Softening Competition through Unilateral Sharing of Customer Data*

Chongwoo Choe,[†] Jiajia Cong,[‡] Chengsi Wang[§]

October 20, 2022

Abstract

We study how a data-rich firm can benefit by unilaterally sharing its customer data with a data-poor competitor when the data can be used for price discrimination. By sharing data on the segment of market that is more loyal to the competitor while keeping the data on the competitor's most loyal segment to itself, the firm can induce the competitor to raise its price for consumers it does not have data on. Such data sharing is an example of a fat-cat strategy as it softens price competition that follows data sharing. Although consumer surplus decreases as a result of data sharing, total surplus can increase when the sharing firm concedes its market share to the competitor, which improves the quality of consumer-firm matching.

Keywords: customer data sharing, price discrimination
JEL Classification Number: D43, L13

1 Introduction

In 2017, *The Economist* famously declared that data is the new oil in the digital economy.[1] Consumer-generated data analyzed with powerful machine-learning

*We thank the Associate Editor and three anonymous reviewers for many valuable comments and suggestions. We are also thankful to Zhijun Chen, Jay Pil Choi, Noriaki Matsushima, Vai-Lam Mui, Andrew Rhodes, Nicolas Schutz, Julian Wright, and Xiaojian Zhao for helpful comments. Choe and Wang gratefully acknowledge financial support from the Australian Research Council (grant number DP210102015). Cong would like to acknowledge the financial support from NSFC (Projects 72003035, 72192845) and Shanghai Pujiang Program (2019PJC007). The usual disclaimer applies.

†Department of Economics, Monash University, chongwoo.choe@monash.edu.

‡School of Management, Fudan University, jjcong@fudan.edu.cn.

§The corresponding author. Department of Economics, Monash University, Clayton, Victoria 3800, Australia. chengsiwang@gmail.com.

[1] "The world's most valuable resource is no longer oil, but data", *The Economist*, May 16, 2017.

tools can facilitate data-enabled learning, which can lead to new or improved products, more target-oriented business models, effective defense against competition, etc. (Hagiu and Wright, 2020). Consumer data can also allow firms to sharpen their pricing tools and extract more surplus from consumers through price discrimination. For example, finer-grained analysis utilizing big data has made personalized pricing closer to becoming a reality in some industries. The flip side is that competition can become more intense when it is based on finer levels of customer data that can be used for price discrimination (Thisse and Vives, 1988; Fudenberg and Tirole, 2000; Choe et al., 2018).

Can a firm benefit by *unilaterally* sharing its customer data with a competitor when the data can be used for price discrimination? As alluded to in the previous paragraph, it would seem that the sharing firm cannot benefit because data sharing will improve the competitor's pricing capabilities and intensify competition by facilitating poaching by the competitor.[2] The purpose of this paper is to provide a model where an informed firm can choose to share customer data with an uninformed competitor and, given the optimally chosen amount of shared data, both firms are better off than without data sharing. The key driver of this result is that the informed firm strategically selects a subset of data to share with a competitor, which induces the competitor to raise its price for consumers it does not have data on. Therefore, the unilateral data sharing considered in this paper is a soft commitment when the subsequent competition is in strategic complements, hence is an example of a fat-cat strategy (Fudenberg and Tirole, 1984).[3] We sketch below our main argument.

Consider a Hotelling model with two firms located at each end. Firm 1 has data on all consumers' locations and can set a personalized price for each consumer. Firm 2 does not have any data to start with and chooses only a uniform price. Given the maximal differentiation, firm 1's informational advantage does not extend to the entire market, implying that there are some consumers served by firm 2 even without data sharing. We call this firm 2's customer base. Since prices are strategic complements,

[2]See, for example, Kim and Choi (2010) or Chen et al. (2022). A number of studies reviewed later provide various models where mutual, rather than unilateral, data sharing can increase industry profits.

[3]In Fudenberg and Tirole (1984), a fat-cat strategy refers to overinvestment by an incumbent that accommodates entry by committing to play less aggressively post-entry. Subsequently, the term has come to be used more generally in two-stage games to refer to any strategic commitment a firm makes in the first stage that softens competition in the second stage when the second-stage competition is in strategic complements.

firm 1 can benefit if it can induce firm 2 to increase its uniform price.

Suppose now firm 1 shares data on some consumers in firm 2's customer base except those who are very close to firm 2's location. Then firm 2 makes pricing decisions separately for consumers it has data on and the rest: for the former whom we call targeted consumers, firm 2 chooses personalized prices and, for the latter, it chooses a uniform price. In the baseline model, we assume firm 2 can engage in search discrimination, which prevents its targeted consumers from choosing its uniform price.[4] Then firm 1 can choose data sharing in such a way that leads firm 2 to set its uniform price to serve only consumers who are very close to firm 2's location.[5] Consequently, firm 2 will raise its uniform price above the level it would choose in the absence of data sharing. This benefits firm 1 by allowing it to increase its personalized prices. But firm 1 does not concede any additional consumers to firm 2 because it shares data only for some consumers in firm 2's customer base.[6] Put together, firm 1 can benefit from data sharing through higher personalized prices at no cost of reduced market share. A clear implication from our analysis can be summarized as this: share data on consumers who are more loyal to the competitor than to yourself; but keep the data on the competitor's most loyal consumers to yourself.

The welfare effects of data sharing depend on market conditions and the nature of competition that follows data sharing. We can decompose the welfare effects into two parts. The first one is how data sharing enables firms to better extract consumer surplus, which we call the surplus-extraction effect. Clearly, the surplus-extraction effect impacts consumer surplus negatively. The second one is how data sharing affects the matching between consumers and firms, called the quality-of-matching effect. Data sharing may have a positive or negative quality-of-matching effect depending on how market shares change after data sharing. Given our assumption of full market coverage, the surplus extraction effect is relevant only to consumer surplus, while the

[4]In Section 4.1, we consider the case where firm 2 cannot engage in search discrimination. Firm 1 can still benefit from data sharing although the mechanism is now different. Firm 2's inability to search discriminate puts a cap on its personalized prices it can charge to targeted consumers. This leads firm 2 to increase its (off-the-path) uniform price, which benefits firm 1.

