 2024; Received in revised form 6 June 2024; Accepted 6 June 2024
Available online 21 June 2024

business practices, seeking more information about the mechanisms governing their data sales (ACCC, 2023). Many questions are still unanswered. How do DBs use the data they collect? Under what terms and conditions (including price) do firms acquire DBs' products and services? Is this done via tender, negotiated contracts, take-it-or-leave-it list prices, or other means? What are the effects on consumer surplus? As of now, there is still little transparency and understanding about how DBs operate, the terms and conditions of the data sale, and the policy implications of their business model.

In this paper, we study the role of different data selling strategies, namely Take-It-Or-Leave-It, auctions with or without reserve prices, affect downstream entry and the related market outcomes, and in turn consumer surplus. Although most DBs do not publicly state their data selling strategies, some of these strategies have been observed online. Major DBs such as Acxiom and Experian offer some of their datasets at Take-It-Or-Leave-It prices on Amazon Web Services marketplace.[2] Conversely, the online data marketplace DataBroker.global allows sellers to either set posted prices, auction their datasets, or auction them with a reserve price.[3] By focusing on these three selling mechanisms, our results can usefully inform the policy debate of potential perils that can occur under specific selling strategies.

To address these questions, we build a model where a monopolistic DB sells consumer data to a downstream Salop (1979) market with free entry. The DB, who owns data about the exact location of all consumers in the downstream market, sells partitions of his dataset about a subset of consumers to downstream firms, which use the data to operate first-degree price discrimination on the identified consumers. We consider three possible data selling mechanisms, namely auctions with or without reserve prices, and Take-It-Or-Leave-It (TIOLI) offers, and assume that the data sale occurs simultaneously with firm entry.

As the DB's only source of revenues is the sale of data to firms, his optimal strategy is to maximize downstream firms' profits, which he extracts through the price of data. Downstream profit maximization requires the DB to trade off the three different effects of data on downstream profits. First, more data leads firms to fiercely compete for the same customers, decreasing downstream profits. Second, data allows firms to identify consumers and charge them with tailored prices, which extract consumer surplus and thus increase downstream profits. Third, an increase in downstream competition leads to lower entry. As market concentration increases, so do firms' profits.

We obtain two main results. First, all three selling mechanisms entail the same level of downstream entry. Indeed, the DB maximizes the third effect of data, reducing entry as much as possible by properly choosing the price and quantity of data. However, among the firms that enter, the number firms that obtain data in equilibrium differs according to the selling mechanisms, and so does the quantity of data sold. Under the auction mechanisms, the DB sells data to only a subset of firms. This strategy increases firms' willingness to pay for data, by threatening them of having to compete without data against informed rivals. Under TIOLI, this threat is ineffective and he sells data to all entering firms. In terms of the quantity of data sold, under the auction with reserve prices, the DB's bargaining power is at its highest, and he can extract most of the firms' profits by selling relatively small partitions of data. As his bargaining power is reduced by first removing reserve prices and then the auction mechanism altogether, the DB floods the downstream market with data to intensify downstream competition. This result extends those of Abrardi et al. (2024), who, in a similar setting, focus instead on a scenario where the data sale occurs through TIOLI offers after firm entry. They find a milder entry barrier effect of data relative to the setup of the present paper, where the data sale occurs at the entry stage. Moreover, by comparing different selling mechanisms, we find that the TIOLI sale is suboptimal from the DB's point of view, as he obtains higher profits under auction-based selling mechanisms.

Second, we highlight how the presence of the DB always leads to consumer harm with respect to the benchmark case where data are not available, especially under the auction mechanisms. This result depends on the trade-off between two opposite effects of data on consumer surplus. On the one hand, data have a direct positive effect on consumers, as the firms' competition over identified consumers decreases downstream prices. On the other hand, the *entry barrier effect* indirectly caused by data increases the downstream market concentration, ultimately harming consumers. While the second – negative – effect of data is constant across selling mechanisms, as they entail the same level of entry in equilibrium, the first – positive – effect is milder under auctions, due to the lower amount of data sold.

We then extend our baseline model along two main directions. First, we relax the assumption of a monopolistic DB, by reducing his market power vis-à-vis downstream firms. Data brokering is a fragmented industry counting thousands of companies, dominated by a few very large players who tend to specialize in specific market segments.[4] Our results show that a lower DB's market power softens the entry barrier effect, thus benefiting consumers. Second, we allow firms to choose how much data to use, and target fewer consumers than those identified by the data purchased. Indeed, the recent literature has highlighted nuanced strategic interactions enabled by data, such as firms choosing to share data with their rivals (Choe et al., 2024) or consumers voluntarily disclosing data to firms (Ali et al., 2022). Our results show that allowing firms to voluntarily limit their use of data entails a reduction of data prices, which in turn increases entry and benefits consumers.

From a policy perspective, our analysis highlights how different selling mechanisms affect consumers. We show that consumers are not indifferent between different selling mechanisms, despite the fact that they entail the same level of entry, and that a policymaker can avoid consumer harm and increase total welfare (relative to the case in which data are absent) by regulating

[2] See https://shorturl.at/aoW45 and https://shorturl.at/qxRY4

[3] See https://www.databroker.global/help/selling-data

[4] The most notable companies in the data brokering industry include Acxiom, Epsilon, and Oracle for consumer data, and Experian, Equifax, and TransUnion as credit reporting agencies. Official data brokers registries currently exist only in Vermont and California and contained 540 companies in July 2021, but according to recent estimates, there are approximately 4000 DBs worldwide. See https://privacyrights.org/resources/registered-data-brokers-united-states-2021 and https://www.webfx.com/blog/internet/what-are-data-brokers-and-what-is-your-data-worth-infographic/

the data sale along two dimensions. First, the DB should sell its data to all entering firms, as this strategy leads to the same market outcomes regardless of the selling mechanism chosen by the DB. Second, the DB should be required to sell data about all consumers, and, at the same time, to limit the overlap between data partitions sold to different firms. Indeed, while allowing all consumers to be reached through tailored offers increases their surplus by leveraging on the (positive) direct effect of data, the overlap between partitions intensifies competition, reducing entry and ultimately causing consumer harm.

The remainder of the paper is organized as follows. In Section 2 we summarize the relevant related literature. In Section 3 we describe the model. In Section 4 we focus on downstream competition and prices. In Section 5 we find the DB's equilibrium strategy under the different selling mechanisms and describe the subsequent market outcomes. In Section 6 we extend the model by introducing different degrees of the DB's market power and allowing firms to choose how much data to use. In Section 7 we discuss the policy implications and the strategy that a policymaker can adopt to avoid consumer harm. Section 8 concludes.

2. Related literature

Our work is related to two main strands of literature. First, we have a conspicuous literature on the effects of price discrimination in horizontally differentiated markets. The seminal work of Thisse and Vives (1988) highlights how price discrimination introduces two opposing effects. On the one hand, we have a *surplus extraction effect*, as firms can extract more surplus from consumers by offering them tailored prices. On the other hand, we have a *competition effect*, as the ability to price discriminate increases the competitive pressure and makes firms' pricing strategies more aggressive. A vast literature has analyzed the interplay between these two effects, both when data are exogenously available to firms (Bester & Petrakis, 1996; Chen et al., 2020; Liu & Serfes, 2004; Shaffer & Zhang, 1995; Shy & Stenbacka, 2016; Taylor, 2003; Taylor & Wagman, 2014) or when firms obtain data by interacting with consumers (Bergemann & Bonatti, 2011; Hagiu & Wright, 2020; Villas-Boas, 2004).[5]

As far as we are aware, the closest work from this strand of literature to our analysis is Liu and Serfes (2005). In their analysis, they focus on a Salop model where firms exogenously have all consumer data and can operate third-degree price discrimination, i.e., consumers are split into a given number of equally-sized segments. Their analysis shows that the number of entering firms exhibits a U-shaped curve with respect to the number of segments, i.e., entry is at its minimum when price discrimination is imperfect. Our analysis complements this work by showing that, if data are instead sold by a DB, firm entry is minimized even under first-degree price discrimination, as the data price acts as a supplementary barrier to entry.

The second and more recent strand of literature focuses instead on strategic data sales by DBs when data are used for price discrimination. Braulin and Valletti (2016) and Montes et al. (2019) focus on a monopolistic DB who sells data regarding all consumers to a duopoly downstream market, and show that the DB maximizes his profits by exclusively selling data to one firm. Bounie et al. (2021, 2022) expand on this scenario by showing that the DB is better off by not selling all consumer information, but instead leaves some consumers unidentified to temper the competition effect. Their analysis shows that this result holds under different selling mechanisms.

The closest work to our analysis is Abrardi et al. (2024), who analyze a monopolistic DB who sells data to a downstream Salop (1979) model. Their results show that the data sale reduces firm entry, leading to consumer harm. This work expands on Abrardi et al. (2024) in two fundamental ways. First, in our setting firm entry and the data sale occur simultaneously instead of sequentially. This allows the DB to anticipate how his data selling strategy affects firm entry and, in turn, firms' profits. Conversely, in Abrardi et al. (2024), the DB treats the number of entering firms as given and does not internalize his effect on entry. Second, we compare multiple selling mechanisms to analyze if and how the DB's selling strategy changes, and the related market outcomes. Moreover, we also focus on the strategies that a policymaker can implement to nullify the consumer harm identified in Abrardi et al. (2024). Indeed, the remedies proposed in our setting would also work in the case where firm entry and the data sale occur sequentially, as the policymaker intervenes on the size of the data partitions.

3. The model

Setting and players. A monopolistic DB (he/him) has data regarding all the consumers in a downstream market, represented by a circular city à la Salop (1979). Consumers (she/her) are uniformly distributed on the circumference and normalized to 1, with location indexed by $x \in [0, 1)$. Symmetric firms (it/its) with marginal cost equal to 0 enter the market by paying a fixed cost F. Firms are indexed by $i \in \{0, 1, 2, \dots, n-1\}$, where n is the number of entering firms. We assume sequential entry to avoid coordination problems and ignore integer constraints on n.[6] The DB sells partitions of his dataset to firms. Each partition of size d_i is centered on the location $\frac{i}{n}$. We assume that, after purchasing a partition d_i, firms locate at the center of that partition. For simplicity, the position of a generic firm i is $\frac{i}{n}$. The partition allows firm i to perform first-degree price discrimination on consumers located in $[\frac{i}{n} - \frac{d_i}{2}, \frac{i}{n} + \frac{d_i}{2}]$. Note that, in our setting, data are not an essential input for firms to compete in the downstream market.[7] Fig. 1 illustrates the downstream market populated by n firms, in which each firm obtains a partition of (potentially different) size d_i.

[5] For recent surveys regarding data markets, refer to Bergemann and Bonatti (2019), Goldfarb and Tucker (2019) and Pino (2022).

[6] This approach is standard in the literature on firm entry, and has been recently implemented in Rhodes and Zhou (2022).

[7] Indeed, as we show in the following sections, uninformed firms can enter and obtain positive profits.

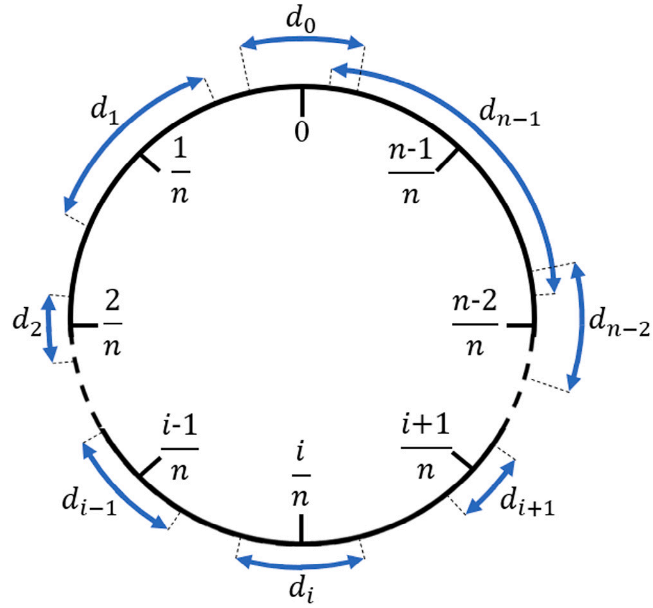

Fig. 1. Data partitions and identified consumers.

A firm i offers location-specific tailored prices $p_i^T(x) \geq 0$ to the consumer x in the identified segment, and a basic price $p_i^B \geq 0$ to unidentified consumers. A consumer located in x purchasing from firm i maximizes her utility

$$U(x,i) = u - p_i(x) - tD(x,i),$$

where u is the gross utility from consumption, $p_i(x) = \{p_i^T(x), p_i^B\}$ is either the tailored or basic price, depending on whether the consumer is identified or not, respectively, $t > 0$ is the transportation cost and $D(x,i)$ is the shortest arch between the consumer and firm i on the Salop circular city. The location of an indifferent consumer between firms i and $i+1$ is $\hat{x}_{i,i+1}$, i.e., $U\left(\hat{x}_{i,i+1}, i\right) = U\left(\hat{x}_{i,i+1}, i+1\right)$.

Firms' profits prior to paying for a data partition are equal to the integral of prices over their market share, and their expression depends on whether a firm identifies only some of its consumers or all of them. If a generic firm i does not identify all the consumers it serves, its profits when buying data, denoted by the superscript W, are

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

where the first term is the profit made on the identified consumers, and the second term is the profit obtained from unidentified consumers. Conversely, if firm i identifies all consumers it serves, its profits are

$$\pi_i^W = \int_{\hat{x}_{i-1,i}}^{\hat{x}_{i,i+1}} p_i^{\mathrm{T}}(x)\,dx - F, \tag{2}$$

as all consumers receive tailored prices. Finally, if firm i does not obtain a data partition, denoted with superscript L, its profits are

$$\pi_i^L = p_i^{\mathrm{B}}\left(\hat{x}_{i,i+1} - \hat{x}_{i-1,i}\right) - F, \tag{3}$$

as it does not identify any consumer.

Timing The timing of the model is as follows[8]:

Stage 1. The DB sells data, and firms choose whether to buy them and enter the market by paying the fixed cost F.
Stage 2. Firms set basic prices p_i^{B} for the unidentified consumers.
Stage 3. Firms set tailored prices $p_i{}^{\mathrm{T}}(x)$ for the identified consumers if they have obtained a partition.
Stage 4. Consumers purchase the product and profits are made.

[8] The sequentiality of Stages 2 and 3 allows the existence of Pure Strategy Nash Equilibria and is supported by managerial practices (Bounie et al., 2021). For example, the use of data results in more frequent price updates, allowing updated prices to be based on the rivals' prices (Fudenberg & Villas-Boas, 2006). See also Montes et al. (2019) for an analogous approach.

Our focus on the case in which the entry decision and data sale occur simultaneously could represent emerging markets where firms' business model is based on data. An example could be the introduction of rental electric scooter platforms, which were rapidly adopted in most metropolitan areas by consumers whose data were likely already available to DBs. In such a situation, companies would want to obtain consumer information as they enter the market to better profile their customer base and gain a competitive edge against their rivals. In particular, the simultaneity of the data sale and firm entry allows the DB to anticipate how his selling strategy affects firm entry and, in turn, his profits. This is different from the model proposed in Abrardi et al. (2024), in which the DB chooses his data selling strategy only after observing the number of entering firms. Thus, in this paper, the DB can strategically use his data sale to purposely affect downstream market concentration.

Data sale We explore three different data selling mechanisms. First, the DB can propose Take-It-Or-Leave-It (TIOLI) offers, as in Abrardi et al. (2024). The DB chooses a partition set $\mathbf{P} = (d_0, d_1, \dots, d_{n-1})$, offering each partition at a price w_i. Then, each firm independently and simultaneously chooses whether to purchase its respective partition.[9]

Second, as in Bounie et al. (2021), Braulin and Valletti (2016), Montes et al. (2019), the DB can sell data through auctions with reserve prices (AR), as described by Jehiel and Moldovanu (2000). The DB chooses the partition set $\mathbf{P}$ and sets up n auctions. In any auction $i \in \{0, \dots, n-1\}$, the DB offers a partition d_i and sets a reserve price v_i equal to the highest Willingness To Pay (WTP) for d_i among all firms. In particular, firm i has the highest valuation of partition d_i, as the partition is centered on its location. The reserve price in auction i is thus equal to the difference in firm i's profits between winning and losing auction i, given $\mathbf{P}$. The vector of reserve prices is $\mathbf{v} = (v_0, v_1, \ v_2, \dots, v_{n-1})$ and both $\mathbf{v}$ and $\mathbf{P}$ are common knowledge. Similar to Bounie et al. (2021), the DB declares the maximum number of auctions he fulfills, k, and only fulfills a subset $\mathbf{J}$ of auctions in equilibrium.[10] After observing all firms' bids, the DB chooses how many and which auctions he wants to fulfill. As some auctions may be unfulfilled, partitions traded in equilibrium may differ from those initially offered in the auctions through $\mathbf{P}$. We denote with $\mathbf{P}^*$ the partition set *traded* in equilibrium.

Third, we focus on auctions without reserve prices (AU), which decrease the DB's bargaining power with respect to the previous mechanism. When the DB cannot set reserve prices, a firm can win its auction simply by bidding above the valuations of the other firms, which are lower than its own owing to their distance. This in turn reduces the price of data with respect to the auction with reserve prices.[11]

Partition sets Abrardi et al. (2024) show that in a Salop model with symmetric firms and a monopolistic DB, firms never serve consumers located after their direct rivals' locations. Then, the downstream market can be seen as a concatenation of Hotelling lines. To simplify the analysis, we assume the DB adopts and replicates a strategy across all lines in equilibrium. Given two adjacent firms, the DB sells data to one or both.[12] This results in the DB either selling partitions to all entering firms ($\mathbf{P}^*_{\mathbf{A}} = (\widetilde{\mathbf{d}_{\mathbf{A}}}, \widehat{\mathbf{d}_{\mathbf{A}}}, \dots, \widehat{\mathbf{d}_{\mathbf{A}}})$), possibly alternating between two different partition sizes, or same-sized partitions to every other entering firm, alternating between informed and uninformed firms ($\mathbf{P}^*_{\mathbf{H}} = (\mathbf{d}_{\mathbf{H}}, \mathbf{0}, \dots, \mathbf{d}_{\mathbf{H}}, \mathbf{0})$). We solve the game through backward induction. As a benchmark, we refer to the standard Salop (1979) model, where the number of entering firms is $\widetilde{n} = \sqrt{\frac{t}{F}}$ and consumer surplus is $\widetilde{CS} = u - \frac{5}{4}\sqrt{tF}$.

4. Downstream equilibrium

4.1. Partition sets

Firm competition is influenced by the partition set offered by the DB under the various selling mechanisms. To streamline the analysis of downstream competition, the following Lemma derives some key characteristics of the partition sets.

Lemma 1. *If the DB sells data to all firms, he always offers* $\mathbf{P_A} = (d_A, d_A, \dots, d_A)$ *regardless of the selling mechanism. If the DB sells data to every other entering firm, he offers:*

- $\mathbf{P^{TIOLI}_H} = (d_H^{TIOLI}, 0, \dots, d_H^{TIOLI}, 0)$ *under TIOLI;*
- $\mathbf{P^{AR}_H} = (d_H^{AR}, 1, \dots, d_H^{AR}, 1)$ *under AR, declares* $k = \frac{n}{2} + 1$ *and fulfills the* $\frac{n}{2}$ *auctions where he offers* d_H^{AR}*;*
- $\mathbf{P^{AU}_H} = (d_H^{AU}, d_H^{AU}, \dots, d_H^{AU}, d_H^{AU})$ *under AU, declares* $k = \frac{n}{2} + 1$ *and fulfills every other auction.*

In equilibrium, firms always pay a data price equal to the difference in firm i*'s profit between obtaining it* (π_i^{W*}) *or not obtaining it* (π_i^{L*}).

Proof. See Appendix. ■

[9] Under TIOLI, the DB solves the coordination problem that is typical of the Salop model. Indeed, once a firm obtains a partition, it locates at its center. By offering partitions centered on the locations $\frac{i}{n}$, the DB ensures that entering firms are equally spaced in the downstream market.

[10] In a duopoly Hotelling setup, where each firm has only one rival as in Bounie et al. (2021), this strategy corresponds to declaring to fulfill exactly $\frac{n}{2}$ auctions. However, our model differs from the one in Bounie et al. (2021) as each firm has two direct rivals instead of one.

[11] Under both auction mechanisms, each firm has the same ex-ante valuation for any partition, as they have not entered the market yet and thus have no location assigned. Firms that win an auction will locate at the center of the partition they win. We instead assume that all losing firms enter sequentially in the market as in Salop (1979), thus avoiding coordination problems.

[12] We discard the case of the DB selling data to neither firm, as it would result in his profits being equal to zero.

If the DB sells data to all firms in equilibrium, then the offered partition set $\mathbf{P}$ must coincide with the equilibrium one $\mathbf{P}^*$. Due to firms' symmetry, we find that the DB maximizes his profits by setting $\widetilde{d_A} = \widehat{d_A} = d_A$. Regarding the sale to every other firm, under TIOLI the DB cannot withdraw any offers, and the partition set offered must again coincide with the equilibrium one. As for AR, the DB maximizes firms' WTP for data. To do so, the DB sets positive reserve prices for the auctions he wants to fulfill in equilibrium and sets reserve prices equal to zero in the auctions he does not want to fulfill, in which he offers the whole dataset. Through this strategy, the DB threatens a firm to sell the whole dataset to its direct rivals if the firm does not match the reserve price. Under AU, the absence of reserve prices implies that a firm wins an auction by beating all bids. To avoid underbidding, the DB prefers offering same-sized partitions in all auctions. In both cases, declaring $k = \frac{n}{2} + 1$ allows the DB to threaten any firm that, if it does not win the auction for its partition, he can sell data to both its direct rivals. Having described the characteristics of the partition sets offered by the DB under the three selling mechanisms, we now shift our attention toward firm competition.

4.2. Firms buy data

Let us focus on firm i obtaining a partition $d_i \in \{d_A, d_H\}$. The indifferent consumers between firms $i-1$ and i, and between i and $i+1$, are located in:

$$\hat{x}_{i-1,i} = \frac{2i-1}{2n} + \frac{p_i^{\mathrm{B}} - p_{i-1}^{\mathrm{B}}}{2t} \qquad \text{and} \qquad \hat{x}_{i,i+1} = \frac{2i+1}{2n} + \frac{p_{i+1}^{\mathrm{B}} - p_i^{\mathrm{B}}}{2t} \tag{4}$$

As firm i obtains a data partition, it offers a tailored price $p_i^{\mathrm{T}}(x)$ to the identified consumers by matching the competitor's basic price in utility level:

$$p_i^{\mathrm{T}}(x) = \begin{cases} p_{i-1}^{\mathrm{B}} + 2tx - \frac{t}{n}(2i-1) & \text{for } x \in [\frac{i}{n} - \frac{d_i}{2}, \frac{i}{n}] \\ p_{i+1}^{\mathrm{B}} - 2tx + \frac{t}{n}(2i+1) & \text{for } x \in [\frac{i}{n}, \frac{i}{n} + \frac{d_i}{2}] \end{cases} \tag{5}$$

By combining (1) with (A.20) and (A.21), we obtain

$$\pi_i^W = \frac{d_i}{2n}\left(2t + np_{i-1}^B + np_{i+1}^B - ntd_i\right) + p_i^B\left(\frac{n\left(p_{i+1}^B + p_{i-1}^B - 2p_i^B\right) + 2t}{2nt} - d_i\right) - F, \tag{6}$$

where the first term is firm i's profits obtained from the identified consumers, and the second term is the profit obtained from the unidentified ones. Eq. (6) shows two effects of d_i. First, it determines the share of identified consumers, positively affecting the first term. Second, it also influences the profits made by firm i on unidentified consumers, negatively affecting the second term.

If, instead, partition d_i is sufficiently large so that firm i identifies all consumers it serves, its profits when it buys data, prior to paying for them, are as in (2). In this case, firm i's basic price only affects its profits through the indifferent consumers' locations $\hat{x}$, as a lower basic price would expand firm i's market share. Firms are thus incentivized to lower their basic prices as much as possible, i.e., they set them to zero. Moreover, once the data partition allows firm i to identify all consumers up to the indifferent one, any additional data do not affect profits as they do not allow to poach additional consumers.

Finally, if firm i does not buy data, it becomes uninformed and competes having $d_i = 0$. By combining (3) with (A.20) and (A.21), we obtain

$$\pi_i^{\mathrm{L}} = p_i^B\left(\frac{n\left(p_{i+1}^B + p_{i-1}^B - 2p_i^B\right) + 2t}{2nt}\right) - F. \tag{7}$$

We then find the firm's reaction function on basic prices by obtaining the first-order condition of (6) and (7) with respect to p_i^{B}:

$$p_{i(\mathrm{W})}^{\mathrm{B}} = \frac{t}{2n} - \frac{td_i}{2} + \frac{p_{i+1}^{\mathrm{B}} + p_{i-1}^{\mathrm{B}}}{4} \tag{8}$$

$$p_{i(\mathrm{L})}^{\mathrm{B}} = \frac{t}{2n} + \frac{p_{i+1}^{\mathrm{B}} + p_{i-1}^{\mathrm{B}}}{4} \tag{9}$$

The reaction function in Eq. (8) presents an additional term with respect to the reaction function of the standard (Salop, 1979) model, namely $-\frac{td_i}{2}$. We conclude that data have two opposite effects on firms' profits. On the one hand, data allow firms to target consumers with tailored prices and extract more surplus from them. This increase in profits is referred to in the literature as*surplus extraction effect* of data (Thisse & Vives, 1988). However, acquiring more data leads to firm i offering its basic price to consumers who are on average farther from its location, requiring the firm to lower it. The reduction in profits stemming from the lower basic price is referred to in the literature as the *competition effect* of data (Thisse & Vives, 1988).

To find the equilibrium basic prices, we solve the system of firms' reaction functions. If the DB sells data to all firms, in equilibrium all firms obtain a partition d_A. then all reaction functions are as in (8). Instead, if the DB sells data to every other entering firm, then the reaction function of firms that obtain d_H are as in (8), whereas the reaction functions of uninformed firms are as in (9). The following Lemma summarizes the effects of data on firms' prices and profits in equilibrium.

Lemma 2. *Suppose firms buy data:*

- *If the DB sells data to all firms, firms' equilibrium prices and profits are decreasing in d_A for $0 < d_A < \frac{1}{n}$, and constant otherwise.*
- *If the DB sells data to every other entering firm, all firms' equilibrium basic prices are decreasing in d_H for $0 < d_H < \frac{3}{2n}$, and constant otherwise. Uninformed firms' equilibrium profits are decreasing in d_H for $0 < d_H < \frac{3}{2n}$, and constant otherwise. Informed firms' equilibrium profits follow an inverse U-shaped curve in d_H for $0 < d_H < \frac{3}{2n}$, and constant otherwise.*

Proof. See Appendix. ■

When all firms obtain data, the competition effect dominates the surplus extraction effect, leading to fiercer competition in the downstream market and a decrease in firms' profits. Due to firms' symmetry, the indifferent consumers are all located equidistantly from firms, i.e., at a distance of $\frac{1}{2n}$ from the two closest firms' locations. Then, once $d_A = \frac{1}{n}$, firms identify all consumers they serve, and additional data do not influence their equilibrium profits. Conversely, when half of the entering firms obtain data, informed firms have an advantage over the uninformed ones. Data thus allow informed firms to expand their market shares. At first, the market share expansion, together with the surplus extraction effect, dominates the competition effect of data, leading to higher profits. However, as the partition size d_H increases, informed firms can extract less profits from newly identified consumers, as they are located farther from their locations. Then, after a threshold, informed firms' profits start decreasing with respect to d_H. Finally, we find that data cannot completely overcome the positional advantage of firms. Indeed, consumers who are located too close to uninformed firms cannot be poached by the informed ones, even if they offer tailored prices equal to zero. Then, for $d_H \geq \frac{3}{2n}$, firms' profits become constant.

4.3. Firms do not buy data

Having analyzed the case where firms buy data, let us focus instead on the case where a given firm i does not obtain its respective partition. As concluded in Lemma 1, the DB's reaction to the firm's refusal to obtain data depends on the selling mechanism. Under TIOLI, if firm i does not buy data, the DB cannot change the partitions offered to all other firms. Thus, under the sale to all firms, firm i becomes the only uninformed firm in the market. Under AU, if firm i does not bid high enough to win its auction, then the DB fulfills the auctions of firms $i+1$ and $i-1$, that in turn obtain d_H. Finally, under AR, if firm i does not win its auction, then its direct rivals obtain the whole dataset. The following Lemma summarizes the effects of the data partitions' sizes on firm i's prices and profits.

Lemma 3. *Suppose firm i does not buy data:*

- *Under TIOLI and AU, firm i's basic price and profits in equilibrium are decreasing in the partition sizes obtained by the rivals, and become constant once $d_{\neq i} \geq \frac{3}{2n}$.*
- *Under AR, firm i's basic price and profits in equilibrium are constant in $d_{\neq i}$ and weakly lower than under TIOLI and AU.*

Proof. See Appendix. ■

Under TIOLI and AU, as firm i's rivals obtain more data, they price more aggressively and expand their market shares, leading to a decrease in firm i's profits. However, as already described in Lemma 2, firm i is always able to serve its closest consumers due to its positional advantage, which implies its profits become constant if the rivals' partitions are large enough. Conversely, under AR, firm i's direct rivals always obtain the whole dataset regardless of the partition size d_H. As they identify all consumers they serve, firm i's direct rival always set their basic prices to zero, resulting in firm i's equilibrium price and profits being constant with respect to d_H.

5. DB's profits, entry and welfare

Having analyzed downstream firms' competition stage, we now focus on the DB's selling strategies under the three selling mechanisms. Given a selling mechanism, the DB sets the equilibrium partition price equal to the difference in a firm's profits between obtaining or not obtaining said partition. Under the sale to all firms, the DB solves the following maximization problem:

$$\pi_{DB_A} = \max_{d_A} \quad n\left(\pi_i^W(\mathbf{P_A}) - \pi_i^L(\mathbf{P_A})\right) \tag{10}$$

$$s.t. \quad \pi_i^L(\mathbf{P_A}) = 0,$$

Where the constraint is the free entry condition. Conversely, as described in Lemma 1, the DB's strategy changes under the three selling mechanisms when he sells data to every other entering firm. By referring with $SM = \{TIOLI, AR, AU\}$ to the specific selling mechanism, the DB solves the following maximization problem:

$$\pi_{DB_H}^{SM} = \max_{d_H^{SM}} \quad \frac{n}{2}\left(\pi_i^W(\mathbf{P_H^{SM}}) - \pi_i^L(\mathbf{P_H^{SM}})\right) \tag{11}$$

$$s.t. \quad \pi_i^L(\mathbf{P_H^{SM}}) = 0.$$

The following proposition describes the DB's equilibrium strategies under the three selling mechanisms.

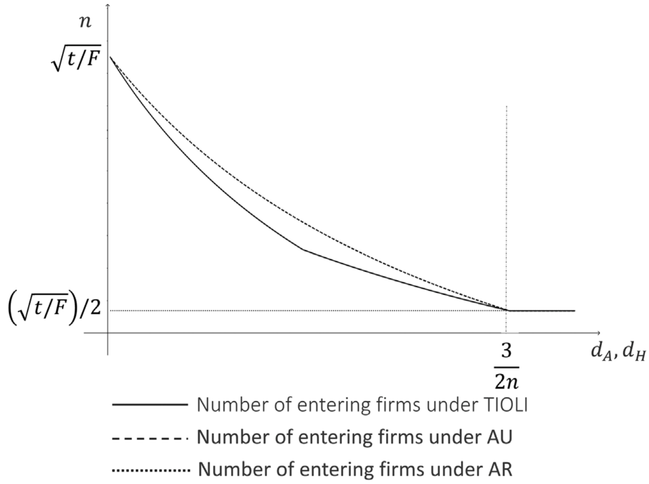

Fig. 2. Number of entering firms as a function of the partition's size.

Proposition 1. *In equilibrium, the DB adopts the following strategies given the selling mechanism:*

- *Under TIOLI, the DB opts for the sale to all firms, and sets* $d_A^{TIOLI^*} \geq \frac{3}{2n}$*;*
- *Under AU, the DB opts for the sale to every other firm, and sets* $d_H^{AU^*} \geq \frac{3}{2n}$*;*
- *Under AR, the DB opts for the sale to every other firm and sets* $d_H^{AR^*} = \frac{6}{7n}$*.*

The number of entering firms is always $n^* = \frac{1}{2}\sqrt{\frac{t}{F}} = \frac{\bar{n}}{2}$*.*

Proof. See Appendix. ■

To better understand the intuition behind the proposition, note that the DB's profits are decreasing in the number of entering firms. Intuitively, a higher number of firms implies fiercer downstream competition, reducing firms' willingness to pay for data. In turn, the number of entering firms is determined by firms' profits after paying for data π_i^L: higher profits entail a higher firm entry. Under both TIOLI and AU, π_i^L is decreasing in the partition's size, as informed rivals price more aggressively. Conversely, under AR, π_i^L is always minimized as, if firm i chooses not to buy data, its direct rivals always obtain the whole dataset. The effects of the partition's size on firms' profits also reflect on the number of entering firms, as we can see from Fig. 2: the number of entering firms is decreasing in the partition's size under TIOLI and AU, whereas it is constant under AR. This implies that the DB has the incentive to increase the partition's size under TIOLI and AU to reduce π_i^L and lower firm entry. Moreover, data have ambiguous effects on the partitions' prices, which are equal to $\pi_i^W - \pi_i^L$. Indeed, π_i^W can be positively or negatively affected by the partition's size, depending on the selling mechanism. Moreover, even if π_i^W is decreasing in the partition's size, firms' willingness to pay can still increase as long as the negative effect on π_i^L is stronger.

To understand how the different selling mechanisms affect the data sale, let us refer to Figs. 3 and 4, which show how firms' profits are influenced by the partition's size.

Under TIOLI, the sale to all firms always dominates the sale to every other firm. This is due to the DB's inability to change the offer made to a firm conditional on another firm's choice and is in line with the previous literature (Abrardi et al., 2024; Bounie et al., 2022).[13] To understand the intuition, suppose that a firm that is offered d_H^{TIOLI} chooses not to buy data. Then, it will be uninformed as it faces informed rivals. As competitive pressure is low, so is firms' willingness to pay for data, and the DB opts for the sale to all firms. Under the sale to all firms, increasing the partition size d_A has three effects. First, firms' profits when buying data decrease, as the *competition effect* of data dominates the *surplus extraction effect* (solid line in Fig. 3). This occurs as firms engage in price wars, dissipating profits. The strong competition effect thus reduces firms' willingness to pay for data. Second, firms' profits when not buying data also decrease, as they face fiercer competition from their informed rivals (dashed line in Fig. 3). This effect increases firms' willingness to pay for data. Third, as firms' profits when not buying data decrease, an increase in d_A reduces the number of entering firms. By reducing competition in the downstream market, this third effect increases firms' willingness to pay for data and, in turn, the DB's profits. The trade-off between the first two effects is the same as in Abrardi et al. (2024): an increase in the partition's size decreases firms' profits both when buying or not buying data, with ambiguous effects on the DB's profits. However, the third effect introduces an additional incentive for the DB to increase the partition size, as lower entry allows him to

[13] The inability of the DB to change the offer made to a firm contingent on another firm's response is due to the simultaneous nature of the TIOLI offers. For an analysis of a monopolistic DB selling partitions through sequential bargaining, see Bounie et al. (2022).

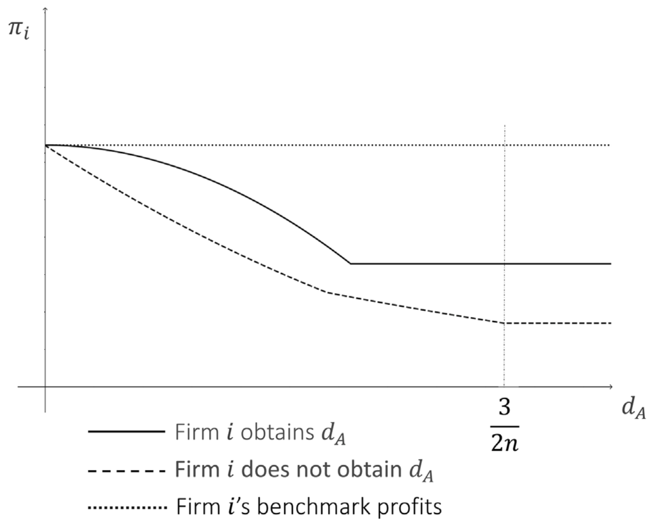

Fig. 3. Firm i's profits under the sale to all firms.