[5]Of course, the segment of firm 2's customer base with shared data needs to be chosen suitably so that firm 2 does not have incentives to set its uniform price to serve some consumers outside its customer base. In Section 3.2, we spell out the precise condition for this.

[6]This is an illustrative example of how firm 1 can benefit from sharing data with firm 2 without losing its market share. Firm 1's optimal data sharing may involve conceding some market share to firm 2 if the benefit of doing so more than offsets the reduced market share.

quality-of-matching effect is relevant to both consumer surplus and total surplus. If firm 1 increases its market share further as a result of data sharing, then both consumer surplus and total surplus decrease. In this case, not only does the surplus-extraction effect hurt consumers, but also the quality-of-matching effect is negative. However, if data sharing results in firm 1 giving away some of its market share to firm 2, then total surplus increases thanks to the positive quality-of-matching effect.

We study several variations of the baseline model and examine how our main insight extends to other economic environments. In Section 4.1, firm 2 is assumed to be unable to use search discrimination, i.e., it cannot prevent its targeted consumers from choosing its uniform price. We show that firm 1 continues to benefit from data sharing, but even more than when search discrimination is possible. In Section 4.2, we analyze the case where firm 2's data anaytics limits use of the shared data to third-degree price discrimination rather than personalized pricing. We show that our main insight continues to hold, but the dampened surplus-extraction effect in this case implies that consumers are better off than when the shared data is used for personalized pricing. In Section 4.3, we discuss the extent to which customer data can be used to deter firm 2's entry to the market. As data sharing always occurs once entry has been accommodated, firm 2's entry becomes more likely when data sharing is possible. Thus the possibility of data sharing mitigates the extent to which data plays a role as an entry barrier. In Section 4.4, we consider the case where firm 1 can be considered a dual-mode platform that charges a fixed sales fee to firm 2. We show that the only effect the sales fee has is to soften competition and raise all prices by the amount of the sales fee, and our results on the optimal data sharing remain robust. Several additional extensions of the baseline model are also discussed in the remainder of Section 4.

Contributions to the literature

Two strands of literature are directly relevant to our work. The first one relates to information sharing among firms and the second one is on competition based on personalized pricing.

Information sharing among firms exists in many industries such as airline, banking and finance, tourism and hospitality, etc. See, for example, Feasey and de Streel (2020) for a comprehensive discussion on information sharing in practice and the related regulatory issues. The earlier academic literature has been focused on sharing information on cost or demand conditions that can facilitate output or pricing decisions (e.g., Gal-Or, 1985; Armantier and Richard, 2003). More recent studies identify various conditions under which mutual information sharing, or information exchange between firms, can increase industry profits. These conditions pertain to firms' asymmetric and imperfect abilities to target customers (Chen et al. 2001), consumer switching costs (Shy and Stenbacka, 2013), or two-dimensional customer information (Jentzsch et al., 2013).[7] In contrast to these studies, our focus is on unilateral sharing of customer information that can be used for price discrimination.

Chen et al. (2001) and Liu and Serfes (2006) provide models where unilateral information sharing subject to side payments can increase industry profits. For example, the main mechanism in Liu and Serfes (2006) is that the firm with a small loyal customer base but a large amount of customer data can share its data with a competitor, thereby allowing the competitor to better extract surplus from the customers whose data is shared. Although this reduces the sharing firm's profit, the reduced profit can be compensated by the side payment. Consequently, unilateral information sharing does not arise in the absence of side payment. Somewhat related studies consider the case where a monopoly data broker sells data to competing firms to maximize its profit (Montes et al., 2019; Bounie et al., 2021). A general insight is that data is sold to only one firm in a way to soften competition, hence data sharing can never take place. In contrast, we show that unilateral information sharing is individually rational even in the absence of side payments. In addition, these studies typically assume consumer preferences are uniformly distributed whereas we consider general distributions.

There is a large and growing body of literature on personalized pricing. Consumer data is shown to intensify competition when it is used for personalized pricing in a static model of horizontal differentiation (Thisse and Vives, 1988) or vertical differentiation (Choudhary et al., 2005).[8] This insight is extended to a dynamic model of behavior-based pricing by Choe et al. (2018) and product personalization by Zhang (2011). Chen and Iyer (2002) studies the firm's decision to invest in customer ad-

[7] Choe et al. (2022) considers a standard model of behavior-based price discrimination, but adds a stage when firms can agree to share customer information before price competition begins. They show that information sharing obtains in equilibrium because it softens competition when information is gathered, although it intensifies competition when shared information is used in subsequent competition.

[8] These results are based on the case where the market is fully covered. When it is not, Rhodes and Zhou (2022) shows the results are reversed.

dressability which enables personalized pricing. Shaffer and Zhang (2002) allows firm heterogeneity and Chen et al. (2020) considers consumers' identity management that can bypass firms' attempt to price discriminate. Fudenberg and Villas-Boas (2012) provides a comprehensive survey of the earlier literature, and Ezrachi and Stucke (2016) provide various examples of personalized pricing in practice. Our work contributes to this literature by analyzing the implications of unilateral data sharing when it is used for personalized pricing, which is not formally studied in the existing literature.

The rest of the paper is organized as follows. The baseline model is described in Section 2 and analyzed in Section 3. Section 4 studies several extensions and variations of the baseline model explained previously. In Section 5, we provide evidence on B2B data data sharing in practice and the implications for management that our analysis sheds light on. Section 6 concludes the paper with discussions on the directions for future research. Appendix contains deferred proofs while the online appendix includes proofs of the results in Section 4 as well as some additional discussions.

2 The model

We adopt a standard location model with a general distribution for consumer preferences. There is a continuum of consumers on $[0, 1]$. A consumer's preference is indexed by her location $x \in [0, 1]$ where x has a smooth, cumulative distribution function F and a strictly positive density f. We assume F satisfies the monotone hazard rate condition, i.e., $h(x) := f(x)/(1 - F(x))$ is strictly increasing in x. We call the consumer located at x simply consumer x. There are two firms. Firm 1 is located at 0 and has full information about consumers' preferences. Firm 2 is located at 1 and initially has no information about consumers' preferences.

A consumer derives a gross value v from a product from either firm and incurs a transportation cost d_i, the distance between her location and firm i's location. So if a consumer buys a product from firm i at price p_i, then she obtains utility $v - p_i - d_i$. We assume $v \geq 2$ so that the market is fully covered.[9] Firm 1 can costlessly share data with firm 2. When a consumer's data is shared, firm 2 knows the exact location of that consumer, called a targeted consumer, based on which to exercise price discrimination.

[9]In Section 4.7, we discuss the case where the market is not fully covered.

We allow firm 1 to choose any market segment for data sharing. For example, it may choose to share data on consumers in several disjoint segments. However, as we show later in Lemma 1, firm 1 weakly prefers sharing data on an interval to any other forms of data sharing. This leads us to a simple form of data sharing: firm 1 chooses an interval $[a, b] \subseteq [0, 1]$ for data sharing. For clarity of exposition, we do not consider any monetary payment between the two firms in the baseline model.[10] With its full information, firm 1 sets personalized price $p_1(x)$ for each consumer $x \in [0, 1]$. Without data sharing, firm 2 can choose only a uniform price, denoted by p_2. Given data sharing on $[a, b]$, firm 2 chooses personalized price $p_2(x)$ for $x \in [a, b]$ and uniform price p_2 for the rest of the consumers. To simplify notation, we normalize the cost of production to zero.[11]

The timing of the game is as follows. First, firm 1 chooses $[a, b]$. Second, firm 2 decides whether or not to accept the shared data. If the shared data is rejected, then we have the status quo. If firm 2 accepts the shared data, then the subsequent pricing game proceeds as follows. Firm 2 moves first by choosing its uniform price where relevant. Following this, the two firms simultaneously make private offers of personalized prices to consumers they have data on. The sequential timing in price offers is standard in the literature on personalized pricing (Thisse and Vives, 1988; Shaffer and Zhang, 2002; Choe et al., 2018). This not only reflects the flexibility in choosing personalized prices, but also allows us to solve for the subgame perfect Nash equilibrium in pure strategies.[12] We assume that firm 2 can prevent consumers who are given its private offers from choosing its uniform price. This can be done, for example, by redirecting the product search by a consumer, a practice called steering or search discrimination. Relaxing this assumption does not change our general insight and further benefits firm 1, as we show in Section 4.1.

Remark 1. Our assumption that firm 1 has full information and firm 2 has no information is made for simplicity. As will become clear, our main insight continues to hold if firm 1 has information on $[0, m]$ with $m \in (z^N, 1)$ where z^N is the location of

[10]In Section 4.4, we extend our model to the case where firm 1 can charge a fixed sales fee to firm 2, a practice often adopted by dual-mode platforms. In Section 4.5, we discuss the case where firm 1 can sell data to firm 2. We show that our main insight remains robust to these extensions.

[11]We can modify the model to allow for cost differences between the two firms. For example, firm i has a constant marginal cost c_i with $c_1 \neq c_2$. This will not change our main insight insofar as both firms are active in equilibrium with or without data sharing.

[12]Following Chen et al. (2020, Section 5.5), we can show that all equilibria in our model are in mixed strategies when firms simultaneously choose uniform and personalized prices.

marginal consumer in the benchmark without data sharing. Then, firm 1 can choose to share data on a subset of $[z^N, m]$, which can benefit both firms. Likewise, mutually beneficial data sharing exists even when firm 2 has information on some consumers close to its location, say, $[n, 1]$ where $n > z^N$. In this case, firm 2's pricing decisions on $[n, 1]$ are delinked from its pricing decisions on $[0, n]$. Thus, firm 1 can choose mutually beneficial data sharing on a subset of $[z^N, n']$ for some $n' < n$.

Remark 2. In our baseline model, the shared data is used for personalized pricing. In Section 4.2, we analyze the case where the shared data is used for third-degree price discrimination. Although we do not consider the general information design problem where firm 1 can choose the informativeness of shared data by choosing a disclosure strategy, which is a mapping from the Hotelling line to a distribution over message space,[13] we conjecture that our main insight may hold for general data sharing strategies. It is because our results hold for the two polar cases, the most informative data sharing analyzed in Section 3 and the least informative data sharing analyzed in Section 4.2. Analysis of the two polar cases can be useful in studying more general cases.

3 Analysis

3.1 Benchmarks

We start with two benchmark cases. First, without data sharing, the marginal consumer z satisfies $p_1(z) + z = p_2 + (1 - z)$. Since firm 1 can lower $p_1(z)$ down to zero, we have $z = (1 + p_2)/2$. Firm 2 chooses $p_2 = 2z - 1$ to maximize profit $\pi_2 = p_2(1 - F(z))$. From the first-order condition,[14] we obtain the following equations for firm 2's optimal price denoted by p_2^N and the marginal consumer's location denoted by z^N:

$$p_2^N := \frac{2}{h(z^N)}, \quad z^N := \frac{1}{2} + \frac{1}{h(z^N)}. \tag{1}$$

Since firm 2 can serve the segment $[z^N, 1]$ even without data sharing, we call it *firm 2's customer base*. Given p_2^N, firm 1 chooses personalized price that leaves all consumers on $[0, z^N]$ indifferent between choosing either firm, i.e., $p_1^N(x) = \max\{p_2^N + (1 -$

[13] See, for example, Bergemann and Morris (2019).

[14] The second-order condition is satisfied due to the monotone hazard rate condition.

$2x), 0\}$.[15] In the benchmark without data sharing, firms earn profits given by

$$\pi_1^N := \int_0^{z^N} (p_2^N + (1-2x))dF(x), \quad \pi_2^N := p_2^N(1-F(z^N)). \tag{2}$$

Second, if firm 1 shares data on all consumers, then we have the outcome in Thisse and Vives (1988): firm 1 serves $[0, 1/2]$ with $p_1(x) = 1 - 2x$ and firm 2 serves the rest with $p_2(x) = 2x - 1$. Clearly, firm 1 is worse off compared to the case without data sharing. Not only firm 1's market share shrinks from $[0, z^N]$ to $[0, 1/2]$ but also personalized prices firm 1 charges to consumers it serves decrease from $p_1^N(x) = p_2^N + (1 - 2x)$ to $p_1(x) = 1 - 2x$. Firm 2 is better off compared to the benchmark without data sharing. It is because firm 2 now serves additional consumers on $[1/2, z^N]$ while serving all consumers on $[z^N, 1]$ at personalized prices higher than its uniform price in the absence of data sharing, i.e., $p_2(x) = 2x - 1 \geq p_2^N = 2z^N - 1$ for all $x \in [z^N, 1]$.