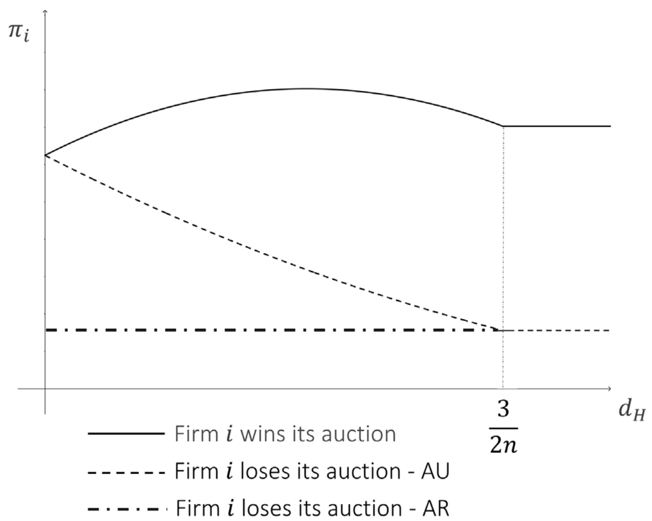

Fig. 4. Firm i's profits under the auction mechanisms.

increase the partitions' prices. The analysis in Abrardi et al. (2024) highlighted how, depending on the number of entering firms, the DB could either sell small or large partitions in equilibrium. Instead, by allowing the DB to anticipate the effects of his data sale on entry, we find that in equilibrium the DB always prefers minimizing firm entry, which he achieves by setting $d_A \geq \frac{3}{2n^*}$.

Under the auction mechanisms, the DB opts for the sale to every other firm instead. Under both auction mechanisms, the DB offers a subset of partitions that are not *traded* in equilibrium, allowing him to simultaneously increase firms' winning profits ($\pi_i^{W^*}$) and decrease firms' losing profits ($\pi_i^{L^*}$), as shown in Fig. 4. Under AU, similarly to TIOLI, the entry barrier effect of data dominates the *surplus extraction* and *competition effects*, and the DB sets $d_H^{AU^*} \geq \frac{3}{2n}$ to minimize entry. Instead, under AR, firms' profits when not buying data are always minimized, as we can see in Fig. 4. Indeed, regardless of $d_H^{AR^*}$, a losing firm under AR always faces direct rivals that obtain the whole dataset and thus exert the maximum competitive pressure. Then, the DB sets $d_H^{AR^*}$ to maximize firms' winning profits, as entry is always minimized.

Proposition 1 also highlights that if the data sale occurs simultaneously with firms' entry, the DB always maximizes his entry barrier effect. As his bargaining power is reduced, the DB floods the downstream market with data to reduce firms' profits after paying for data and, in turn, their entry.

Having analyzed the DB's equilibrium strategies, we now focus on the effects of the data sale on consumer surplus and welfare. The surplus of consumers buying from firm i is equal to the integral of consumers' utility:

$$CS_i = \int_{\hat{x}_{i-1,i}}^{\hat{x}_{i,i+1}} U(x,i)dx. \tag{12}$$

Total consumer surplus is the sum across all firms. Note that, as consumer surplus is computed in equilibrium, we need to separately analyze the cases of the sale to all firms and of the sale to every other firm. Under the sale to all firms, all firms buy data in equilibrium, and additional data have no effect when $d_A \geq \frac{1}{n}$ (see Fig. 3). Moreover, as priorly described, the data partition size sold by the DB d_A is not influenced by the selling mechanism under the sale to all firms. Thus, we can express consumer surplus as

$$CS_A = \begin{cases} v - \frac{5t}{4n} + \frac{ntd_A^2}{2} & \text{for } d_A \in [0, \frac{1}{n}) \\ v - \frac{5t}{4n_A} + \frac{t}{2n} & \text{for } d_A \in [\frac{1}{n}, 1]. \end{cases} \tag{13}$$

Conversely, under the sale to every other firm, only one of every two firms obtains data in equilibrium, and additional data have no effect when $d_H \geq \frac{3}{2n}$ (see Fig. 4). Then, consumer surplus is equal to

$$CS_H^{SM} = \begin{cases} v - \frac{5t}{4n} + \frac{ntd_H^{SM^2}}{9} & \text{for } d_H^{SM} \in [0, \frac{3}{2n}) \\ v - \frac{5t}{4n} + \frac{t}{4n} & \text{for } d_H^{SM} \in [\frac{3}{2n}, 1], \end{cases} \tag{14}$$

Where $SM = \{TIOLI, AR, AU\}$ is the selling mechanism adopted by the DB under the sale to half of the firms.

With regards to total welfare TW, we define it as the weighted sum of consumer surplus $CS \in \{CS_A, CS_H^{SM}\}$, firm profits and the DB' profits:

$$TW = CS + \alpha\left(\sum_{i=0}^{n-1}(\pi_i^L) + \pi_{\text{DB}}\right), \tag{15}$$

where $\alpha \in [0,1]$ is the weight assigned to industry profits in the welfare function, representing the lower weight a policymaker attributes to industry profits because of, for example, their international reach (Baron & Myerson, 1982). Moreover, in equilibrium, $\pi_i^{L*} = 0$ by the free-entry condition. Recall that $\widetilde{CS} = \widetilde{TW} = v - \frac{5}{4}\sqrt{tF}$ are the consumer surplus and total welfare in the standard Salop (1979) model, respectively. The following proposition summarizes the effects of the DB's data sale on consumer surplus and welfare under the three selling mechanisms.

Proposition 2. *In equilibrium, both consumer surplus and total welfare are higher under TIOLI than under AU or AR. Moreover,* $CS_A^* < \widetilde{CS}$. *There exists a threshold* $\bar{\alpha}$ *such that, iff* $\alpha \geq \bar{\alpha}$, $TW_A^* \geq \widetilde{TW}$.

Proof. See Appendix. ■

As shown in (13) and (14), CS is a function of the number of entering firms and the amount of data sold in equilibrium. As the number of entering firms is the same across selling mechanisms, as concluded in Proposition 1, this variable does not influence the results. Instead, an increase in the quantity of data drive firms' prices downwards, benefiting consumers. As the quantity of data sold is maximized under TIOLI, consumer surplus is maximized under this selling mechanism. However, consumers are harmed with respect to the standard Salop model. Indeed, the entry barrier effect posed by the DB's data sale increases downstream market concentration, leading to higher prices and lower consumer utility.

With regard to total welfare, we find that it is maximized under TIOLI. Indeed, the data sale only transfers surplus between firms and the DB, and the only net losses of total welfare derive from the fixed entry cost F and the transportation cost t. As previously described, the number of entering firms is constant across selling mechanisms. Instead, transportation costs are minimized under TIOLI, as all firms obtain data in equilibrium, leading to a symmetric outcome in which all consumers buy from the closest firm. We also find that, if the weight α of the industry profits in the welfare function is sufficiently high, total welfare is higher than in the benchmark case.[14] Indeed, the DB's equilibrium strategy under all selling mechanisms solves the excessive entry problem identified by Salop (1979). However, while being higher, total welfare is mostly appropriated by the DB, potentially raising redistributive concerns from a policymaking point of view.

6. Extensions

The previous analysis highlighted how the TIOLI mechanism, while providing the maximum level of consumer surplus relative to auctions, still makes consumers worse off relative to a situation in which a DB is absent and no data is sold, as in the standard Salop model. The harm to consumers stems from the entry barrier effect of data, which leads to higher market concentration and, in turn, higher prices. The entry barrier effect is particularly strong in our model due to the DB's monopolistic position and to the sale of the whole dataset, which intensifies firms' competition and depletes their profit. In this section we extend our baseline model by analyzing the role of the DB's market power (in Section 6.1) and by assuming that firms may decide to use a subset of the data purchased to soften competition (in Section 6.2).

[14] In the Appendix, we show that total welfare increases iff $\alpha \geq \frac{1}{2}$ under the TIOLI selling mechanism.

6.1. DB's market power

In the baseline model, we assume that the DB is a monopolist, who is thus able to extract all of the firms' willingness to pay for data. We now extend our model by analyzing how market outcomes change in response to a reduction of the DB's market power.[15]

To do so, we introduce a parameter $\theta \in (0,1]$, which measures the DB's market power.[16] Given a generic firm i willingness to pay for data, which is equal to $w_i = \pi_i^W - \pi_i^L$, we assume that the DB can only post a price equal to θw_i. Such reduction in market power could stem from the presence of competing DBs, or the presence of other alternatives available to firms to obtain data analytics. We focus on the equilibrium under the sale to all firms. Indeed, in the presence of competition, the threat posed to firms under the sale to half of the firms would no longer be available, as those firms could instead buy data from a competing DB.

The following proposition describes the effects of θ on the DB's equilibrium strategy and the subsequent market outcomes.

Proposition 3. *Under TIOLI, for any $\theta \in (0,1]$, the DB's equilibrium strategy is to offer $d_A \geq \frac{3}{2n}$ to all entering firms, and DB's profits are decreasing in θ. Consumer surplus is increasing in θ and there exists a threshold $\bar{\theta}$ such that, iff $\theta \geq \bar{\theta}$, $CS_A^* \geq \widetilde{CS}$. Total welfare is decreasing in θ iff $\alpha < \frac{3}{4}$.*

Proof. See Appendix. ■

The main result of Proposition 3 is that a reduction in market power does not affect the DB's equilibrium strategy. Indeed, regardless of θ, the DB always prefers to sell large data partitions to maximize competitive pressure in the downstream market and, in turn, the entry barrier effect. Intuitively, the level of the DB's market power proportionally scales down all of the DB's profits as a function of d_A, leaving his equilibrium strategy unaltered.

Looking at consumer surplus, the reduction in the DB's market power softens the entry barrier effect, which in turn decreases the prices posted by firms, thus benefiting consumers. In particular, a low enough θ allows the pro-competitive effect of data to dominate the reduction in entry, leading to a better outcome for consumers when compared with the benchmark case where data are absent.

Finally, we find that a lower level of DB's market power can increase total welfare, provided that industry profits have a sufficiently low weight in the welfare function. In such a case, the excessive entry that is typical of the Salop model is dominated by the lower transportation cost paid by consumers, as more firms enter the market.

6.2. Firms' ability to use less data

In Section 5, we have shown how the DB has the incentive, under TIOLI, to sell large partitions to all entering firms to increase downstream competition and, in turn, the price of data. A question then naturally arises: do firms have the incentive to use all the data they purchased or would they rather prefer to only use a part of them, if given the choice? Indeed, using less data would lower the competitive pressure firms face, benefiting their profits. At the same time, however, using less data when facing informed rivals could put a firm at a disadvantage and decrease its profits. In this Section we analyze this trade-off.

We focus on the equilibrium under TIOLI, in which the DB sells large partitions to all entering firms, for two reasons. First, under TIOLI firms obtain large partitions, so that allowing firms to strategically use only a part of the purchased data might have a high impact on their strategy. Second, as shown in Section 5, TIOLI is the selling mechanism under which consumers are less harmed and is thus the ideal starting point to understand whether the strategic use of data by firms can make the data sale beneficial for consumers. Moreover, many DBs adopt TIOLI sales online,[17] making this case particularly interesting for its practical implications.

To analyze firms' incentive to use fewer data than those purchased, we focus on the TIOLI equilibrium described in Section 5, in which the DB offers $d_A^* \geq \frac{3}{2n}$. Then, we allow firms to individually and simultaneously choose $d_{A,i} \in [0, d_A^*]$. The following proposition describes firms' equilibrium strategies and how they affect market outcomes.

Proposition 4. *Suppose firms can choose to use a subset $d_{A,i} \leq d_A^*$ of the data they purchase. There exist only two equilibria. In the first one, all firms purchase data and choose $d_{A,i}^* = d_{low}^* < \frac{1}{n}$. In the second one, all firms purchase data and choose $d_{A,i}^* = d_{high}^* \geq \frac{1}{n}$.*

Proof. See Appendix. ■

Firms' ability to choose how much data to use affects their strategies, generating a coordination game. To understand the intuition, let us focus on a generic firm i. Suppose that all of firm i's rivals choose to use only a small subset of the acquired partitions. Then, competitive pressure is low and firms charge positive basic prices, leading to higher profits with respect to those obtained in the baseline model. In such a situation, also firm i has the incentive to use a small subset of the partition. Indeed, using

[15] Although the assumption of a monopolistic DB might fit specific situations, where one DB has exclusive access to data relating to a specific consumer segment, little is known regarding the concentration level in the DB industry (ACCC, 2023).

[16] This approach allows us to simply yet effectively introduce a proxy for competition in the data industry, without modeling more complex interactions, such as data complementarities, that would be outside the scope of the paper.

[17] Large DBs such as Acxiom and Experian offer some of their datasets at posted prices on Amazon Web Services marketplace. See https://shorturl.at/aoW45 and https://shorturl.at/qxRY4

all available data would induce a strategic response from its rivals, which would lower their basic prices in response to the increased competitive pressure. This strategic response would reduce firm i's profits, making the deviation unprofitable.

Conversely, if all of firm i's rivals choose to use all available data, firm i faces the highest level of competitive pressure. Then, data become the only way for firm i to defend its market share from the rivals' aggressive pricing strategies. In particular, any partition that allows firm i to identify the consumer on which it has a positional advantage is an equilibrium, as it ensures that firm i serves those consumers.

Having analyzed the two possible equilibria, we now describe how a change in firms' strategies affects market outcomes. In continuity with the equilibrium discussed in previous sections, in which firms use overlapping partitions, we focus on the equilibrium where firms use the d^*_{high} amount of data.

Proposition 5. *Consider the equilibrium in which firms use d^*_{high} data. In equilibrium, data partition prices are lower, entry is higher, and $CS^* > \widetilde{CS}$ with respect to the baseline model.*

Proof. See Appendix. ■

To understand the implication of firms' ability to strategically choose how much data to use, it is useful to separately analyze the cases where all firms buy data and where one firm chooses not to buy data. First, as described in Proposition 4, if all other firms choose to use a large subset of the partition, then firm i's best response is to choose $d_{i,A} = d^*_{high} \geq \frac{1}{n}$, i.e., to offer tailored prices to all consumers on which it has a positional advantage. Indeed, firm i faces high competitive pressure from its direct rivals, who can identify all consumers. Then, due to the threat of consumers being poached, firm i chooses to defend its turf by offering tailored prices to all the consumers it serves, even though such choice maximizes competitive pressure. Thus, if all firms purchase data, they face a prisoner's dilemma: even though they would be better off not using data, the threat posed by their rivals forces them into using them.

Suppose instead that firm i chooses not to purchase data. Then, its direct rivals $i+1$ and $i-1$ use all available data on the arches they do not share with firm i, as they face informed firms on those arches. However, firms $i+1$ and $i-1$ can choose the amount of consumers they want to identify on the arch they share with firm i. We find that, in this subgame equilibrium, firms $i+1$ and $i-1$ choose to leave some consumers unidentified on the arches they share with firm i to temper competition with firm i and extract more profits.

The change in firms' strategy directly impacts the data price. Indeed, if a firm chooses not to buy data, it faces a lower competitive pressure with respect to the baseline model, resulting in a lower willingness to pay for data. DB's profits thus decrease under firms' strategic use of data, and entry increases. In particular, the reduction of the entry barrier effect is so strong that, in equilibrium, consumer surplus is higher than in the standard Salop model, as opposed to the result described in Section 5. Thus, our analysis shows that allowing firms to have agency over the amount of data they use can have positive effects on consumers.

7. Policy discussion

A central result of our analysis is that the DB's data sale, by drastically reducing entry, leads to consumer harm. This finding is in contrast with that obtained in most of the literature on price discrimination (e.g., the studies following the seminal paper by Thisse & Vives, 1988), highlighting that price discrimination by symmetric firms benefits consumers. In this section, we analyze possible policy interventions to avoid the consumer harm caused by the data sale. A first channel of policy intervention is the regulation of the data price. By mandating a lower data price, a policymaker could redistribute profits from the DB towards firms, increasing entry and, in turn, consumer surplus. However, implementing price regulation on data may be challenging in practice, due to the elusive nature of the traded commodity.[18] Other types of interventions have less obvious effects than direct price regulation and deserve a deeper analysis. In what follows we allow a policymaker to regulate the selling mechanism and the quantity of data sold, but not the prices posted by the DB. The following proposition describes the optimal size of data partitions and selling mechanisms from a social point of view, denoted with superscript P.

Proposition 6. *For any $\theta \in [0,1]$, consumer surplus is maximized under the sale to all firms with $d_A^{P^*} = \frac{1}{n}$. This results in $CS^{P^*} \geq \widetilde{CS}$ and $TW^{P^*} \geq \widetilde{TW}$. Total welfare is maximized with $d_A^{P^*} = \frac{1}{n}$ iff $\alpha < \bar{\bar{\alpha}}$, and with $d_A^{P^*} \geq \frac{3}{2n}$ otherwise.*

Proof. See Appendix. ■

Our analysis highlights how a policymaker can obtain a Pareto improvement with respect to the benchmark Salop (1979) model by regulating two features of the data sale. First, by imposing the sale to all firms, the regulator ensures that all firms can access data regardless of the selling mechanism adopted by the DB. This, in turn, creates a leveled playing field in the downstream market, as firms engage in price wars that benefit consumers. Second, by controlling the amount of data sold, the policymaker can strike a balance between the negative entry barrier effect and the positive competition effect. To better understand the intuition, let us refer to Fig. 5.

[18] Two datasets of equal size might carry vastly different amounts of usable information. Moreover, the value of the information depends on the firm using it.

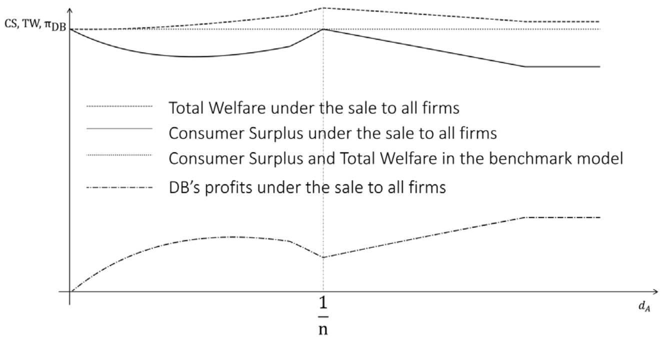

Fig. 5. Consumer Surplus (CS) and Total Welfare (TW) as a function of d_A. $t = 10, F = 0.1, u = 5, \alpha = 0.6, \theta = 1$.

As already noted, for $d_A < \frac{1}{n}$, an increase in d_A has two opposing effects on consumer surplus. On the one hand, the decrease in prices given by the competition effect benefits consumers. On the other hand, the reduction in firms' profits decreases entry, harming consumers. When d_A is small, identified consumers are located close to firms' locations. Then, thanks to their positional advantage, firms can extract most of the surplus from consumers, resulting in a weak competition effect. Thus, for low values of d_A, the entry barrier effect dominates the competition effect, and consumer surplus decreases. However, as d_A increases, informed consumers start being located farther from the firms' locations, leading to a stronger competition effect that dominates over the entry barrier effect. Notably, for $d_A = \frac{1}{n}$, consumer surplus reaches its maximum. Indeed, after this point, additional data does not affect firms' strategies in equilibrium, as they already identify all consumers they serve. However, additional data still affect the firms' willingness to pay for data, as they reduce firms' profits when not buying a partition (dashed line in Fig. 3). Thus, after $d_A = \frac{1}{n}$, an increase in data does not affect the competition effect, whereas it increases the entry barrier effect, harming consumers. Setting $d_A^{P*} = \frac{1}{n}$ thus limits the entry barrier effect, leading to the same level of consumer surplus of the standard Salop (1979) model if $\theta = 1$.[19] Intuitively, a decrease in the DB's market power θ decreases the data price and, in turn, the entry barrier effect for any level of d_A. Thus, although θ does not influence the regulator's equilibrium choice, the increase in firm entry benefits consumers. In particular, for any $\theta < 1$, we find that consumers are better off under the presence of the DB.

With regard to total welfare we find that, if the weight of industry profits α is low enough, welfare is maximized under the same conditions that maximize consumer surplus, namely $d_A^{P*} = \frac{1}{n}$. However, for higher values of α, total welfare is maximized for $d_A^{P*} \geq \frac{3}{2n}$, as it is increasing in d_A. Indeed, in such a case, the welfare loss given by the fixed entry cost is minimized, benefiting the DB who appropriates most of the surplus.

Our analysis thus highlights two possible strategies for a policymaker. If a policymaker gives a sufficiently low weight to industry profits, he should mandate that the DB sells data to all entering firms, *and* sell all available data, without however selling the same data points to different firms. Indeed, mandating the sale of all data maximizes the (positive) direct effect of data, whereas banning the sale of data about any given consumer to multiple firms mitigates the (negative) entry barrier effect of data. Instead, if a policymaker gives a sufficiently high weight to industry profits, it is sufficient to mandate the sale of data to all firms.

8. Conclusions

Digital markets have been substantially growing over the years, prompting policymakers to intensify their focus on regulating data markets both at the data collection stage and in the possible uses of data. Examples include the European GDPR, the Digital Markets Act, the Digital Service Act, and the more recent Data Act. However, such regulations focus on firms that directly collect data from consumers and only tangentially influence the DB industry, which is a cornerstone of the data markets.

Our analysis highlights how an unregulated DB's data sale can indeed lead to consumer harm, especially if the DB is allowed to sell data through a system of auctions. Such harm mainly stems from the *entry barrier effect* of data: as firms must obtain data to remain competitive, they reduce their profits, leading to lower entry and a more concentrated downstream market.

However, we also show that the DB's data sale, when properly regulated, can be Pareto improving. First, the policymaker should enforce that all downstream firms can buy data, as it hinders the DB's ability to extract surplus from firms and, in turn, decrease entry. Second, by imposing the sale of all available data and simultaneously banning the sale of the same data points to multiple firms, the policymaker can limit the entry barrier effect and maximize downstream firm competition, restoring the same level of consumer surplus that is achieved under the Salop (1979) model and increasing total welfare. In our analysis, we have specifically focused on the competitive effects of data, abstracting from the privacy trade-offs related to consumer data collection and use and from explicitly modeling the strategic interaction between competing DBs. Promising avenues for future research could thus be

[19] Our analysis abstracts from the possible privacy costs faced by consumers when their data are used for price discrimination. Intuitively, such a privacy cost would lead the policymaker to reduce the quantity of data sold in the downstream market. For a comprehensive survey on the economics of privacy, see Acquisti et al. (2016).

focusing on the interactions between the privacy and competitive effects of data, to achieve a more comprehensive picture of the trade-offs stemming from the use of data in online markets, and on the effects that competition in the DB industry, together with data complementarities, can have on downstream firm entry.

CRediT authorship contribution statement

Laura Abrardi: Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Carlo Cambini:** Writing – review & editing, Validation, Supervision, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Flavio Pino:** Writing – review & editing, Writing – original draft, Visualization, Resources, Investigation, Formal analysis, Conceptualization.

Acknowledgment

This study was carried out within the “Cyber Resilience ”project - funded by European Union - Next Generation EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell’Università e della Ricerca). This manuscript reflects only the authors’ views and opinions and the Ministry cannot be considered responsible for them.

Appendix

Proof of Lemma 1. The proof proceeds in three steps. In Step 1, we show that if the DB sells data to all firms, he offers a partition set $\mathbf{P} = (\widetilde{d_A}, \widehat{d_A}, \dots, \widehat{d_A})$, regardless of the selling mechanism. In Step 2, we show that under the sale to all firms, the DB sets $\widetilde{d_A} = \widehat{d_A} = d_A$. In Step 3, we find the general partition set offered by the DB under the three selling mechanisms if in equilibrium he wants to sell to half of the entering firms.

Step 1. Suppose that the DB sells to all firms in equilibrium under TIOLI, so that $\mathbf{P}^*_{\mathbf{TIOLI}} = (\widetilde{d_A}, \widehat{d_A}, \dots, \widehat{d_A})$. As under TIOLI each firm independently and simultaneously chooses whether to purchase their respective partition, the DB cannot change the offers he makes to one firm based on another firm’s response. Then, we must have $\mathbf{P_{TIOLI}} = \mathbf{P}^*_{\mathbf{TIOLI}}$.

Suppose instead that the DB sells data to all firms under one of the auction mechanisms, i.e., $\mathbf{P}^*_{\mathbf{AR}} = \mathbf{P}^*_{\mathbf{AU}} = (\widetilde{d_A}, \widehat{d_A}, \dots, \widehat{d_A})$. Then, the DB sets up n auctions, with the aim of concluding all of them. Note that the DB cannot change the partitions he puts up for auction based on firms’ bids, as he can only choose the number of auctions he wants to fulfill. Thus, the only way to obtain $\mathbf{P}^*$ is to offer $\mathbf{P} = \mathbf{P}^*$, and then concluding all of the auctions. This strategy corresponds to the one for TIOLI, implying that the DB’s equilibrium strategy and market outcomes under the auction mechanisms are the same than under the TIOLI sale.

Step 2. Suppose the DB sets $\widehat{d} \neq \widetilde{d}, \widehat{d} > 0, \widetilde{d} > 0$. We show that this strategy is always dominated by a strategy where the DB sets $\widehat{d} = \widetilde{d} = d_A > 0$. To do so, we solve the model under the first strategy. The DB offers a partition set $\mathbf{P} = \left(\widetilde{d}, \widehat{d}, \widetilde{d}, \dots, \widehat{d}\right)$. Without loss of generality, we focus on a generic firm i, to which the DB offers a partition $\widetilde{d}$. The indifferent consumers between firms i, i+1 and i−1 can be obtained by equating utility levels, and they are:

$$\hat{x}_{i-1,i} = \frac{2i-1}{2n} + \frac{p_i^{\mathrm{B}} - p_{i-1}^{\mathrm{B}}}{2t} \qquad \text{and} \qquad \hat{x}_{i,i+1} = \frac{2i+1}{2n} + \frac{p_{i+1}^{\mathrm{B}} - p_i^{\mathrm{B}}}{2t} \tag{A.1}$$

Firm i offers a tailored price $p_i^{\mathrm{T}}(x)$ to the identified consumers, matching the competitor’s offer in utility level. It sets a tailored price for each arc where it competes, resulting in

$$p_i^{\mathrm{T}}(x) = \begin{cases} p_{i-1}^{\mathrm{B}} + 2tx - \frac{t}{n}(2i-1) & \text{for } x \in [\frac{i}{n} - \frac{d_i}{2}, \frac{i}{n}] \\ p_{i+1}^{\mathrm{B}} - 2tx + \frac{t}{n}(2i+1) & \text{for } x \in [\frac{i}{n}, \frac{i}{n} + \frac{d_i}{2}] \end{cases} \tag{A.2}$$

Firm i’s profits are thus given by:

$$\pi_i^{\mathrm{W}}(\mathbf{P}) = \int_{\frac{i}{n}-\frac{\widetilde{d}}{2}}^{\frac{i}{n}} p_i^{\mathrm{T}}(x)\,dx + \int_{\frac{i}{n}}^{\frac{i}{n}+\frac{\widetilde{d}}{2}} p_i^{\mathrm{T}}(x)\,dx + p_i^{\mathrm{B}}(\mathbf{P})\left(\hat{x}_{i,i+1} - \hat{x}_{i-1,i} - \widetilde{d}\right) - F \tag{A.3}$$

Using the expression of the indifferent consumers from (A.1) and of the tailored prices in (A.2), we can rewrite the profits of the generic informed firm i in (A.3) as

$$\begin{aligned} \pi_i^{\mathrm{W}}(\mathbf{P}) = {} & \frac{\widetilde{d}}{2n}\left(2t + \mathrm{np}_{i-1}^{\mathrm{B}}(\mathbf{P}) + \mathrm{np}_{i+1}^{\mathrm{B}}(\mathbf{P}) - nt\widetilde{d}\right) \\ & + p_i^{\mathrm{B}}(\mathbf{P})\left(\frac{n\left(p_{i+1}^{\mathrm{B}}(\mathbf{P}) + p_{i-1}^{\mathrm{B}}(\mathbf{P}) - 2p_i^{\mathrm{B}}(\mathbf{P})\right) + 2t}{2nt} - \widetilde{d}\right) - F \end{aligned} \tag{A.4}$$

Similarly, the profits of its rival i+1 firm are

$$\pi_{i+1}^{\mathrm{W}}(\mathbf{P}) = \frac{\widehat{d}}{2n}\left(2t + np_i^{\mathrm{B}}(\mathbf{P}) + np_{i+2}^{\mathrm{B}}(\mathbf{P}) - nt\widehat{d}\right)$$

$$+p_{i+1}^{\mathrm{B}}(\mathbf{P})\left(\frac{n\left(p_i^{\mathrm{B}}(\mathbf{P})+p_{i+2}^{\mathrm{B}}(\mathbf{P})-2p_{i+1}^{\mathrm{B}}(\mathbf{P})\right)+2t}{2nt}-\hat{d}\right)-F \tag{A.5}$$

By taking the first-order condition of (A.4) with respect to $p_i^{\mathrm{B}}(\mathbf{P})$ and of (A.5) with respect to $p_{i+1}^{\mathrm{B}}(\mathbf{P})$, we obtain firms' reaction function on basic prices

$$\begin{gathered} p_i^{\mathrm{B}}(\mathbf{P})=\frac{t}{2n}-\frac{t\tilde{d}}{2}+\frac{p_{i+1}^{\mathrm{B}}(\mathbf{P})+p_{i-1}^{\mathrm{B}}(\mathbf{P})}{4} \\ \text{and} \\ p_{i+1}^{\mathrm{B}}(\mathbf{P})=\frac{t}{2n}-\frac{t\hat{d}}{2}+\frac{p_{i}^{\mathrm{B}}(\mathbf{P})+p_{i+2}^{\mathrm{B}}(\mathbf{P})}{4} \end{gathered} \tag{A.6}$$

The system of Eqs. (A.6) for all$i=0,\dots,n-1$ allows us to obtain the equilibrium basic prices and, by replacing them in (A.4), firm i's profits. In matrix form we have $\mathbf{A}*\mathbf{p}=\mathbf{b}$, where $\mathbf{p}$ is the price vector, and $\mathbf{b}$ is the known terms vector. Assuming that the DB offers $\tilde{d}$ to even indexed firms, we obtain

$$\begin{bmatrix} 4 & -1 & \dots & 0 & 0 & 0 & \dots & -1 \\ -1 & 4 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 4 & -1 & 0 & \dots & 0 \\ 0 & 0 & \dots & -1 & 4 & -1 & \dots & 0 \\ 0 & 0 & \dots & 0 & -1 & 4 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -1 & 0 & \dots & 0 & 0 & 0 & \dots & 4 \end{bmatrix} * \begin{bmatrix} p_0^{\mathrm{B}}(\mathbf{P}) \\ p_1^{\mathrm{B}}(\mathbf{P}) \\ \dots \\ p_{i-1}^{\mathrm{B}}(\mathbf{P}) \\ p_i^{\mathrm{B}}(\mathbf{P}) \\ p_{i+1}^{\mathrm{B}}(\mathbf{P}) \\ \dots \\ p_{n-1}^{\mathrm{B}}(\mathbf{P}) \end{bmatrix} = \begin{bmatrix} \frac{2t}{n}-2t\tilde{d} \\ \frac{2t}{n}-2t\hat{d} \\ \dots \\ \frac{2t}{n}-2t\hat{d} \\ \frac{2t}{n}-2t\tilde{d} \\ \frac{2t}{n}-2t\hat{d} \\ \dots \\ \frac{2t}{n}-2t\hat{d} \end{bmatrix}$$

Matrix $\mathbf{A}$ is circulant, tridiagonal and symmetric. The inverse of this type of matrix has been computed by Searle (1979). We obtain

$$A^{-1}=\begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-1} & a_0 & \dots & a_{n-2} \\ \dots & \dots & \dots & \dots \\ a_1 & a_2 & \dots & a_0 \end{bmatrix}$$

where, in our specific case, $a_j=-\frac{1}{2\sqrt{3}}*\left(\frac{\left(2+\sqrt{3}\right)^j}{1-\left(2+\sqrt{3}\right)^n}-\frac{\left(2-\sqrt{3}\right)^j}{1-\left(2-\sqrt{3}\right)^n}\right)$. A property of this type of matrix is that $a_j=a_{n-j}\ \forall j\neq 0,\frac{n}{2}$ if n is even. Moreover, in our particular case, $\sum_{j=0}^{n-1}a_j=\frac{1}{2}$. We can now write $\mathbf{p}=\mathbf{A^{-1}}*\mathbf{b}$. We obtain

$$\begin{bmatrix} p_0 \\ p_1 \\ \dots \\ p_{n-1} \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-1} & a_0 & \dots & a_{n-2} \\ \dots & \dots & \dots & \dots \\ a_1 & a_2 & \dots & a_0 \end{bmatrix} * \begin{bmatrix} \frac{2t}{n}-2t\tilde{d} \\ \frac{2t}{n}-2t\hat{d} \\ \dots \\ \frac{2t}{n}-2t\hat{d} \end{bmatrix}$$

Thus, we can write

$$p_i^{\mathrm{B}}=\left(\frac{2t}{n}*\sum_{j=0}^{n-1}a_j\right)-2t\sum_{j=0}^{\frac{n-2}{2}}\tilde{d}a_{2j}-2t\sum_{j=0}^{\frac{n-2}{2}}\hat{d}a_{2j+1}$$

Since $\sum_{j=0}^{n-1}a_j=\frac{1}{2}$, we can simplify and obtain

$$p_i^{\mathrm{B}}=\frac{t}{n}-2t\sum_{j=0}^{\frac{n-2}{2}}\tilde{d}a_{2j}-2t\sum_{j=0}^{\frac{n-2}{2}}\hat{d}a_{2j+1}$$

Due to the symmetry properties of the coefficients a_j, we also obtain a similar form for p_{i-1}^{B} and p_{i-1}^{B}:

$$p_{i-1}^{\mathrm{B}}=p_{i+1}^{\mathrm{B}}=\frac{t}{n}-2t\sum_{j=0}^{\frac{n-2}{2}}\hat{d}a_{2j}-2t\sum_{j=0}^{\frac{n-2}{2}}\tilde{d}a_{2j+1}$$

We find that in our case $\sum_{j=0}^{\frac{n-2}{2}}a_{2j}=\frac{1}{3}$ and $\sum_{j=0}^{\frac{n-2}{2}}a_{2j+1}=\frac{1}{6}$. Thus, we can rewrite basic prices as

$$p_i^{\mathrm{B}}=\frac{t}{n}-\frac{2}{3}t\tilde{d}-\frac{1}{3}t\hat{d} \qquad \text{and} \qquad p_{i-1}^{\mathrm{B}}=p_{i+1}^{\mathrm{B}}=\frac{t}{n}-\frac{2}{3}t\hat{d}-\frac{1}{3}t\tilde{d} \tag{A.7}$$

By replacing the basic prices from (A.7) in firms' profits functions (A.4) and (A.5), we obtain

$$\pi_i^{\mathrm{W}}(\mathbf{P}) = \frac{t}{9n^2}\left(9 - 2n\left(n\widehat{d}\widetilde{d} + \widetilde{d}\left(\frac{7}{4}n\widetilde{d} - 3\right) - n\widehat{d}^2 + 3\widehat{d}\right)\right) - F \tag{A.8}$$

$$\pi_{i-1}^{\mathrm{W}}(\mathbf{P}) = \pi_i^{\mathrm{W}}(\mathbf{P}) = \frac{t}{9n^2}\left(9 - 2n\left(n\widehat{d}\widetilde{d} + \widehat{d}\left(\frac{7}{4}n\widehat{d} - 3\right) - n\widetilde{d}^2 + 3\widetilde{d}\right)\right) - F \tag{A.9}$$

We now compute firms' profits when they do not obtain their partition. Suppose that firm i does not obtain its partition: as such, at the equilibrium $d_i = 0$. By imposing it in (A.4), we obtain that firm i's profits are

$$\pi_i^{\mathrm{L}}(\mathbf{P}) = p_i^{\mathrm{B}}(\mathbf{P})\left(\frac{n\left(p_{i+1}^{\mathrm{B}}(\mathbf{P}) + p_{i-1}^{\mathrm{B}}(\mathbf{P}) - 2p_i^{\mathrm{B}}(\mathbf{P})\right) + 2t}{2nt}\right) - F \tag{A.10}$$

We can again compute firms' basic prices by solving the n-equations system. The only difference from the already analyzed subgame is that firm i's known term has $d_i = 0$ instead of $d_i = \widetilde{d}$. As such, we can compute the new basic prices by simply subtracting $\widetilde{d}a_{i-j}$ from the basic prices p_j^{B} computed in (A.7). Thus, we obtain

$$p_i^{\mathrm{B}} = \frac{t}{n} - 2t\widetilde{d}\left(\frac{1}{3} - a_0\right) - \frac{1}{3}t\widehat{d} \quad \text{and} \quad p_{i-1}^{\mathrm{B}} = p_{i+1}^{\mathrm{B}} = \frac{t}{n} - \frac{2}{3}t\widehat{d} - 2t\widetilde{d}\left(\frac{1}{6} - a_1\right) \tag{A.11}$$

By replacing the basic prices of (A.11) in (A.10), we obtain

$$\pi_i^{\mathrm{L}}(\mathbf{P}) = \frac{t}{9n^2}\left(6a_0n\widetilde{d} - 2n\widetilde{d} - n\widehat{d} + 3\right)\left(-6a_0n\widetilde{d} + 6a_1n\widetilde{d} + n\widetilde{d} - n\widehat{d} + 3\right) - F \tag{A.12}$$

Following the same procedure, we obtain firm i+1's profits in the subgame where it does not obtain data:

$$\pi_{i+1}^{\mathrm{L}}(\mathbf{P}) = \frac{t}{9n^2}\left(6a_0n\widehat{d} - 2n\widehat{d} - n\widetilde{d} + 3\right)\left(-6a_0n\widehat{d} + 6a_1n\widehat{d} + n\widehat{d} - n\widetilde{d} + 3\right) - F \tag{A.13}$$

Finally, we compute DB's profits. We can write them as

$$\pi_{\mathrm{DB}} = \frac{n}{2}\left(\pi_i^{\mathrm{W}}(\mathbf{P}) - \pi_i^{\mathrm{L}}(\mathbf{P})\right) + \frac{n}{2}\left(\pi_{i+1}^{\mathrm{W}}(\mathbf{P}) - \pi_{i+1}^{\mathrm{L}}(\mathbf{P})\right) \tag{A.14}$$

Replacing firms' profits from (A.12) and (A.13) and simplifying, we obtain

$$\begin{aligned}\pi_{\mathrm{DB}} = {} & \frac{t}{3}\left(6na_0^2\left(\widetilde{d}^2 + \widehat{d}^2\right) - 6na_0a_1\left(\widetilde{d}^2 + \widehat{d}^2\right) - 3na_0\left(\widetilde{d}^2 + \widehat{d}^2\right)\right. \\ & \left. + 2na_1\left(\widetilde{d}^2 + \widehat{d}^2 + \widetilde{d}\widehat{d}\right) - 3a_1\left(\widetilde{d} + \widehat{d}\right) - \frac{n}{4}\left(\widetilde{d}^2 + \widehat{d}^2\right) - n\widetilde{d}\widehat{d} + \frac{3}{2}\widetilde{d} + \frac{3}{2}\widehat{d}\right)\end{aligned} \tag{A.15}$$

By computing FOCs of (A.15) for both $\widetilde{d}$ and $\widehat{d}$, we find that both partitions have the same effect on DB's profits; to maximize them, the DB would set $\widetilde{d} = \widehat{d}$. However, since $\widetilde{d} \neq \widehat{d}$ by hypothesis, we find that setting $\widetilde{d} \neq \widehat{d}$ is suboptimal for the DB.