3.2 Mutually beneficial data sharing

We now turn to the full game and analyze when data sharing can benefit both firms, hence is accepted by firm 2. Let us start with the following observations. First, given data sharing on $[a, b]$, firm 2's problem of choosing its uniform price can be delinked from its choice of personalized prices, due to our assumption that firm 2 can exercise search discrimination. This intensifies competition on $[a, b]$. Consequently, firm 1's profit from the segment $[a, b]$ cannot be higher after data sharing. Second, firm 1 will set its personalized price that will leave each consumer it serves indifferent between choosing either firm. That is, given firm 2's uniform price p_2, firm 1 will set $p_1(x) = \max\{p_2 + (1 - 2x), 0\}$. Thus, higher p_2 benefits firm 1 by allowing it to raise its personalized prices. Put together, these observations imply that firm 1 can benefit from data sharing only if it softens competition for consumers whose data is not shared and induces firm 2 to raise its uniform price above p_2^N. From these observations, we can show that it is sufficient to focus on the case in which firm 1 shares data on an interval.

Lemma 1 *Firm 1 weakly prefers sharing data on an interval to any other forms of data sharing.*

[15] In case of indifference, we assume consumers choose firm 1.

Proof: See the appendix.

Given that $[z^N, 1]$ is firm 2's customer base, it is easy to see that, for data sharing on $[a, b]$ to benefit firm 1, we must have $a > 1/2$ and $b > z^N$. The reasoning is as follows. First, sharing data on $[a, b]$ with $b \leq z^N$ intensifies competition on this segment without the benefit of raising firm 2's uniform price above p_2^N. Second, sharing data on $[a, b]$ with $a \leq 1/2$ is dominated by sharing data on only $[1/2, b]$ because the former intensifies competition on $[a, 1/2]$ without increasing firm 2's uniform price. Henceforth, we restrict analysis to the following set of (a, b).

$$\mathcal{S} := \{(a, b) \mid 1/2 < a < b,\ z^N < b\} \tag{3}$$

Following data sharing, firm 2's choice of uniform price depends on which segment of consumers it intends to serve with uniform price. There are two possibilities. First, if firm 2 serves only $[b, 1]$, then it chooses $p_{21} = 2b - 1$ and earns profit

$$\pi_{21}(b) = (2b - 1)(1 - F(b)). \tag{4}$$

Second, if it serves $[z', a] \cup [b, 1]$ for some $z' \leq a$, then it chooses $p_{22} = 2z' - 1$ and earns profit $(2z' - 1)(1 - F(b) + F(a) - F(z'))$. Maximizing this leads to

$$p_{22} = \frac{2\,(1 - F(b) + F(a) - F(z'))}{f(z')}, \quad z' = \frac{1}{2} + \frac{1 - F(b) + F(a) - F(z')}{f(z')} > \frac{1}{2}. \tag{5}$$

Denote the resulting profit by $\pi_{22}(a, b)$. The next lemma shows that, for firm 1 to benefit from sharing data on $[a, b]$, firm 2's uniform price must be $p_{21} = 2b - 1$.

Lemma 2 *For data sharing on* $[a, b]$ *to be profitable for firm 1, it should necessarily induce firm 2 to choose its uniform price to serve consumers on* $[b, 1]$ *only.*

Proof: See the appendix.

We provide below the intuition for Lemma 2. Given $a > 1/2$, data sharing on $[a, b]$ implies that firm 2 will serve the entire segment $[a, b]$. Therefore, the only way firm 1 can benefit from data sharing is to induce firm 2 to raise its uniform price above p_2^N. This is possible when firm 2 chooses its uniform price to serve consumers on $[b, 1]$ only. However, if firm 2 chooses its uniform price to serve additional consumers on $[z', a]$, then firm 2's uniform price is lower than p_2^N. It is because F satisfies the monotone hazard rate condition, which implies $z' < z^N$ as shown in the proof of Lemma 2. In this case, not only does firm 1 concede its market share to firm 2, but its personalized prices also decrease due to a lower uniform price chosen by firm 2. To summarize, Lemma 2 implies firm 2's incentive compatibility constraint that any data sharing by firm 1 needs to satisfy:

$$\text{(IC) } \pi_{21}(b) \geq \pi_{22}(a, b). \tag{6}$$

In addition, data sharing needs to make firm 2 weakly better off than in the benchmark without data sharing; or else, firm 2 will ignore the shared data. When firm 2 serves consumers on $[a, b]$ with personalized prices $p_2(x) = 2x - 1$ and those on $[b, 1]$ with uniform price $p_{21} = 2b - 1$, its profit is $\pi_2(a, b) = \int_a^b (2x - 1)dF(x) + \pi_{21}(b)$. Thus we have the following individual rationality constraint for firm 2:

$$\text{(IR) } \pi_2(a, b) = \int_a^b (2x - 1)dF(x) + \pi_{21}(b) \geq \pi_2^N. \tag{7}$$

When data sharing on $[a, b]$ satisfies (IC) and (IR), firm 1 serves all consumers on $[0, a]$ with personalized prices $p_1(x) = p_{21} + (1 - 2x) = 2(b - x)$, while firm 2 serves the rest of the market. Thus, firm 1's profit is given by

$$\pi_1(a, b) = \int_0^a (p_{21} + (1 - 2x))dF(x) = 2bF(a) - 2\int_0^a xdF(x). \tag{8}$$

Differentiating $\pi_1(a, b)$ in (8), we find $\partial\pi_1(a, b)/\partial a > 0$ and $\partial\pi_1(a, b)/\partial b > 0$ for all $b > a > 1/2$. Thus, firm 1's profit increases in a because it shares less data, and it increases in b because higher b leads to a higher uniform price by firm 2. Likewise, differentiating $\pi_2(a, b)$ in (7), we find $\partial\pi_2(a, b)/\partial a < 0$ and $\partial\pi_2(a, b)/\partial b > 0$ for all $b > a > 1/2$. Thus, firm 2's profit decreases in a because of less data sharing, and it increases in b because of more data sharing. Moreover, in the proof of Proposition 1 below, we show $\pi_1(a, b) > \pi_1^N$ and $\pi_2(a, b) > \pi_2^N$ for all $b > a = z^N$. This implies that, insofar as firm 2's (IC) is satisfied, there exists data sharing $[z^N, b']$ with $b' \in (z^N, 1)$, under which both firms have higher profits than in the benchmark without data sharing.