Step 3. As in Step 1, the only strategy for the DB to sell $\mathbf{P}^*_{\mathbf{TIOLI}}$ in equilibrium is to set $\mathbf{P}_{\mathbf{TIOLI}} = \mathbf{P}^*_{\mathbf{TIOLI}}$. Thus, the DB sets $\mathbf{P}^{\mathbf{TIOLI}}_{\mathbf{H}} = (d_H^{TIOLI}, 0, \dots, d_H^{TIOLI}, 0)$. However, we prove that such a strategy is suboptimal for the DB. Abrardi et al. (2024) have already shown that in the same setting, except that the data sale occurs after firm entry, in equilibrium the DB opts for the sale to all firms under TIOLI. If the data sale and entry occur simultaneously, like in our model, the DB takes into account how his data sale affects his profits, and maximizes them accordingly. Moreover, Abrardi et al. (2024) show that (i) the DB's profits are inversely proportional to the number of entering firms, (ii) the number of entering firms is directly proportional to their profits after paying for data, and (iii) the firms' profits after paying for data are decreasing with the total quantity of data sold in the downstream market. Then, to maximize his profits, the DB has an incentive to reduce firm entry, which he can do by selling more data in the downstream market. Then, we conclude that the timing of our model further exacerbates the advantages of selling to all firms than to half of the firms under TIOLI.

With regards to AR, in equilibrium, the DB offers a partition set $\mathbf{P}^*_{\mathrm{AR}} = (d_H, 0, d_H, 0, \dots, 0)$. However, the DB can still set up an auction for firms who will not obtain data in equilibrium, and then he does not fulfill them. We show that the DB offers the whole dataset in these auctions, and thus the offered partition set is $\mathbf{P}_{\mathrm{AR}} = (d_H, 1, d_H, 1, \dots, 1)$.

The DB's profits are

$$\pi_{\mathrm{DB}}(\mathbf{P}_{\mathrm{AR}}, \mathbf{J}) = \sum_{i \in \mathbf{J}}\left(\pi_i^{\mathrm{W}}(\mathbf{P}_{\mathrm{AR}}) - \pi_i^{\mathrm{L}}(\mathbf{P}_{\mathrm{AR}})\right). \tag{A.16}$$

Consider the set of data partitions $\{d_j\}_{j \notin \mathbf{J}}$. Since these partitions are offered in the auction, but are not actually sold in equilibrium, then $\pi_i^{\mathrm{W}}(\mathbf{P}_{\mathrm{AR}}) = \pi_i^{\mathrm{W}}(\mathbf{P}^*_{\mathrm{AR}})$ for all $i \in \mathbf{J}$, i.e., $\pi_i^{\mathrm{W}}(\mathbf{P}_{\mathrm{AR}})$ does not depend on $\{d_j\}_{j \notin \mathbf{J}}$. Hence, the DB chooses $\{d_j\}_{j \notin \mathbf{J}}$ so as to minimize $\pi_i^{\mathrm{L}}(\mathbf{P}_{\mathrm{AR}})$, given that doing so does not affect $\pi_i^{\mathrm{W}}(\mathbf{P}_{\mathrm{AR}})$.

Firm i's losing profits are

$$\pi_i^{\mathrm{L}}(\mathbf{P}_{\mathrm{AR}}) = p_i^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}})\left(\frac{n\left(p_{i+1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}}) + p_{i-1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}}) - 2p_i^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}})\right) + 2t}{2nt}\right) - F. \tag{A.17}$$

Eq. (A.17), for any given p_i^{B}, is minimized when both $p_{i-1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}})$ and $p_{i+1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}})$ are minimized, that is, due to the non-negative constraint on prices,

$$p_{i-1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}}) = p_{i+1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}}) = 0.$$

This can be achieved by setting $d_{i-1} = d_{i+1} = 1$. In fact, focus on firm $i+1$. Then, its profit can be expressed as

$$\pi_{i+1}^{\mathrm{W}}(\mathbf{P}_{\mathrm{AR}}) = \int_{\hat{x}_{i,i+1}}^{\hat{x}_{i+1,i+2}} p_{i+1}^{\mathrm{T}}(x)\,dx - F. \tag{A.18}$$

Firm $i+1$ chooses $p_{i+1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}})$ so as to maximize (A.18). To this aim, note that function (A.18) is strictly decreasing in $p_{i+1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}})$. In fact, $p_{i+1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}})$ affects $\hat{x}_{i,i+1}$ and $\hat{x}_{i+1,i+2}$ but not $p_{i+1}^{\mathrm{T}}(\mathbf{P}_{\mathrm{AR}})$ in (A.18). In particular, a decrease of $p_{i+1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}})$ expands firm $i+1$'s market share by moving further away the two indifferent consumers. Given that function (A.18) is strictly decreasing in $p_{i+1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}})$, then he sets $p_{i+1}^{\mathrm{B}}(\mathbf{P}_{\mathrm{AR}}) = 0$. The same argument holds for $i-1$ by symmetry.

With regard to AU, the DB can again offer a partition set $\mathbf{P}_{\mathrm{AU}}$ which is different from $\mathbf{P}_{\mathrm{AU}}^{*}$, as he can decide to not fulfill some of the auctions he sets up. However, without reserve prices, firms can win their auctions by beating their rivals' offers. Hereafter, we prove that the DB offers a partition set $\mathbf{P}_{\mathrm{AU}} = (d_H, d_H, \dots, d_H)$.

The DB can offer a generic partition set $\mathbf{P}_{\mathrm{AU}} = (d_H, d_1, d_H, d_3, \dots, d_H, d_{n-1})$. If firm 0 deviates, the DB would want to fulfill the auctions where he offers d_{n-1} and d_1. As d_{n-1} and d_1 have the same effect on firm 0's profits, the DB would set $d_{n-1} = d_1$. The same holds for any deviating firm, and thus the DB sets $d_1 = d_3 = \cdots = d_{n-1} = d$. If no even-indexed firm deviates, in equilibrium the DB fulfills the auctions where he offered d_H. If one even-indexed firm deviates, the DB can instead fulfill all the auctions where he offered d. Without loss of generality, we focus on firms 0 and 1. Firms' willingness to pay for data are equal to

$$\pi_0^{\mathrm{W}}\left(d_H, 0, d_H, 0, \dots, d_H, 0\right) - \pi_0^{\mathrm{L}}(0, d, 0, d, \dots, 0, d)$$
and
$$\pi_1^{\mathrm{W}}(0, d, 0, d, \dots, 0, d) - \pi_1^{\mathrm{L}}(d_H, 0, d_H, 0, \dots, d_H, 0).$$

Suppose that $d > d_H$ (the opposite case is solved similarly): then it is straightforward to show that

$$\pi_1^{\mathrm{W}}(0, d, 0, d, \dots, 0, d) - \pi_1^{\mathrm{L}}(d_H, 0, d_H, 0, \dots, d_H,)0 > \\ \pi_0^{\mathrm{W}}\left(d_H, 0, d_H, 0, \dots, d_H, 0\right) - \pi_0^{\mathrm{L}}(0, d, 0, d, \dots, 0, d),$$

that is, firm 1's willingness to pay is higher than firm 0's one. Because there are no reserve prices, firm 1 can win its auction by offering firm 0's willingness to pay plus ε, where ε is an arbitrary small number. As such, the DB chooses d as low as possible to maximize firm 0's willingness to pay and, in turn, firm 1's: that is, he chooses $d = d_H$.

Proof of Lemma 2. We start our analysis from the sale to all firms. Suppose that the indifferent consumers are not identified. A generic firm i's profits are

$$\pi_i^{\mathrm{W}}(\mathbf{P}_{\mathbf{A}}) = \int_{\frac{i}{n}-\frac{d_A}{2}}^{\frac{i}{n}+\frac{d_A}{2}} p_i^{\mathrm{T}}(x)\,dx + p_i^{\mathrm{B}}\left(\hat{x}_{i,i+1} - \hat{x}_{i-1,i} - d_A\right) - F, \tag{A.19}$$

where

$$\hat{x}_{i-1,i} = \frac{2i-1}{2n} + \frac{p_i^{\mathrm{B}} - p_{i-1}^{\mathrm{B}}}{2t} \qquad \text{and} \qquad \hat{x}_{i,i+1} = \frac{2i+1}{2n} + \frac{p_{i+1}^{\mathrm{B}} - p_i^{\mathrm{B}}}{2t} \tag{A.20}$$

Are the indifferent consumer locations, and

$$p_i^{\mathrm{T}}(x) = \begin{cases} p_{i-1}^{\mathrm{B}} + 2tx - \frac{t}{n}(2i-1) & \text{for } x \in [\frac{i}{n} - \frac{d_i}{2}, \frac{i}{n}] \\ p_{i+1}^{\mathrm{B}} - 2tx + \frac{t}{n}(2i+1) & \text{for } x \in [\frac{i}{n}, \frac{i}{n} + \frac{d_i}{2}] \end{cases} \tag{A.21}$$

Are the tailored prices set by firm i to match the direct rivals' basic prices, adjusted for the transportation cost.

By computing FOC of (A.19) with respect to p_i^B, we find

$$4p_i^{\mathrm{B}} - p_{i-1}^{\mathrm{B}} - p_{i+1}^{\mathrm{B}} = \frac{2t}{n} - 2td_A,$$

$\forall i \in \{0, \dots, n-1\}$. The firms' equilibrium basic prices are obtained by solving the system composed by the above n equations. The system in matricial form is expressed by $\mathbf{A} * \mathbf{p} = \mathbf{b}$, where $\mathbf{p}$ is the vector containing basic prices, and $\mathbf{b}$ is the vector containing the known terms:

$$\begin{bmatrix} 4 & -1 & \dots & 0 & 0 & 0 & \dots & -1 \\ -1 & 4 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 4 & -1 & 0 & \dots & 0 \\ 0 & 0 & \dots & -1 & 4 & -1 & \dots & 0 \\ 0 & 0 & \dots & 0 & -1 & 4 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -1 & 0 & \dots & 0 & 0 & 0 & \dots & 4 \end{bmatrix} * \begin{bmatrix} p_0^{\mathrm{B}} \\ p_1^{\mathrm{B}} \\ \dots \\ p_{i-1}^{\mathrm{B}} \\ p_i^{\mathrm{B}} \\ p_{i+1}^{\mathrm{B}} \\ \dots \\ p_{n-1}^{\mathrm{B}} \end{bmatrix} = \begin{bmatrix} \frac{2t}{n} - 2td_A \\ \frac{2t}{n} - 2td_A \\ \dots \\ \frac{2t}{n} - 2td_A \\ \frac{2t}{n} - 2td_A \\ \frac{2t}{n} - 2td_A \\ \dots \\ \frac{2t}{n} - 2td_{n-1} \end{bmatrix}$$

Matrix **A** is circulant, tridiagonal and symmetric. Given a general circulant tridiagonal matrix of form $\mathbf{M} = (a, b, 0, 0, \dots, 0, c)$, where a, b, c express the non-null elements of the first line, the general expression of its inverse is provided by Searle (1979), and it is given by

$$\mathbf{A}^{-1} = \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-1} & a_0 & \dots & a_{n-2} \\ \dots & \dots & \dots & \dots \\ a_1 & a_2 & \dots & a_0 \end{bmatrix}$$

where $a_j = \frac{z_1 z_2}{b(z_1 - z_2)} \left(\frac{z_1^j}{1 - z_1^n} - \frac{z_2^j}{1 - z_2^n} \right)$ and $z_1, z_2 = \frac{\sqrt{-a \pm (a^2 - 4bc)}}{2c}$. In our case, as $a = 4, b = -1$ and $c = -1$, we obtain that $a_j = -\frac{1}{2\sqrt{3}} \left(\frac{\left(2+\sqrt{3}\right)^j}{1-\left(2+\sqrt{3}\right)^n} - \frac{\left(2-\sqrt{3}\right)^j}{1-\left(2-\sqrt{3}\right)^n} \right)$.

Using $\mathbf{A}^{-1}$, we obtain the equilibrium basic prices through $\mathbf{p} = \mathbf{A}^{-1} * \mathbf{b}$:

$$\begin{bmatrix} p_0^{\mathrm{B}*} \\ p_1^{\mathrm{B}*} \\ \dots \\ p_{n-1}^{\mathrm{B}*} \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-1} & a_0 & \dots & a_{n-2} \\ \dots & \dots & \dots & \dots \\ a_1 & a_2 & \dots & a_0 \end{bmatrix} * \begin{bmatrix} \frac{2t}{n} - 2td_A \\ \frac{2t}{n} - 2td_A \\ \dots \\ \frac{2t}{n} - 2td_A \end{bmatrix} \tag{A.22}$$

Thus, equilibrium basic prices are

$$p_i^{\mathrm{B}*} = \left(\frac{2t}{n} * \sum_{j=0}^{n-1} a_j \right) - 2t \sum_{j=0}^{n-1} d_A a_j. \tag{A.23}$$

A useful property of the matrix $\mathbf{A}^{-1}$ in our framework is that $\sum_{j=0}^{n-1} a_j = \frac{1}{2}$, so that (A.23) can be simplified as

$$p_i^{\mathrm{B}*}(\mathbf{P_A}) = \frac{t}{n} - td_A \tag{A.24}$$

As all basic prices are equal, indifferent consumers are located in the middle between firms' locations, i.e., $\hat{x}_{i,i+1} = \frac{2i+1}{2n}$. By substituting the equilibrium prices in (A.19), we obtain firms' profits under non-overlapping partitions:

$$\pi_i^{W*}(\mathbf{P_A}) = \frac{t}{n^2} - \frac{td^2}{2} - F. \tag{A.25}$$

Firms identify all consumers they serve when $\frac{i}{n} + \frac{d}{2} \geq \hat{x}_{i,i+1}$, which we can rewrite as $d \geq \frac{1}{n}$. In this case, firm i's profits are

$$\pi_i^{\mathrm{W}}(\mathbf{P_A}) = \int_{\hat{x}_{i-1,i}}^{\hat{x}_{i,i+1}} p_i^{\mathrm{T}}(x)\, dx - F. \tag{A.26}$$

Each firm's best strategy is then to minimize p_i^B to try and expand its market share. Then, all firms set $p_i^{B*} = 0$, and we obtain

$$\pi_i^{W*}(\mathbf{P_A}) = \frac{t}{2n^2} - F. \tag{A.27}$$

Let us now consider the sale to half of the firms under AR and AU. As in equilibrium the DB offers the same partition set, where half of the firms are informed, this step of the analysis is the same for both mechanisms. Without loss of generality, we focus on the AR case, but the same applies for the AU case.

In equilibrium, the DB sells a partition set $\mathbf{P}_{\mathrm{AR}}^* = (d, 0, d, 0, \dots, 0)$. Suppose that all firms who can obtain a partition of size d_{AR} centered on their location win their respective auctions, and that firm i is one of those. By following the same procedure as in Step 3, equilibrium basic prices are

$$p_{\mathrm{i}}^{\mathrm{B}*}\left(\mathbf{P}_{\mathrm{AR}}^*\right) = \frac{t}{n} - 2td_{\mathrm{AR}} \left(a_0 + a_{\frac{n}{2}} + 2 \sum_{j=1}^{\frac{n}{4}-1} a_{2j} \right) = \frac{t}{n} - \frac{2}{3} td_{\mathrm{AR}} \tag{A.28}$$

and

$$p_{i+1}^{B*}\left(\mathbf{P}_{AR}^{*}\right)=p_{i-1}^{B*}\left(\mathbf{P}_{AR}^{*}\right)=\frac{t}{n}-\frac{1}{3}td_{AR}. \tag{A.29}$$

Substituting (A.28) and (A.29) in (A.20), we obtain the indifferent consumers' locations:

$$\hat{x}_{i-1,i}=\frac{2i-1}{2n}-\frac{d_{AR}}{6} \quad\text{and}\quad \hat{x}_{i,i+1}=\frac{2i+1}{2n}+\frac{d_{AR}}{6}. \tag{A.30}$$

We can compute firm i's profits by substituting (A.28), (A.29) and (A.30) in (A.19), obtaining

$$\pi_i^{W*}\left(\mathbf{P}_{AR}^{*}\right)=\frac{t}{n^2}+\frac{2d_{AR}t}{3n}-\frac{7td_{AR}^2}{18}-F. \tag{A.31}$$

Similarly, firm $i+1$'s profits are

$$\pi_{i+1}^{L*}\left(\mathbf{P}_{AR}^{*}\right)=\frac{t}{n^2}-\frac{2d_{AR}t}{3n}+\frac{td_{AR}^2}{9}-F. \tag{A.32}$$

We now focus on the case where winning firms only serve identified consumers. By comparing $\frac{i}{n}+\frac{d_{AR}}{2}$ and $\hat{x}_{i,i+1}$, we find that firms only serve identified consumers when

$$d_H \geq \frac{3}{2n}. \tag{A.33}$$

When (A.33) holds, firm i sets its basic price equal to 0 and its profits are given by (A.18). We find $p_{i+1}^{B*}\left(\mathbf{P}_{AR}^{*}\right)$ by solving the FOCs of (A.17) with $p_{i+2}^{B*}\left(\mathbf{P}_{AR}^{*}\right)=p_i^{B}\left(\mathbf{P}_{AR}^{*}\right)=0$, obtaining $p_{i+1}^{B*}\left(\mathbf{P}_{AR}^{*}\right)=\frac{t}{2n}$. The same logic applies to firm $i-1$ by symmetry. By substituting the basic prices in the profits functions, we obtain

$$\pi_i^{W*}\left(\mathbf{P}_{AR}^{*}\right)=\frac{9t}{8n^2}-F \quad\text{and}\quad \pi_{i+1}^{L*}\left(\mathbf{P}_{AR}^{*}\right)=\frac{t}{4n^2}-F.$$

Proof of Lemma 3. The proof proceeds in two steps. First, we focus on firms' losing profits under the sale to all firms. Second, we focus on firms' losing profits under AR and AU.

Step 1. We now focus on the subgame in which firm i does not buy data. Its profits are then

$$\pi_i^{L}(\mathbf{P_A})=p_i^{B}\left(\hat{x}_{i,i+1}-\hat{x}_{i-1,i}\right)-F, \tag{A.34}$$

while all other firms' profits are as in (A.19). We have to consider three separate cases: i) all informed firms serve both identified and unidentified consumers, ii) all informed firms except firm i's direct rivals only serve identified consumers and iii) all informed firms only serve identified consumers.

(i) When firm i does not buy data, it becomes the only uninformed firm in the market, and its profits are

$$\pi_i^{L}(\mathbf{P_A})=p_i^{B}\left(\hat{x}_{i,i+1}-\hat{x}_{i-1,i}\right)-F. \tag{A.35}$$

To find equilibrium prices, we again solve the system of FOCs. The only difference with respect to the previous case is that the ith component of vector $\mathbf{b}$ becomes $\frac{2t}{n}$. We thus obtain

$$p_i^{B\ L*}\left(\mathbf{P_A}\right)=\frac{t}{n}-td_A+2td_Aa_0 \quad\text{and}\quad p_{i-j}^{B\ L*}\left(\mathbf{P_A}\right)=p_{i+j}^{B\ L*}\left(\mathbf{P_A}\right)=\frac{t}{n}-td_A+2td_Aa_j. \tag{A.36}$$

By substituting (A.36) in (A.35), we find

$$\pi_i^{L*}\left(\mathbf{P_A}\right)=\left(\frac{t}{n}-td_A+2td_Aa_0\right)\left(2d_A\left(a_1-a_0\right)+\frac{1}{n}\right)-F. \tag{A.37}$$

From (A.36), we know that informed firms set different equilibrium basic prices, depending on their distance from firm i. In particular, basic prices are higher, the closer a firm is to firm i. Let us focus on the indifferent consumer between firms $i-2$ and $i-1$. Using (A.36), we obtain

$$\hat{x}_{i-2,i-1}=\frac{2i-3}{2n}+d(a_1-a_2).$$

Firm $i-2$ can identify consumers up to $\frac{i-2}{n}+\frac{d}{2}$. Then, if

$$d\geq d_1\equiv\frac{1}{2n(\frac{1}{2}+a_1-a_2)},$$

firm $i-2$ only serves identified consumers and sets its basic price equal to 0. As (a_j-a_{j+1}) decreases with j, all other informed firms except $i+1$ and $i-1$ also set their basic prices equal to 0. Thus, this case only holds as long as $d<d_1$.

(ii) Without loss of generality, we focus on firms $i-1$ and i. If $d\geq d_1$, firm $i-1$ identifies all consumers on the arch it shares with firm $i-2$, whereas it still serves some unidentified consumers on the arch it shares with firm i. We can write firm $i-1$'s profits as

$$\pi_{i-1}^{W}\left(\mathbf{P_A}\right)=\int_{\hat{x}_{i-2,i-1}}^{\frac{i-1}{n}}p_{i-1,i-2}^{T}(x)\,dx+\int_{\frac{i-1}{n}}^{\frac{i-1}{n}+\frac{d_A}{2}}p_{i-1,i}^{T}(x)\,dx$$

$$+p_{i-1}^{\mathrm{B}}\left(\mathbf{P_A}\right)\left(\hat{x}_{i-1,i}-\frac{i-1}{n}-\frac{d_A}{2}\right)-F. \tag{A.38}$$

Firm i's profits are given by (A.35). The FOCs of (A.38) and (A.35) give us the equilibrium basic prices:

$$p_{i-1}^{\mathrm{B}*}\left(\mathbf{P_A}\right)=\frac{t\left(3-2nd_A\right)}{5n}\quad\text{and}\quad p_i^{\mathrm{B\ L}*}\left(\mathbf{P_A}\right)=\frac{t\left(4-nd_A\right)}{5n}. \tag{A.39}$$

Substituting (A.39) into (A.35), we obtain

$$\pi_i^{\mathrm{L}*}\left(\mathbf{P_A}\right)=\frac{t(nd_A-4)^2}{25n^2}-F. \tag{A.40}$$

From (A.39), informed firms set positive basic prices as long as $d_A<\frac{3}{2n}$. After this threshold, firms $i-1$ and $i+1$ identify all the consumers they serve, and thus set their equilibrium basic prices equal to zero.

(iii) Suppose $d_A\geq\frac{3}{2n}$. Then, firms $i+1$ and $i-1$ identify all the consumers they serve, and set their basic prices equal to zero. In turn, firm i's profits are given by (A.35), and only depend on its basic price. The FOC of (A.35) with respect to p_i^B and imposing $p_{i+1}^B=p_{i-1}^B=0$ leads to

$$p_i^{\mathrm{B}*}\left(\mathbf{P_A}\right)=\frac{t}{2n}\quad\text{and}\quad\pi_i^{\mathrm{L}*}\left(\mathbf{P_A}\right)=\frac{t}{4n^2}-F. \tag{A.41}$$

Step 2.

Suppose that firm i loses its auction under AR. As the DB can fulfill up to $k=\frac{n}{2}+1$ auctions, he can now fulfill both firm $i+1$ and $i-1$'s auctions, that thus obtain the whole dataset. This subgame is the same as the case where firm i wins the auction and $d_{\mathrm{AR}}\geq\frac{3}{2n}$. As such, firm i's basic price and profits are equal to firm $i+1$'s ones in the previous subgame, leading to

$$p_i^{\mathrm{B}*}\left(\mathbf{P}_{\mathrm{AR}}\right)=\frac{t}{2n}\quad\text{and}\quad\pi_i^{\mathrm{L}*}\left(\mathbf{P}_{\mathrm{AR}}\right)=\frac{t}{4n^2}-F.$$

With regard to AU, when firm i loses its auction, its basic price and profits are the same as firm $i+1$'s in the auction with reserve prices when firm i wins its auction (see (A.32)).

Proof of Proposition 1. The proof proceeds in three steps. In Step 1, we solve the game when the DB sells data to all entering firms, which corresponds to the equilibrium under TIOLI. In Step 2, we assess the DB's strategy under auction with reserve prices (AR). In Step 3, we assess the DB's strategy under auction without reserve prices (AU).

Step 1. The DB solves the following maximization problem

$$\max_{d_A}\pi_{DB}(\mathbf{P_A})=n\left(\pi_i^{\mathrm{W}*}\left(\mathbf{P_A}\right)-\pi_i^{\mathrm{L}*}\left(\mathbf{P_A}\right)\right) \tag{A.42}$$

$$s.t.\quad\pi_i^{L*}(\mathbf{P_A})=0.$$

By combining the expressions for $\pi_i^{\mathrm{W}*}(\mathbf{P_A})$ (in (A.28) and (A.29)) and $\pi_i^{\mathrm{L}*}(\mathbf{P_A})$ (in (A.37), (A.40) and (A.41)) above, we obtain the function of DB's profits:

$$\max_{d_A}\pi_{\mathrm{DB}}=\begin{cases}n\left(\frac{t}{n^2}-\frac{td_A^2}{2}-\left(\frac{t}{n}-td_A+2td_Aa_0\right)\left(2d_A\left(a_1-a_0\right)+\frac{1}{n}\right)\right) & \text{for } d_A<d_1\\ n\left(\frac{t}{n^2}-\frac{td_A^2}{2}-\frac{t(nd_A-4)^2}{25n^2}\right) & \text{for } d_1\leq d_A<\frac{1}{n}\\ n\left(\frac{t}{2n^2}-\frac{t(nd_A-4)^2}{25n^2}\right) & \text{for } \frac{1}{n}\leq d_A<\frac{3}{2n}\\ n\left(\frac{t}{4n^2}\right) & \text{for } d_A\geq\frac{3}{2n},\end{cases} \tag{A.43}$$

which is continuous in d_A.

First, we show that the strategies where the DB sets $d_1\leq d_A<\frac{3}{2n}$ are always dominated by the strategy where the DB sets $d_A\geq\frac{3}{2n}$. DB's profits are given by firms' willingness to pay times the number of entering firms. The DB chooses his strategy by anticipating how d_A influences firm entry and, in turn, his profits. From the FOC of DB's profits, they are decreasing in n. Firm entry is determined by firms' profits when losing, and is minimized for $d_A\geq\frac{3}{2n}$. Thus, strategies where $d_1\leq d_A<\frac{3}{2n}$ can only be dominant if they allow to extract more surplus from individual firms. Suppose that n is given. By computing the FOC of the second part with respect to d_A, we find that it is monotonically decreasing in d_A over its domain. By computing the FOC of the third part with respect to d_A, we find that it is monotonically increasing in d_A over its domain. Thus, as the DB's profits are continuous, the DB profits in the fourth part are constant and higher than those in the third one. Finally, by comparing DB's profits for $d_A=d_1$ and $d_A\geq\frac{3}{2n}$, we find that DB's profits are maximized for $d_A\geq\frac{3}{2n}$ if n is given. As the reduction in entry further improves this strategy's profitability, we conclude that in equilibrium the DB either sells $d_A<d_1$ or $d_A\geq\frac{3}{2n}$.

Second, we show that the strategy where the DB sets $d_A\geq\frac{3}{2n}$ always dominates the one where he sets $d_A\leq d_1$. We refer to the DB setting $d_A\geq\frac{3}{2n}$ as d_{high}, whereas we refer to the DB setting $d_A<d_1$ as d_{low}. When the DB sets $d_A\geq\frac{3}{2n}$, he maximizes

$$\pi_{\mathrm{DB}}(\mathbf{P}_{\mathbf{A}}^{*\,d_{\mathrm{high}}})=\frac{t}{4n}\quad s.t.\quad\frac{t}{4n^2}-F\geq 0. \tag{A.44}$$

By binding the constraint we obtain

$$n^*_{d_{\text{high}}} = \frac{1}{2}\sqrt{\frac{t}{F}}. \tag{A.45}$$

By replacing (A.45) in (A.44) we obtain

$$\pi_{\text{DB}}(\mathbf{P}^{*\,d_{\text{high}}}_{\mathbf{A}}) = \frac{1}{2}\sqrt{tF}. \tag{A.46}$$

When the DB sets $d = d_{\text{low}}$, he maximizes

$$\begin{aligned} &\pi_{\text{DB}}(\mathbf{P}^{*\,d_{\text{low}}}_{\mathbf{A}}) = t\left(d_{\text{low}}\left(1-2a_1\right) - n\frac{d^2_{\text{low}}}{2}\left(1+4\left(1-2a_0\right)\left(a_0-a_1\right)\right)\right) \\ &\quad s.t. \quad \left(\frac{t}{n} - td_{\text{low}} + 2td_{\text{low}}a_o\right)\left(2d_{\text{low}}\left(a_1-a_0\right)+\frac{1}{n}\right) - F \geq 0. \end{aligned} \tag{A.47}$$

We want to show that

$$\pi_{\text{DB}}({\mathbf{P}_{\mathbf{A}}}^{d_{\text{low}}}) < \pi_{\text{DB}}({\mathbf{P}_{\mathbf{A}}}^{d_{\text{high}}}), \tag{A.48}$$

for all relevant values of d_{low}, t and F. In particular, we recall that $0 \leq d_{\text{low}} < \frac{1}{n}$, $t > 0$, $F > 0$, $t > F$. It is useful to express $\frac{F}{t} = k$, with $0 < k < 1$. We can rewrite (A.48) as

$$2d_{\text{low}}\left(1-2a_1\right) - nd^2_{\text{low}}\left(1+4\left(1-2a_0\right)\left(a_0-a_1\right)\right) < \sqrt{k}. \tag{A.49}$$

To solve (A.49), we bind the constraint in (A.47), find the number of entering firms and substitute it in (A.49). The constraint in (A.47) has no explicit solution. We thus want to find an approximated solution of n that overestimates the left-side of (A.49). By showing that the left side is smaller than the right side even after the round ups, we prove that also the original inequality holds.

The number of entering firms is given by binding the constraint in (A.47):

$$\pi_i{}^{\text{L}}({\mathbf{P}_{\mathbf{A}}}^{d_{\text{low}}}) = \left(\frac{t}{n} - td_{\text{low}} + 2td_{\text{low}}a_o\right)\left(2d_{\text{low}}\left(a_1-a_0\right)+\frac{1}{n}\right) - F = 0. \tag{A.50}$$

By substituting the explicit forms of a_0 and a_1 in (A.50), we can rewrite it as

$$\frac{1}{n^2} - \frac{2d_{\text{low}}}{\sqrt{3}n}f(n) + \frac{1}{3}d^2_{\text{low}}f(n)^2 - \frac{F}{t} = 0, \tag{A.51}$$

where

$$f(n) = \frac{\left(\sqrt{3}-1\right)\left(2+\sqrt{3}\right)^n + \left(1+\sqrt{3}\right)\left(2-\sqrt{3}\right)^n - 2\sqrt{3}}{\left(2+\sqrt{3}\right)^n + \left(2-\sqrt{3}\right)^n - 2}. \tag{A.52}$$

Although there is no explicit solution $n(d_{\text{low}})$, the expression is a second-order polynomial in d_{low}: then, we can obtain an explicit solution for $d(n)$. Solving (A.51) with respect to d we obtain

$$d^*_1(n) = \frac{\sqrt{3}\left(n\sqrt{k}+1\right)}{nf(n)} \quad \text{and} \quad d^*_2(n) = -\frac{\sqrt{3}\left(1-n\sqrt{k}\right)}{nf(n)}. \tag{A.53}$$

From Salop (1979) we know that, if a DB is absent, the number of entering firms is $n = \sqrt{\frac{t}{F}}$. As such, our solution must satisfy $d_{\text{low}}\left(\sqrt{\frac{t}{F}}\right) = 0$, which gives us $d_{\text{low}} = d^*_2(n)$. Having found $d_{\text{low}}(n)$, we need to invert the function to obtain $n(d_{\text{low}})$. To do so, we approximate $f(n)$ to find an explicit form of $n(d_{\text{low}})$. We recall that we want to round up $\pi_{\text{DB}}({\mathbf{P}_{\mathbf{A}}}^{d_{\text{low}}})$, which is inversely proportional to n. As such, we need to round down $n(d_{\text{low}})$, which requires rounding up $f(n)$. We find

$$f(n) \approx 0.6197\frac{1.0489n - 1.0566}{0.7806n - 0.4757}, \tag{A.54}$$

which overestimates $f(n)\ \forall\ n \geq 2$.

We can now substitute (A.53) in (A.51), obtaining

$$\begin{aligned} &n^2\left(0.6197 + 1.0489d_{\text{low}} + 0.7806\sqrt{3k}\right) \\ &- n\left(0.6197 * 1.0566d_{\text{low}} + 0.7806\sqrt{3} + 0.4757\sqrt{3k}\right) + 0.4757\sqrt{3} = 0, \end{aligned}$$

which has two solutions:

$$\begin{aligned} n\left(d_{\text{low}}\right) \;=\; &\frac{0.66d_{\text{low}} + 0.4757\sqrt{3k} + 0.7806\sqrt{3}}{1.3d + 1.56\sqrt{3k}} \\ &\pm \frac{1.31\sqrt{-0.277\sqrt{3}\left(2.6d_{\text{low}} + 3.12\sqrt{3k}\right) + \left(0.5d_{\text{low}} + 0.363\sqrt{3k} - 0.597\sqrt{3}\right)^2}}{2.6d_{\text{low}} + 1.56\sqrt{3k}}. \end{aligned}$$

Given that $n(0)=\sqrt{\frac{t}{F}}=\sqrt{\frac{1}{k}}$, the correct solution is the one with the positive sign. We have obtained $n\left(d_{\text{low}}\right)$ rounded down, which in turn rounds up $\pi_{\text{DB}}(\mathbf{P_A}^{d_{\text{low}}})$. Next, we round up the exponential terms present in the left side of (A.49) to increase it. We first focus on $\left(1-2a_1\right)$. This function is monotonically increasing in n, and its limit is

$$\lim_{n\to\infty}\left(1-2a_1\right)=2-\frac{2}{\sqrt{3}}.$$

As $\left(1-2a_1\right)$ increases the left side of (A.49), we approximate

$$\left(1-2a_1\right)\approx 2-\frac{2}{\sqrt{3}}. \tag{A.55}$$

Next, we focus on $\left(1+4\left(1-2a_0\right)\left(a_0-a_1\right)\right)$. This function is monotonically increasing in n, and decreases the left side of (A.49). We find

$$\left(1+4\left(1-2a_0\right)\left(a_0-a_1\right)\right)\approx 1.36\frac{8n-1}{8n+2}, \tag{A.56}$$

which underestimates the function and thus overestimates the left-side of (A.49). By replacing (A.55) and (A.56) in (A.49) and setting $n=n\left(d_{\text{low}}\right)$ we obtain

$$\left(2-\frac{2}{\sqrt{3}}\right)2d_{\text{low}}-n\left(d_{\text{low}}\right)d_{\text{low}}^2\left(1.36\frac{8n\left(d_{\text{low}}\right)-1}{8n\left(d_{\text{low}}\right)+2}\right)-\sqrt{k}<0, \tag{A.57}$$

which is always satisfied for $0<d_{\text{low}}<\frac{1}{n(d_{\text{low}})}$, and $0<k<1$. Thus, the DB's equilibrium strategy under TIOLI is setting $d_A^{TIOLI*}\geq\frac{3}{2n^*}$.