Proposition 1 *For any distribution F that satisfies the monotone hazard rate condition, we have $\pi_1(a, b) > \pi_1^N$ and $\pi_2(a, b) > \pi_2^N$ for all $b > a = z^N$. In addition, there* *exists $b' \in (z^N, 1)$ such that firm 2's (IC) is satisfied under data sharing on $[z^N, b']$, hence sharing data on $[z^N, b']$ makes both firms better off compared to the benchmark without data sharing.*

Proof: See the appendix.

The intuition for the above proposition is as follows. Firm 1 can benefit from sharing data on consumers who are in firm 2's customer base, whom firm 1 cannot serve even without data sharing. That is, sharing data on $[z^N, b']$ does not cost firm 1 any market share. But firm 1 should keep data on consumers who have high loyalty to firm 2. By keeping data on $[b', 1]$ to itself, firm 1 can induce firm 2 to raise its uniform price to $p_{21} = 2b' - 1 > p_2^N$.

Sharing data in this way softens competition as firm 2 will try to extract surplus from the consumers with high loyalty by charging a high uniform price. This allows firm 1 to raise its own personalized prices for consumers whom firm 1 continues to serve after data sharing. Thus firm 1 benefits from data sharing through higher personalized prices but at no cost of reduced market share. Firm 2 also benefits from such data sharing because it serves the same set of consumers but at higher prices for all of them than without data sharing.

3.3 Optimal data sharing

In the previous section, we showed that mutually beneficial data sharing exists for any distribution F. We now turn to the discussion of the optimal data sharing that firm 1 will choose. Firm 1's problem can be stated as follows:

$$\max_{(a,b)\in\mathcal{S}} \pi_1(a, b) \text{ subject to (IC) and (IR).} \tag{9}$$

We already know that the following conditions must be met for the optimal data sharing. First, data sharing has to be on an interval $[a, b]$. Second, we must have $1/2 < a$ and $z^N < b < 1$. Third, data sharing should induce firm 2 to choose a uniform price $2b - 1$, which is accepted by consumers on $[b, 1]$ only. These necessary conditions characterize the optimal data sharing in reasonable details, and they do not depend on the underlying distribution F. But one remaining issue is whether $a < z^N$ or $a \geq z^N$ under the optimal data sharing, which depends on F as we discuss below.

When $a \geq z^N$, data sharing does not entail any cost of reduced market share for firm 1. When $a < z^N$, firm 1 sacrifices some market share if doing so increases firm 2's uniform price, which more than compensates for the reduced market share. Clearly, whether a is above or below z^N depends on the shape of distribution F. Recall that firm 2 chooses between a high uniform price p_{21} to serve $[b, 1]$ and a low uniform price p_{22} to serve $[z', a] \cup [b, 1]$. Firm 2's (IC) requires that the former is more profitable for firm 2 than the latter. Given that $b > z^N > 1/2$, this is likely to be the case when F is left-skewed with a concentrated density around $[z^N, 1]$. Moreover, firm 1's profit strictly increases in b when firm 2 chooses the high uniform price p_{21}. Therefore, when F is left-skewed, firm 1 can maintain large b, hence induce the high uniform price p_{21}, without losing market share by setting $a < z^N$.

Firm 1 may concede additional consumers to firm 2 by setting $a < z^N$ if F is not highly left-skewed. Many distributions are not highly left-skewed, including all distributions symmetric around 1/2 and hence neither left- nor right-skewed. Firm 1 then optimally reduces a below z^N (i.e., losing the marginal consumer) in order to maintain a large b (i.e., a high profit margin on all inframarginal consumers) without violating firm 2's (IC). To show this argument more formally, we first fully characterize the optimal data sharing under the uniform distribution, showing that the optimal a is indeed set strictly below z^N. We then numerically solve for the optimal data sharing for a family of distribution functions to further support our main insight.

Suppose F is a uniform distribution, hence symmetric around 1/2. Then, $F(x) = x$, and it is straightforward to calculate the equilibrium outcome in the benchmark without data sharing: $z^N = 3/4$, $p_2^N = 1/2$, $\pi_1^N = 9/16$, and $\pi_2^N = 1/8$. Given data sharing on $[a, b]$, we have $\pi_{21}(b) = (2b - 1)(1 - b)$, $\pi_{22}(a, b) = (1 + 2a - 2b)^2/8$, and $\pi_1 = 2ab - a^2$. Based on the above, we can solve firm 1's problem in (9), leading to the following equilibrium outcome for the data sharing game.

Proposition 2 *If F is a uniform distribution, then firm 1's optimal choice is given by $[a^*, b^*] \approx [0.71, 0.97]$.*

- *Firm 1 serves consumers on $[0, a^*]$ with personalized price $p_1^*(x) = 2b^* - 2x$ and earns profit $\pi_1^* \approx 0.87 > 9/16 = \pi_1^N$.*

- *Firm 2 serves consumers on $[a^*, b^*]$ with personalized price $p_2^*(x) = 2x - 1$, consumers on $[b^*, 1]$ with uniform price $p_2^* = 2b^* - 1$, and earns profit $\pi_2^* \approx 0.21 > 1/8 = \pi_2^N$.*

Proof: See the appendix.

Notice that we have $a^* < z^N$ when F is uniform. This confirms our earlier observation that the optimal data sharing involves choosing $a < z^N$. That is, firm 1 sacrifices its market share by sharing data on some consumers outside firm 2's customer base. But firm 1's loss from giving away market share $[a^*, z^N]$ is smaller than the gain from inducing a higher uniform price from firm 2. This is because $a^* > 1/2$ and firm 1 can set higher personalized prices for all consumers on $[0, a^*]$ but its loss is limited to $[a^*, z^N]$.