Step 2.

DB's profits under AR can be computed as

$$\max_{d_{\text{AR}}}\pi_{\text{DB}}=\frac{n}{2}\left(\pi_i^{\text{W}*}\left(\mathbf{P}_{\text{AR}}\right)-\pi_i^{\text{L}*}\left(\mathbf{P}_{\text{AR}}\right)\right), \tag{A.58}$$

$$s.t.\quad \pi_i^{L*}(\mathbf{P_{AR}})=0.$$

where

$$\pi_i^{\text{W}*}\left(\mathbf{P}_{\text{AR}}\right)=\begin{cases}\frac{t}{n^2}+\frac{2d_{\text{AR}}t}{3n}-\frac{7td_{\text{AR}}^2}{18}-F & \text{for } d_{\text{AR}}<\frac{3}{2n}\\ \frac{9t}{8n^2}-F & \text{for } d_{\text{AR}}\geq\frac{3}{2n}\end{cases} \tag{A.59}$$
$$\pi_i^{\text{L}*}\left(\mathbf{P}_{\text{AR}}\right)=\frac{t}{4n^2}-F\quad\text{for } 0\leq d_{\text{AR}}\leq 1.$$

We can rewrite DB's profits by substituting (A.59) in (A.58), obtaining

$$\max_{d_{\text{AR}}}\pi_{\text{DB}}=\begin{cases}\frac{n}{2}\left(\frac{3t}{4n^2}+\frac{2d_{\text{AR}}t}{3n}-\frac{7td_{\text{AR}}^2}{18}\right) & \text{for } d_{\text{AR}}<\frac{3}{2n}\\ \frac{n}{2}\left(\frac{7t}{8n^2}\right) & \text{for } d_{\text{AR}}\geq\frac{3}{2n}.\end{cases}$$

When the DB sets $d_{\text{AR}}<\frac{3}{2n}$, the FOC of π_{DB} gives $d_{\text{AR}}^*=\frac{6}{7n}$, resulting in profits equal to $\pi_{\text{DB}}^*=\frac{29t}{56n}$. Conversely, when the DB sets $d_{\text{AR}}\geq\frac{3}{2n}$, then his profits are constant with respect to d_{AR} and equal to $\pi_{\text{DB}}^*=\frac{7t}{16n}$. By comparing the two results, we find that the DB maximizes his profits by setting $d_{\text{AR}}^*=\frac{6}{7n}$.[20] The number of entering firms will be such that their profits after paying for entry and data are 0. We obtain the number of entering firms by solving the free-entry condition

$$\pi_i^{\text{L}*}\left(\mathbf{P_H}\right)=\frac{t}{4n^2}-F=0,$$

from which we find $n_{\text{AR}}^*=\frac{1}{2}\sqrt{\frac{t}{F}}$. When the DB sells data to all firms, his profits are equal to $\pi_{\text{DB}}^{\text{TIOLI}*}(\mathbf{P}_{\text{TIOLI}})=\frac{t}{4n}$. By direct comparison, we find that the DB prefers selling data to half of the entering firms.

Step 3. Under AU, the DB offers same-sized partitions in all the auctions. Without loss of generality, we focus our analysis on firm 0.

Firm 0's profits when winning and losing are the same as in Lemmas 2 and 3. Thus, DB's profits are

$$\max_{d_{\text{AU}}}\pi_{\text{DB}}^{\text{AU}}=\begin{cases}\frac{d_{\text{AU}}t(8-3d_{\text{AU}}n)}{12} & \text{for } d_{\text{AU}}<\frac{3}{2n}\\ \frac{7t}{16n} & \text{for } d_{\text{AU}}\geq\frac{3}{2n},\end{cases} \tag{A.60}$$

[20] Note that, under AR, firms' losing profits do not depend on d_H, and thus the number of entering firms does not affect the DB's strategy.

given that

$$\begin{cases} \frac{t}{n^2} - \frac{2d_{\text{AU}}t}{3n} + \frac{\text{td}_{\text{AU}}^2}{9} - F \geq 0 & \text{for } d_{\text{AU}} < \frac{3}{2n} \\ \frac{t}{4n^2} - F \geq 0 & \text{for } d_{\text{AU}} \geq \frac{3}{2n}. \end{cases} \tag{A.61}$$

When $d_{\text{AU}} < \frac{3}{2n}$, we obtain the number of entering firms by binding the first part of the piecewise function (A.61), obtaining

$$n_{\text{AU}}^* = \frac{9\sqrt{tF} - 3d_{\text{AU}}t}{9F - td_{\text{AU}}^2}. \tag{A.62}$$

When $d_{\text{AU}} \geq \frac{3}{2n}$, the number of entering firms is constant and given by binding the second part of the piecewise function (A.61), obtaining

$$n_{\text{AU}}^* = \frac{1}{2}\sqrt{\frac{t}{F}}. \tag{A.63}$$

By substituting (A.62) and (A.63) in (A.60), we obtain

$$\max_{d_{\text{AU}}} \pi_{\text{DB}}^{\text{AU}} = \begin{cases} \frac{d_{\text{AU}}t\left(72F + td_{\text{AU}}^2 - 27d_{\text{AU}}\sqrt{tF}\right)}{108F - 12td_{\text{AU}}^2} & \text{for } d_{\text{AU}} < \frac{3}{2n_{\text{AU}}^*} \\ \frac{7}{8}\sqrt{tF} & \text{for } d_{\text{AU}} \geq \frac{3}{2n_{\text{AU}}^*}. \end{cases} \tag{A.64}$$

Computing FOCs of (A.64) for $d_{\text{AU}} < \frac{3}{2n_{\text{AU}}^*}$ with respect to d_{AU}, we find that DB's profits are monotonically increasing in d_{AU}. As such, the DB sets $d_{\text{AU}}^* \geq \frac{3}{2n_{\text{AU}}^*}$ and obtains profits

$$\pi_{\text{DB}}^{\text{AU}*} = \frac{7}{8}\sqrt{tF}. \tag{A.65}$$

As profits in (A.65) are higher than when selling data to all firms, in equilibrium the DB opts for selling to every other firm.

Proof of Proposition 2. The proof proceeds in two steps. In Step 1, we compute consumer surplus in equilibrium under the three selling mechanisms. In Step 2, we instead compute total welfare.

Step 1. We start our analysis with TIOLI, where in equilibrium all consumers are identified. Without loss of generality, let us focus on the arch between firms i and $i+1$. The indifferent consumer in the middle of the arch is located in $\frac{2i+1}{2n^*}$. Firm i serves all its consumers in $[\frac{i}{n^*}, \frac{2i+1}{2n^*}]$ through its tailored price. Consumer surplus on this semi-arch is given by integrating consumer net utility between $\frac{i}{n^*}$ and $\frac{2i+1}{2n^*}$. The total consumer surplus on all $2n^*$ semi-arches of the market is:

$$CS^{TIOLI^*} = 2n^*\left(\int_{\frac{i}{n^*}}^{\frac{2i+1}{2n^*}} u - p_i^{\text{T}*}(x) - t\left(x - \frac{i}{n^*}\right)dx\right), \tag{A.66}$$

where

$$p_i^{\text{T}}(x) = -2tx + \frac{t}{n^*}(2i+1). \tag{A.67}$$

By replacing (A.67) in (A.66) we obtain

$$CS^{TIOLI^*} = u - \frac{3t}{4n^*} = u - \frac{3}{2}\sqrt{tF},$$

i.e. $CS^{TIOLI^*} < \widetilde{CS} = u - \frac{5}{4}\sqrt{tF}$.

As for AR, Suppose that firm i wins its auction in equilibrium: as such, firm i and $i+1$ basic prices are given by (A.28) and (A.29), with $d_{\text{AR}}^* = \frac{6}{7n_{\text{AR}}^*}$. Firm i's tailored price is $p_{i,i+1}^{\text{T}*}(x) = p_{i+1}^{\text{B}*}\left(\mathbf{P}_{\text{AR}}^*\right) - 2tx + \frac{t}{n_{\text{AR}}^*}(2i+1)$, and the indifferent consumer is located in $\widehat{x^*}_{i,i+1} = \frac{2i+1}{2n_{\text{AR}}^*} + \frac{d_{\text{AR}}^*}{6}$. Consumer surplus is thus equal to

$$\begin{aligned} CS \; = \; n_{\text{AR}}^* \Bigg(& \int_{\frac{i}{n_{\text{AR}}^*}}^{\frac{i}{n_{\text{AR}}^*}+\frac{d_{\text{AR}}^*}{2}} u - p_{i,i+1}^{\text{T}*}(x) - t\left(x - \frac{i}{n_{\text{AR}}^*}\right)dx + \int_{\frac{i}{n_{\text{AR}}^*}+\frac{d_{\text{AR}}^*}{2}}^{\widehat{x^*}_{i,i+1}} u - p_i^{\text{B}*}\left(\mathbf{P}_{\text{AR}}^*\right) \\ & - t\left(x - \frac{i}{n_{\text{AR}}^*}\right)dx + \int_{\widehat{x^*}_{i,i+1}}^{\frac{i+1}{n_{\text{AR}}^*}} u - p_{i+1}^{\text{B}*}\left(\mathbf{P}_{\text{AR}}^*\right) - t\left(\frac{i+1}{n_{\text{AR}}^*} - x\right)dx \Bigg). \end{aligned} \tag{A.68}$$

By substituting the prices and the indifferent consumer's location in (A.68), we obtain

$$CS = u - \frac{5t}{4n_{\text{AR}}^*} + \frac{\text{nt}{d_{\text{AR}}^*}^2}{9}, \tag{A.69}$$

which we can rewrite as

$$CS = u - \frac{229}{98}\sqrt{tF}. \tag{A.70}$$

Consumer surplus where data are absent is equal to $\widetilde{CS} = u - \frac{5t}{4\tilde{n}}$, with $\tilde{n} = \sqrt{\frac{t}{F}}$ (Salop, 1979). This implies $\widetilde{CS} = u - \frac{5}{4}\sqrt{tF}$, which is always higher than CS in (A.70).

Finally, with regard, to AU, we can use (A.68) to compute consumer surplus, as the difference between AR and AU with regard to consumer surplus only lies in the quantity of data sold. By replacing d^*_{AU} in (A.68), we obtain

$$CS^{\mathrm{AU}} = u - \frac{5t}{4n^*_{\mathrm{AU}}} + \frac{n^*_{\mathrm{AU}} t d^{*2}_{\mathrm{AU}}}{9}. \tag{A.71}$$

where $n^*_{\mathrm{AU}} = \frac{1}{2}\sqrt{\frac{t}{F}}$ and $d^*_{\mathrm{AU}} = \frac{3}{2n^*_{\mathrm{AU}}}$, as it is the limit case after which data exhaust their marginal effect. We can rewrite (A.71) as

$$CS^{\mathrm{AU}} = u - \frac{t}{n^*_{\mathrm{AU}}} = u - 2\sqrt{tF},$$

which is lower than $\widetilde{CS}$.

Step 2.

To compute Total Welfare, we need to add consumer surplus, firms' profits and the DB's profits. Under TIOLI and AU, all firms remaining profits after paying for data are equal to $\pi_i^{L^*} = \frac{t}{4n^{*2}}$. As all firms' profits are equal, they are all dissipated by the entry cost, and thus firms' equilibrium total profits are equal to zero. First, suppose that $\alpha = 0$. Then, total welfare is equal to consumer surplus, and we find that total welfare is always lower than in the benchmark case. Instead, suppose that $\alpha = 1$. Then, by adding consumer surplus and the DB's profits, we find

$$TW^{TIOLI^*} = u - \sqrt{tF}$$

$$TW^{AU^*} = u - \frac{9}{8}\sqrt{tF},$$

which are both higher than $\widetilde{TW}$. Then, by continuity, we conclude that under TIOLI and AU there exists an α^{SM^*} such that, if $\alpha \geq \alpha^{SM^*}$, then $TW^{SM^*} \geq \widetilde{TW}$. By simple calculations, $\alpha^{TIOLI^*} = \frac{1}{2}$ and $\alpha^{AU^*} = \frac{15}{16}$.

Instead, under AR, firms' losing profits are higher than firms' winning profits after paying for data. We can write total firms' equilibrium profits as[21]

$$\pi^*_{firms} = \frac{n^*}{2}\left(\frac{t}{4n^{*2}} - F\right) + \frac{n^*}{2}\left(\frac{t}{n^{*2}} - \frac{2d^*_{AR}t}{3n^*} + \frac{td^{*2}_{AR}}{9} - F\right). \tag{A.72}$$

By replacing $d^*_{AR} = \frac{6}{7n^*}$ and $n^* = \frac{1}{2}\sqrt{\frac{t}{F}}$, we obtain

$$\pi^*_{firms} = \frac{43}{196}\sqrt{tF}. \tag{A.73}$$

We can thus write Total Welfare under AR as

$$TW^*_{AR} = u - \frac{229}{98}\sqrt{tF} + \alpha\left(\frac{43}{196}\sqrt{tF} + \frac{29}{28}\sqrt{tF}\right). \tag{A.74}$$

By simple calculations, we find that if $\alpha \geq \alpha^*_{AR} = \frac{71}{82}$, then $TW^*_{AR} \geq \widetilde{TW}$.

Proof of Proposition 3. To find how the market power level θ affects the DB's equilibrium strategy, we rewrite DB's profits as

$$\max_{d_A} \pi_{\mathrm{DB}} = \begin{cases} n\theta\left(\frac{t}{n^2} - \frac{td_A^2}{2} - \left(\frac{t}{n} - td_A + 2td_A a_0\right)\left(2d_A\left(a_1 - a_0\right) + \frac{1}{n}\right)\right) & \text{for } d_A < d_1 \\ n\theta\left(\frac{t}{n^2} - \frac{td_A^2}{2} - \frac{t(nd_A-4)^2}{25n^2}\right) & \text{for } d_1 \leq d_A < \frac{1}{n} \\ n\theta\left(\frac{t}{2n^2} - \frac{t(nd_A-4)^2}{25n^2}\right) & \text{for } \frac{1}{n} \leq d_A < \frac{3}{2n} \\ n\theta\left(\frac{t}{4n^2}\right) & \text{for } d_A \geq \frac{3}{2n}. \end{cases} \tag{A.75}$$

[21] Note that, while firms' losing profits are greater than zero, an increase of n by one unit would always make them negative. Thus, the number of entering losing firms is equal to the number of entering winning firms.

Firms' profits after paying for data are equal to $\pi_i^W - \theta(\pi_i^W - \pi_i^L)$, which can be written as

$$\pi_i^W - \theta(\pi_i^W - \pi_i^L) = \begin{cases} \frac{t}{n^2} - \frac{td_A^2}{2} - \theta\left(\frac{t}{n^2} - \frac{td_A^2}{2} - \left(\frac{t}{n} - td_A + 2td_A a_0\right)\left(2d_A\left(a_1 - a_0\right) + \frac{1}{n}\right)\right) & \text{for } d_A < d_1 \\ \frac{t}{n^2} - \frac{td_A^2}{2} - \theta\left(\frac{t}{n^2} - \frac{td_A^2}{2} - \frac{t(nd_A-4)^2}{25n^2}\right) & \text{for } d_1 \le d_A < \frac{1}{n} \\ \frac{t}{2n^2} - \theta\left(\frac{t}{2n^2} - \frac{t(nd_A-4)^2}{25n^2}\right) & \text{for } \frac{1}{n} \le d_A < \frac{3}{2n} \\ \frac{t}{2n^2} - \theta\left(\frac{t}{4n^2}\right) & \text{for } d_A \ge \frac{3}{2n}. \end{cases} \tag{A.76}$$

To find the DB's equilibrium strategy, we find the number of entering firms as a function of d_A by binding (A.76) to zero, and then maximize the DB's profits with respect to d_A. As for the previous proofs, we cannot find a closed-form solution for the number of entering firms due to the exponential nature of a_0 and a_1. Thus, we follow the same approximation approach of previous proofs, obtaining $a_0 \approx \frac{1}{2\sqrt{3}}$ and $a_1 \approx \frac{2-\sqrt{3}}{2\sqrt{3}}$. Applying these approximations to (A.76) and solving for n we obtain:

$$n^*(d_A) = \begin{cases} \frac{\sqrt{\left(4\sqrt{3}d_A\theta t - 12d_A\theta t\right)^2 - 24t\left(-4\sqrt{3}d_A^2\theta t + 11d_A^2\theta t - 3d_A^2 t - 6F\right)} - 12d_A\theta t + 4\sqrt{3}d_A\theta t}{2\left(-11d_A^2\theta t + 4\sqrt{3}d_A^2\theta t + 3d_A^2 t + 6F\right)} & \text{for } d_A < d_1^* \\ \frac{8d_A\theta t - 5\sqrt{2}\sqrt{11d_A^2\theta^2 t^2 - 36d_A^2\theta t^2 + 25d_A^2 t^2 - 18F\theta t + 50Ft}}{27d_A^2\theta t - 25d_A^2 t - 50F} & \text{for } d_1 \le d_A < \frac{1}{n^*(d_A)} \\ \frac{8d_A\theta t - 5\sqrt{2}\sqrt{d_A^2\theta^2 t^2 - d_A^2\theta t^2 + 7F\theta t + 25Ft}}{2\left(d_A^2\theta t - 25F\right)} & \text{for } \frac{1}{n^*(d_A)} \le d_A < \frac{3}{2n^*(d_A)} \\ \frac{\sqrt{2t - \theta t}}{2\sqrt{F}} & \text{for } d_A \ge \frac{3}{2n^* d_A}. \end{cases} \tag{A.77}$$

By replacing (A.77) in (A.75) and maximizing with respect to d_A, we find that the DB sets $d_A^* \ge \frac{3}{2n^*}$ irrespective of θ. From (A.75) it is also immediate that DB's equilibrium profits are increasing in θ.

With regard to consumer surplus, we can use the same formula obtained in the proof of Proposition 2, i.e., $CS^* = u - \frac{3t}{4n^*}$, with $n^* = \frac{\sqrt{2t-\theta t}}{2\sqrt{F}}$. By comparing it with $\widetilde{CS} = u - \frac{5}{4}\sqrt{tF}$, we find that $CS^* \ge \widetilde{CS}$ iff $\theta \le \frac{14}{25}$.

With regard to total welfare, we compute it as $TW^* = \alpha\pi_{DB}^* + CS^*$. By deriving it with respect to θ, we obtain

$$\frac{\delta TW^*}{\delta\theta} = \frac{\sqrt{F}t^2(\alpha\theta - 3)}{4((\theta-2)(-t))^{3/2}} + \frac{\alpha\sqrt{F}t}{2\sqrt{(\theta-2)(-t)}}, \tag{A.78}$$

which is lower than zero iff $\alpha < \frac{3}{4}$.

Proof of Proposition 4. The proof proceeds in two steps. First, we show that if partitions do not overlap, there exists a unique equilibrium. Second, we show that if partitions do overlap, there exists another unique equilibrium.

Step 1. Suppose all firms buy data and choose to use a partition of size $d_{A,i} < \frac{1}{n}$. Then, all firms' profits are as in (A.19). However, we now assume that each firm can individually choose their partition's size $d_{A,i}$. Following the same reasoning of Lemma 1, we obtain that equilibrium prices are equal to

$$p_i^{B*} = \frac{t}{n} - 2t\sum_{j=0}^{n-1} d_{A,i}a_j. \tag{A.79}$$

To proceed with the analysis, it is useful to isolate the terms $d_{A,i}$ from the equilibrium basic prices of firms $i, i+1$ and $i-1$. We obtain:

$$p_i^{B*} = \frac{t}{n} - 2td_{A,i}a_0 - 2t\sum_{j=1}^{n-1} d_{A,i}a_j = \frac{t}{n} - 2td_{A,i}a_0 - 2th, \tag{A.80}$$

$$p_{i+1}^{B*} = \frac{t}{n} - 2td_{A,i}a_1 - 2t\sum_{j=0, j\neq 1}^{n-1} d_{A,i}a_j = \frac{t}{n} - 2td_{A,i}a_1 - 2tk, \tag{A.81}$$

$$p_{i-1}^{B*} = \frac{t}{n} - 2td_{A,i}a_1 - 2t\sum_{j=0,}^{n-2} d_{A,i}a_j = \frac{t}{n} - 2td_{A,i}a_1 - 2tl, \tag{A.82}$$

where h, k and l represent the sum of all the partitions which are not those of firm i. By replacing (A.80), (A.81) and (A.82) in (A.19) and maximizing for $d_{A,i}$, we obtain

$$d_{A,i}^* = \frac{1 - 2a_1 + 2hn - 8a_0hn + 4a_1hn - kn + 2a_0kn - ln + 2a_0ln}{n(1 - 4a_0 + 8a_0^2 + 4a_1 - 8a_0a_1)}. \tag{A.83}$$

As the equilibrium partitions' sizes used by firms are the same due to symmetry, we can replace $h + d_{A,i}a_0 = \frac{d_{A,i}}{2}$, $k + d_{A,i}a_1 = l + d_{A,i}a_1 = \frac{d_{A,i}}{2}$, and obtain

$$d^*_{A,i} = d^*_{low} = \frac{1-2a_1}{n}, \tag{A.84}$$

which is lower than $\frac{1}{n}$.

Step 2.

Suppose that all firms purchase a partition $d_A \geq \frac{3}{2n}$ and choose to use all of them so that partitions overlap. We focus on a generic firm i's incentive to deviate from using all the data it purchased. As firms $i+1$ and $i-1$ can identify all the consumers they can potentially serve, they offer all consumers tailored prices, as they allow them to extract all available surplus from consumers and thus always outperform basic prices. Conversely, firm i chooses the partition size $d_{A,i} \in [0, d_A]$ it wants to use. Suppose firm i offers a basic price to a consumer who is identified by firm $i+1$. Then, in order for that consumer to buy from firm i, we must have that $U(x,i) \geq U(x,i+1)$, with $p^B_{i+1} = 0$. Indeed, the only instance in which a basic price offer beats a tailored price offer is if the tailored price becomes negative, as the firm would prefer not to serve that consumer. We thus find the indifferent consumers' location by replacing $p^B_{i+1} = p^B_{i-1} = 0$ in (A.20). By maximizing firm i's profits with respect to p^B_i, we find $p^B_i = \frac{t}{2n^2}(1-nd)$. By replacing the equilibrium basic prices in firm i's profits (6), and maximizing with respect to d, we find that firm i's profits are maximized for any $d_{A,i} = d^*_{high} \geq \frac{1}{n}$. This implies that firm i has no incentive to deviate from using all data, as its profits are the same for any $d_{A,i} = d^*_{high} \geq \frac{1}{n}$.

Proof of Proposition 5. The proof proceeds in two steps. First, we show that, if a firm chooses not to purchase data, its direct rivals choose to identify fewer consumers on the arch they share with the uninformed firm. Second, we compute consumer surplus under the new equilibrium.

Step 1.

Suppose instead that firm i does not purchase data, whereas all other firms obtain $d_A \geq \frac{3}{2n}$. First, let us focus on firm i's direct rivals ($i+1$ and $i-1$) and how they compete with their informed rivals ($i+2$ and $i-2$). As both firms $i+1$ and $i+2$ can target all consumers they serve, they only offer tailored prices on the arch they share, as described in Step 1. The same reasoning also applies to firms $i-1$ and $i-2$. Conversely, on the arch they share with firm i, they can choose the amount of data d_{i+1} and d_{i-1} they want to use.

As firm i's direct rivals only offer their basic prices on the arch they share with firm i, we only focus on the profits they make on said arch. We thus obtain:

$$\pi^W_{i-1} = \frac{p^B_{i-1} + d_{i-1}t}{2n} - \frac{2p^B_{i-1}(p^B_{i-1} - p^B_i) + 2d_{i-1}t(p^B_{i-1} - p^B_i) + d^2_{i-1}t^2}{4t} - F \tag{A.85}$$

$$\pi^L_i = \frac{p^B_i}{n} + \frac{p^B_i(p^B_{i+1} - 2p^B_i + p^B_{i-1})}{2t} - F \tag{A.86}$$

$$\pi^W_{i+1} = \frac{p^B_{i+1} + d_{i+1}t}{2n} - \frac{2p^B_{i+1}(p^B_{i+1} - p^B_i) + 2d_{i+1}t(p^B_{i+1} - p^B_i) + d^2_{i+1}t^2}{4t} - F \tag{A.87}$$

Note that, to remain consistent with the notation of the baseline model, d_{i+1} and d_{i-1} represent the amount of data used on both arches. Thus, only a share of $\frac{d_{i+1}}{2}$ and $\frac{d_{i-1}}{2}$ are used on the arch shared with firm i.

Equilibrium basic prices are

$$p^{B*}_{i-1} = \frac{12t - 7ntd_{i-1} - ntd_{i+1}}{12n}, \quad p^{B*}_i = \frac{6t - nt(d_{i-1} + d_{i+1})}{6n}, \quad p^{B*}_{i+1} = \frac{12t - 7ntd_{i+1} - ntd_{i-1}}{12n}. \tag{A.88}$$

Finally, by maximizing firms $i+1$ and $i-1$'s profits with respect to d_{i+1} and d_{i-1}, we obtain $d^*_{i+1} = d^*_{i-1} = \frac{15}{13n}$. Thus, firm i's direct rivals only identify a segment of length $\frac{15}{26n}$ on the arch they share with firm i. As this amount is lower than $frac34n$, we conclude that firm i's direct rivals would use less data than in the baseline model. The resulting firm i's profits are

$$\pi^{L*}_i = \frac{64t}{169n^2} - F. \tag{A.89}$$

As these profits are higher than in the baseline model, it is immediate to conclude that equilibrium partition prices are lower.

Step 2.

Firms' strategic use of data changes their profits if they choose not to buy data and, in turn, the number of entering firms. We obtain the equilibrium number of entering firms by binding (A.89) to zero, which results in

$$n^* = \frac{8}{13}\sqrt{\frac{t}{F}}, \tag{A.90}$$

which is higher than the number of entering firms in the baseline model. With regard to consumer surplus, in equilibrium, all consumers are served through tailored prices, and we can thus use the same expression as in the proof of Proposition 2, i.e., $CS^* = u - \frac{3t}{4n^*}$, with $n^* = \frac{8}{13}\sqrt{\frac{t}{F}}$. We thus obtain $CS^* = u - \frac{39}{32}\sqrt{tF}$, which is always higher than $\widetilde{CS}$.

Proof of Proposition 6. We analyze possible combinations of interventions by a policymaker. First, suppose that the policymaker only intervenes by imposing a selling mechanism on the DB. From the results of Proposition 2, the policymaker would impose the TIOLI mechanism, as it entails the highest level of consumer surplus and total welfare (regardless of α). Second, suppose that the policymaker can also impose a cap on the maximum partition size that can be sold by the DB. To understand the effects of this cap on consumer surplus and total welfare, it is useful to find DB's profits and the number of entering firms as a function of d_A. DB's profits as a function of d_A are as in (A.43). With regard to the number of entering firms, we find it by binding to zero firms' profits after paying for entry and data, which are equal to:

$$\pi_i^{L^*} = \begin{cases} \left(\frac{t}{n} - td_A + 2td_A a_0\right)\left(2d_A\left(a_1 - a_0\right) + \frac{1}{n}\right) - F & \text{for } d_A < d_1 \\ \frac{t(nd_A-4)^2}{25n^2} - F & \text{for } d_1 \le d_A < \frac{3}{2n} \\ \frac{t}{4n^2} - F & \text{for } d_A \ge \frac{3}{2n}. \end{cases} \tag{A.91}$$

By solving for n (and applying the same approximation method used in the proof of Proposition 1 for the first part of $\pi_i^{L^*}$ due to lack of a closed form solution), we find:

$$n^{TIOLI^*} = \begin{cases} \frac{-3d_At+\sqrt{3}d_At+3\sqrt{F}\sqrt{t}}{-4d_A^2t+2\sqrt{3}d_A^2t+3F} & \text{for } d_A < d_1(n^{TIOLI^*}) \\ \frac{4\left(d_At-5\sqrt{F}\sqrt{t}\right)}{d_A^2t-25F} & \text{for } d_1(n^{TIOLI^*}) \le d_A < \frac{3}{2n^{TIOLI^*}} \\ \frac{1}{2}\sqrt{\frac{t}{F}} & \text{for } d_A \ge \frac{3}{2n^{TIOLI^*}}. \end{cases} \tag{A.92}$$

Consumer surplus as a function of d_A can be written as

$$CS^{TIOLI} = \begin{cases} u - \frac{5t}{4n^{TIOLI^*}} + \frac{n^{TIOLI^*}td_A^2}{2} & \text{for } d_A < \frac{1}{n^{TIOLI^*}} \\ u - \frac{3}{2}\sqrt{tF} & \text{for } d_A \ge \frac{1}{n^{TIOLI^*}}. \end{cases} \tag{A.93}$$

By combining (A.92) and (A.93), we find that CS^{TIOLI} is continuous in d_A, is decreasing in d_A for $d_A < d_1(n^{TIOLI^*})$, increasing in d_A for $d_1(n^{TIOLI^*}) \le d_A \le \frac{1}{n^{TIOLI^*}}$, decreasing in d_A for $\frac{1}{n^{TIOLI^*}} < d_A < \frac{3}{2n^{TIOLI^*}}$ and constant in d_A for $d_A \ge \frac{3}{2n^{TIOLI^*}}$. Thus, it is straightforward that consumer surplus has two local maxima in $d_A = 0$, resulting in $n^{TIOLI^*} = \sqrt{\frac{t}{F}}$ and $d_A = \frac{1}{n^{TIOLI^*}}$, resulting in $n^{TIOLI^*} = \frac{3}{5}\sqrt{\frac{t}{F}}$. The case of $d_A = 0$ is the standard Salop model, where $\widetilde{CS} = u - \frac{5}{4}\sqrt{tF}$. We also find that $CS(d_A = \frac{1}{n^{TIOLI^*}}) = \widetilde{CS}$. However, from the proof of Proposition 1, we know that DB's profits have a local maximum for $d_A^* < d_1 < \frac{1}{n}$. Thus, if a policymaker would impose a maximum partition size, the DB would instead set $d_A = d_A^* < \frac{1}{n}$. Thus, to achieve the desired outcome, the policymaker should impose the DB must sell all available data, without selling the same data points to more than one firm. By doing so, the policymaker can obtain the same consumer surplus as in the standard Salop model, and increase total welfare as DB's profits are positive (see (A.43)). With regards to the effects of θ, we repeat the same analysis on CS^{TIOLI} with the number of entering firms in equilibrium as in (A.77), and again find that it is maximized for $d_A^* = \frac{1}{n}$ and decreasing in θ. Thus, any $\theta < 1$ makes consumer surplus higher than in the benchmark case.

From a total welfare perspective, recall that under TIOLI all firms obtain symmetric partitions. Thus, transportation costs are minimized and not influenced by d_A, as all consumers buy from the closest firm. Then, total welfare is only influenced by the fixed cost of entry and the weight of industry profits α. Suppose $\alpha = 0$. Then, $TW^{TIOLI^*} = CS^{TIOLI^*}$, and to maximize total welfare the policymaker would set $d_A^P = \frac{1}{n^{TIOLI^*}}$. Instead, suppose $\alpha = 1$. Then, the policymaker minimizes the number of entering firms by setting $d_A^P \ge \frac{3}{2n^{TIOLI^*}}$. Thus, by continuity of the total welfare function with respect to α, we conclude that there exists a threshold $\bar{\alpha}$ such that, if $\alpha < \bar{\alpha}$, then setting $d_A^P = \frac{1}{n^{TIOLI^*}}$ also maximizes total welfare. Instead, if $\alpha \ge \bar{\alpha}$, a policymaker that aims to maximize total welfare sets $d_A^P \ge \frac{3}{2n^{TIOLI^*}}$.

References

Abrardi, L., Cambini, C., Congiu, R., & Pino, F. (2024). User data and endogenous entry in online markets. *The Journal of Industrial Economics, forthcoming*.

ACCC (2023). *Digital platform services inquiry–March 2024 report on data brokers: issues paper*: *Tech. rep.*, Australian Competition and Consumer Commission.

Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature*, *54*(2), 442–492.

Ali, S. N., Lewis, G., & Vasserman, S. (2022). Voluntary disclosure and personalized pricing. *Review of Economic Studies*, *90*(2), 538–571.

Baron, D. P., & Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica*, 911–930.

Bergemann, D., & Bonatti, A. (2011). Targeting in advertising markets: implications for offline versus online media. *Rand Journal of Economics*, *42*(3), 417–443.

Bergemann, D., & Bonatti, A. (2019). Markets for information: An introduction. *Annual Review of Economics*, *11*, 85–107.

Bester, H., & Petrakis, E. (1996). Coupons and oligopolistic price discrimination. *International Journal of Industrial Organization*, *14*(2), 227–242.

Bounie, D., Dubus, A., & Waelbroeck, P. (2021). Selling strategic information in digital competitive markets. *Rand Journal of Economics*, *52*(2), 283–313.

Bounie, D., Dubus, A., & Waelbroeck, P. (2022). *Market for Information and Selling Mechanisms*: *Economics Working Paper Series 22/367*, CER-ETH – Center of Economic Research at ETH Zurich.

Braulin, F. C., & Valletti, T. (2016). Selling customer information to competing firms. *Economics Letters*, *149*, 10–14.

Chen, Z., Choe, C., & Matsushima, N. (2020). Competitive Personalized Pricing. *Management Science*, *66*(9), 4003–4023.

Choe, C., Cong, J., & Wang, C. (2024). Softening competition through unilateral sharing of customer data. *Management Science*, *70*(1), 526–543.
Christl, W. (2017). *Corporate Surveillance in Everyday Life*: *Tech. rep.*, Australian Competition and Consumer Commission, Cracked Labs.
Cramer-Flood, E. (2023). *Worldwide Ecommerce Forecast 2023*: *Tech. rep.*, Insider Intelligence. URL: https://www.insiderintelligence.com/content/worldwide-ecommerce-forecast-2023.
FTC (2014). *Data Brokers: A Call for Transparency and Accountability*: *Tech. rep.*, Washington, DC: Federal Trade Commission.
Fudenberg, D., & Villas-Boas, J. M. (2006). Behavior-based price discrimination and customer recognition. *Vol. 1*, In *Handbook on economics and information systems* (pp. 377–436). Citeseer.
Goldfarb, A., & Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, *57*(1), 3–43.
Hagiu, A., & Wright, J. (2020). *Data-Enabled Learning, Network Effects and Competitive Advantage*: *Working Paper*.
Jehiel, P., & Moldovanu, B. (2000). Auctions with Downstream Interaction among Buyers. *Rand Journal of Economics*, *31*(4), 768–791.
Krämer, J., Colangelo, G., Richter, H., & Schnurr, D. (2023). *Data Act: Towards A Balanced EU Data Regulation*: *Tech. rep.*, Centre on Regulation in Europe, URL: https://cerre.eu/wp-content/uploads/2023/03/230327_Data-Act-Book.pdf.
Liu, Q., & Serfes, K. (2004). Quality of information and oligopolistic price discrimination. *Journal of Economics & Management Strategy*, *13*(4), 671–702.
Liu, Q., & Serfes, K. (2005). Imperfect price discrimination, market structure, and efficiency. *Canadian Journal of Economics/Revue Canadienne d'économique*, *38*(4), 1191–1203.
Montes, R., Sand-Zantman, W., & Valletti, T. (2019). The value of personal information in online markets with endogenous privacy. *Management Science*, *65*(3), 955–1453.
Pino, F. (2022). The microeconomics of data–a survey. *Journal of Industrial and Business Economics*, *49*(3), 635–665.
Rhodes, A., & Zhou, J. (2022). Personalized pricing and competition. Available at SSRN 4103763.
Ruschemeier, H. (2022). Data brokers and European digital legislation. *EDPL European Data Protection Law Review*, *9*(2023), 27–38.
Salop, S. C. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics*, 141–156.
Searle, S. R. (1979). On inverting circulant matrices. *Linear Algebra and its Applications*, *25*, 77–89.
Shaffer, G., & Zhang, Z. J. (1995). Competitive coupon targeting. *Marketing Science*, *14*(4), 395–416.
Shy, O., & Stenbacka, R. (2016). Customer privacy and competition. *Journal of Economics & Management Strategy*, *25*(3), 539–562.
Taylor, C. R. (2003). Supplier Surfing: Competition and consumer behavior in subscription markets. *Rand Journal of Economics*, *34*(2), 223–246.
Taylor, & Wagman, L. (2014). Consumer privacy in oligopolistic markets: Winners, losers, and welfare. *International Journal of Industrial Organization*, *34*, 80–84.
Thisse, J.-F., & Vives, X. (1988). On the strategic choice of spatial price policy. *The American Economic Review*, *78*(1), 122–137.
Villas-Boas, J. M. (2004). Price cycles in markets with customer recognition. *Rand Journal of Economics*, *35*(3), 486–501.