We now consider a family of distribution functions, which all satisfy the monotone hazard rate condition, to demonstrate that $a^* \geq z^N$ unless F is too highly left-skewed, and $a^* < z^N$, otherwise. Specifically, we consider the beta distribution with shape parameters (α, β), i.e., $x \sim Beta(\alpha, \beta)$. The probability density function of $Beta(\alpha, \beta)$ can take on various shapes and it nests many continuous distributions as special or limiting cases, including uniform, exponential, power, and normal distributions. In Table 1, we present the numerical solution (a^*, b^*) to firm 1's optimal data sharing problem for five examples of beta distribution. As shown in Table 1, the optimal data sharing involves $a^* < z^N$ in all cases except when $(\alpha, \beta) = (5, 1)$, the latter being the case with an increasing density function, or a highly left-skewed distribution. This supports our claim that $a^* \geq z^N$ is likely when F is highly left-skewed. To examine this further, we plot a^* and z^N in Figure 1 for the case where $\beta = 1$ and α increases from 1 to 5. The solid line labeled z^N and the dotted line labeled a^* cross once when $\alpha \approx 2.1$. Thus, as α increases, hence F becomes more left-skewed, we have $a^* > z^N$ for all values of α greater than 2.1.[16]

— Insert Table 1 and Figure 1 about here. —

3.4 Welfare implications

How does data sharing affect consumer surplus and total surplus? Consider first total surplus. Given full market coverage, total surplus can be proxied by the average distance travelled by a consumer, which is minimized when the marginal consumer

[16]When F is a beta distribution, we can numerically show (i) $a^* < z^N$ for all $\alpha = \beta$, i.e., F is symmetric around 1/2, (ii) $a^* > z^N$ if $s < -1$ where s is the skewness of F, (iii) a^* monotonically increases as α increases holding β fixed, or as β decreases holding α fixed, i.e., F becomes more left-skewed. The details are available from the authors.

is at 1/2 because firms are located at 0 and 1.[17] It then follows that total surplus increases if $1/2 < a^* < z^N$ and decreases if $a^* > z^N$. In the former case, the reduced preference mismatch by firms and consumers is a salutary effect of data sharing. If $a^* < z^N$, then data sharing has a positive quality-of-matching effect by allowing consumers on $[a^*, z^N]$ to be matched to firm 2, their preferred choice. If $a^* > z^N$, then data sharing has a negative quality-of-matching effect by leading consumers on $[z^N, a^*]$ to be matched to firm 1, their less preferred choice. The effect of data sharing on total surplus is shown in the last column in Table 1 for various beta distributions.

Consider next consumer surplus. Observe that all consumers staying with the same firm after data sharing are worse off because they incur the same transportation cost, but pay higher price after data sharing. For these consumers, the quality-of-matching effect stays the same but the surplus-extraction effect works against them. But consumers who switch firms may or may not be worse off depending on different equilibria and how the quality-of-matching and surplus-extraction effects are balanced against each other.

Consider the case where $a^* \geq z^N$. Then, consumer $x \in [z^N, a^*]$ switches to firm 1 after data sharing. Thus, for consumer $x \in [z^N, a^*]$, the quality-of-matching effect is given by $(1 - x) - x = 1 - 2x < 0$, while the surplus-extraction effect is given by $p_1^*(x) - p_2^N$, which may or may not be negative. But the overall effect is negative since her total cost is $p_1^*(x) + x = 2b^* - x$ after data sharing and $p_2^N + (1 - x) = 2z^N - x$ before data sharing. The former is larger because $b^* > z^N$, hence all switching consumers are worse off after data sharing. Put together, data sharing reduces consumer surplus if $a^* \geq z^N$.

If $a^* < z^N$, then consumer $x \in [a^*, z^N]$ switches to firm 2 after data sharing, so incurs a total cost $p_2^*(x) + (1 - x) = x$. Before data sharing, her total cost is $p_1^N(x) + x = 2z^N - x$. The latter is larger because $x \leq z^N$, hence all switching consumers are better off after data sharing. The effect on consumer surplus then depends on comparing the losses for consumers who stay with the same firm and the gains for those who switch firms. We show in the proof of Proposition 3 that the losses outweigh the gains so that data sharing also reduces consumer surplus when $a^* < z^N$. For example, when F is a uniform distribution, we have $a^* < z^N$ as shown in Proposition 2, which implies that the optimal data sharing decreases consumer

[17] Let z be the location of the marginal consumer. Then the average distance travelled by a consumer is $\int_0^z x dF(x) + \int_z^1 (1 - x) dF(x)$, which is minimized when $z = 1/2$.

surplus but increases total surplus. We summarize the above discussions below.

Proposition 3 *Suppose firm 1's optimal data sharing is given by* $[a^*, b^*]$. *Then, for any distribution* F *that satisfies the monotone hazard rate condition, consumer surplus decreases unambiguously, and total surplus decreases if and only if* $a^* \geq z^N$.

Proof: See the appendix.

Our analysis offers clear implications for competition policy in regards to data sharing between competitors. While data sharing unambiguously hurts consumers,[18] it has a salutary effect of increasing total surplus if a dominant, data-rich firm concedes some of its market share to its competitor as a result of sharing data. The increase in total surplus in this case is due to reduced preference mismatch: holding prices equal, firm 2 is a preferred choice for consumers who switch from firm 1 to firm 2 after data sharing. However, if data sharing enables the dominant, data-rich firm to further increase its market share, i.e., $a^* \geq z^N$, then both consumer surplus and total surplus are adversely affected.[19]

4 Extensions

4.1 No search discrimination

In this section, we discuss the case where firm 2 cannot engage in search discrimination. That is, firm 2 cannot prevent a consumer who is offered its personalized price from choosing its uniform price. Given data sharing on $[a, b]$, consumer $x \in [a, b]$ then chooses the lower of the two prices offered by firm 2, i.e., $\min\{p_2(x), p_2\}$. Thus, firm 2's uniform price acts as a cap on its personalized prices.

Suppose firm 1 chooses to share data on $[a, b]$. It is easy to see that, when firm 2 cannot search discriminate, it does not have incentives to choose a low uniform price to serve consumers on $[z', a]$. If it does, then all consumers on $[a, b]$ will choose the

[18] Using the competition-in-utility-space approach, de Cornière and Taylor (2020) obtains a similar result that giving firms more data can be anti-competitive as it makes surplus extraction more efficient.

[19] Somewhat related, Armstrong and Zhou (2022) shows that more data can reduce preference mismatch but soften competition by amplifying perceived product differentiation. In our model, the effect of data sharing on match quality is driven entirely by price competition because product differentiation is exogenously given.

lower of the uniform price and personalized prices. As a result, firm 2's profit is at most $(2z' - 1)(1 - F(z'))$. But this is smaller than $\pi_2^N = (2z^N - 1)(1 - F(z^N))$ by the definition of z^N, which firm 2 can guarantee by ignoring the shared data. This implies that, when firm 2 cannot search discriminate, its optimal pricing decision automatically satisfies (IC). Consequently, if firm 2 accepts the shared data on $[a, b]$, then it will choose a uniform price $p_2 = 2b-1$ for consumers on $[b, 1]$ and personalized prices $p_2(x) = 2x - 1$ for consumers on $[a, b]$. In sum, although firm 2's inability to search discriminate gives consumers on $[a, b]$ an extra option of purchasing at uniform price, firm 2 instead charges a high uniform price, rendering the extra option unattractive. This leads to firm 2's profit given by

$$\pi_2(a, b) = \int_a^b (2x - 1)dF(x) + (2b - 1)(1 - F(b)).$$

Because firm 1 benefits when $p_2 = 2b-1$ increases, the solution to firm 1's problem is then $(\hat{a}, \hat{b})$ such that $\hat{b}$ is as large as possible and $\pi_2(\hat{a}, \hat{b}) = \pi_2^N$. This gives us $\hat{b} = 1$, and $\hat{a}$ such that $\pi_2(\hat{a}, 1) = \pi_2^N$. On the equilibrium path, firm 2 uses only personalized prices. But the equilibrium is supported by its off-the-path uniform price $p_2 \geq 1$; if $p_2 < 1$, then consumers on $[(p_2 + 1)/2, 1]$ will deviate and choose firm 2's uniform price. Because $p_2 \geq p_2^* = 2b^* - 1$, firm 1's personalized prices increase for all its consumers. Moreover, we can show $\hat{a} > a^*$ from $\pi_2(\hat{a}, 1) = \pi_2^N$. Thus, firm 1 is strictly better off in the absence of search discrimination.

Compared to the case without data sharing, it is straightforward to check that all consumers are worse off. The total surplus is also lower simply because $\hat{a} > z^N$, implying that there is more preference mismatch. Compared to the case with data sharing and search discrimination, consumer surplus is lower because prices rise for all consumers and total surplus is also lower because $\hat{a} > a^*$. To summarize, banning search discrimination can enable the dominant, data-rich firm to fine-tune its data sharing strategy, which is then used by its competitor to further soften competition than when search discrimination is allowed. This reduces both consumer surplus and total surplus.

Proposition 4 *Suppose firm 2 cannot engage in search discrimination.*

- *Firm 1's optimal data sharing is given by* $[\hat{a}, 1]$ *where* $\hat{a}$ *solves* $\int_{\hat{a}}^1 (2x-1)dF(x) = \pi_2^N$*, and* $\hat{a} > \max\{z^N, a^*\}$.

- *Firm 1 serves consumers on* $[0, \hat{a}]$ *with personalized prices* $\hat{p}_1(x) = \hat{p}_2 + (1-2x) \geq 2 - 2x$ *and earns strictly higher profit than the case with search discrimination, where* $\hat{p}_2 \geq 1$ *is firm 2's uniform price.*

- *Firm 2 serves consumers on* $[\hat{a}, 1]$ *with personalized prices* $\hat{p}_2(x) = 2x - 1$*, and earns the same profit as in the benchmark without data sharing.*

- *Both consumer surplus and total surplus are lower compared to the case without data sharing, as well as the case with data sharing and search discrimination.*

Proof: See the appendix.

4.2 Data sharing for third-degree price discrimination

Even when firm 1 shares data, firm 2 may not have capabilities such as data analytics that may be necessary to process the shared data for personalized pricing. This reduces the value of data for firm 2 and may allow it to exercise only third-degree price discrimination, choosing one uniform price for the segment with shared data and another uniform price for the rest of the market. But firm 1 continues to use personalized pricing for all consumers. We analyze this case in this section.

As discussed in the previous section, firm 1 can benefit from data sharing only when it induces firm 2 to increase uniform price above p_2^N. Given data sharing on $[a, b]$ with $a \geq 1/2$, denote firm 2's uniform price for the segment $[a, b]$ by q_{2l}, and that for the rest of the market by q_{2h}. On $[a, b]$, the marginal consumer's location, denoted by z_l, satisfies $p_1(z_l) + z_l = q_{2l} + 1 - z_l$ with $p_1(z_l) = 0$, implying $q_{2l} = 2z_l - 1$. Suppose $b < z^N$. Then we have $q_{2l} = 2z_l - 1 < 2z^N - 1 = p_2^N$. In addition, firm 2 chooses $q_{2h} = p_2^N$ and serves $[z^N, 1]$ by the definition of p_2^N. Thus, when $b < z^N$, firm 1 cannot induce firm 2 to increase its price above p_2^N. In what follows, we focus on the case with $b \geq z^N$.

Consider the segment $[a, b]$. Firm 2 chooses q_{2l} to maximize profit from this segment given by $\pi_{2l} = q_{2l}\left(F(b) - F(z_l)\right)$. From the first-order condition for an interior solution, we obtain

$$q_{2l} = \frac{2\left(F(b) - F(z_l)\right)}{f(z_l)} \text{ where } z_l = \frac{1}{2} + \frac{F(b) - F(z_l)}{f(z_l)} \in (a, b).$$

Otherwise, we have a boundary solution, $q_{2l} = 2a - 1$ and $z_l = a$, hence $\pi_{2l} =$

$q_{2l}(F(b) - F(z_l))$. Clearly, it is not optimal for firm 2 to charge $q_{2l} < 2a-1$, in which case it serves the same set of consumers but at a lower price. Likewise, firm 2 never charges price $q_{2l} \geq 2b - 1$ because it leads to $\pi_{2l} = 0$.

For the rest of the market, firm 2 chooses q_{2h}. As before, firm 2's choice of q_{2h} depends on whether it serves only $[b, 1]$, hence $q_{2h} = 2b - 1$ with corresponding profit $\pi_{21}(b)$, or it serves additional consumers on $[z', a]$ for some $z' < a$ with corresponding profit $\pi_{22}(a, b)$. The following lemma shows that firm 1's optimal data sharing should necessarily satisfy firm 2's (IC): $\pi_{21}(a) \geq \pi_{22}(a, b)$. In addition, it is not optimal for firm 1 to share data on $[a, b]$ and then serve some consumers on that segment. It is because firm 1's personalized prices on $[a, b]$ depend on firm 2's low uniform price q_{2l}; firm 1 is better off by not sharing data on these consumers so that it can charge personalized prices that depend on firm 2's high uniform price q_{2h}. Lemma 3 summarizes these observations.

Lemma 3 *Firm 1's optimal data sharing on $[a, b]$ should necessarily satisfy firm 2's (IC) $\pi_{21}(b) \geq \pi_{22}(a, b)$, and induce firm 2 to serve the entire segment $[a, b]$ with uniform price $q_{2l} = 2a - 1$.*

Proof: See the online appendix.

By Lemma 3, we can write firm 2's profit under the optimal data sharing as $\pi_2(a, b) = (2a - 1)(F(b) - F(a)) + \pi_{21}(b)$. Thus we have the following individual rationality constraint for firm 2:

$$\text{(IR) } \pi_2(a, b) = (2a - 1)(F(b) - F(a)) + (2b - 1)(1 - F(b)) \geq \pi_2^N. \tag{10}$$

Then, firm 1's profit under the optimal data sharing is

$$\pi_1(a, b) = \int_0^a (2b - 2x) dF(x).$$

Firm 1's problem can be stated as follows:

$$\max_{(a,b)} \pi_1(a, b) \text{ subject to } b \geq z^N, \ \pi_{21}(b) \geq \pi_{22}(a, b), \text{ and (10)}. \tag{11}$$

Denote the solution to (11) by (a^{**}, b^{**}), and the solution to firm 1's problem under personalized pricing stated in (9) by (a^*, b^*). Clearly, firm 1's profit function is the same in both (9) and (11). Firm 2's (IC) is also the same. The only difference is that firm 2's (IR) is harder to satisfy in (11) than in (9). It is because third-degree price discrimination is a less effective tool than personalized pricing for firm 2 to extract surplus from targeted consumers: firm 2's profit from $[a, b]$ is smaller when it chooses $q_{2l} = 2a - 1$ to serve the entire segment than when it can choose personalized price $p_2(x) = 2x - 1$ for each targeted consumer. Therefore, an immediate conclusion is that, if (a^*, b^*) satisfies (10), then we must have $(a^{**}, b^{**}) = (a^*, b^*)$. In this case, neither firm 1's profit nor total surplus changes from those in Section 3.3, but firm 2's profit is smaller, which implies that consumer surplus is larger.

Proposition 5 *Suppose firm 2 uses the shared data for third-degree price discrimination and (10) holds under the optimal data sharing under personalized pricing. Then, compared to the case where firm 2 can use the shared data for personalized pricing,*

- *firm 1's optimal data sharing does not change,*
- *firm 1's profit under the optimal data sharing does not change but firm 2's profit decreases;*
- *total surplus does not change but consumer surplus increases.*

We can show that condition (10) is satisfied at (a^*, b^*) for many distributions satisfying the monotone hazard rate condition. For example, (10) holds for the uniform distribution, power distribution, triangular distribution when the peak point of $f(x)$ is to the right of $x = 0.56$, beta distribution with increasing $f(x)$, U-shaped $f(x)$, and unimodal $f(x)$ when $\alpha - \beta$ is positive and relatively large. Thus the following corollary is immediate.

Corollary *Suppose F is a uniform distribution and firm 2 uses the shared data for third-degree price discrimination. Then, firm 1's optimal choice of $[a, b]$ is the same as in Proposition 2, with the only difference that consumer surplus is higher under third-degree price discrimination.*

However, (10) may not hold at (a^*, b^*) when $f(x)$ has a very thin right tail. Even in this case, we can still show that there exists $b \in (z^N, 1)$ such that sharing data on $[z^N, b]$ is mutually beneficial, similar to Proposition 1. But we would need to solve numerically for the optimal data sharing for specific distribution functions, as in Section 3.2. In Figure 2, we plot the optimal data sharing under third-degree price discrimination for various beta distributions with $\alpha = 2$, which we label (a^{**}, b^{**}). Given $\alpha = 2$, condition (10) is satisfied at (a^*, b^*) for all $\beta \leq 1.9$. Thus, we have $(a^{**}, b^{**}) = (a^*, b^*)$ for $\alpha = 2$ and $\beta \leq 1.9$. But when $\beta > 1.9$, condition (10) fails at (a^*, b^*) and, as shown in Figure 2, we now have $a^* < a^{**} < z^N < b^{**} < b^*$. Thus, compared to the case with personalized pricing, firm 1 shares less data.

— Insert Figure 2 about here. —

4.3 Customer data as an entry barrier

Customer data can be a source of a firm's competitive advantage if it is utilized for data-enabled learning. Moreover, customer data is often an important input in digital markets and, therefore, can act as a barrier to entry. A data-rich incumbent can cement its position by offering better-targeted products thanks to data-enabled learning, which attracts more customers and more customer data, hence creating a self-enforcing loop. This has led to active policy discussions related to data mobility and data openness.[20] Our analysis sheds light on this by clarifying to what extent customer data can work as an entry barrier.

We modify our model as follows. The game is played over two periods, indexed by $t = 1, 2$. Firm 1, an incumbent, has data on all consumers in both periods while firm 2, a potential entrant, has access to customer data depending on how the game evolves. Firm 1's location is fixed at point zero in both periods and firm 2's location upon entry is fixed at point one.[21] In $t = 1$, firm 2 makes an entry decision, followed by firm 1's data sharing decision.[22] There is a fixed entry cost $E > 0$. If firm 2 enters, then the two firms compete as in our baseline model. At the end of $t = 1$, firm 2 collects data on customers it serves in addition to the data shared by firm 1, if any. In $t = 2$, following firm 2's entry, the two firms compete using the customer data they have for price discrimination. We focus on the case where F is a uniform

[20] See Crémer et al. ("EU Report", 2019), Furman et al. ("Furman Report", 2019), or the ACCESS Act in the US that passed through the House Committee on June 24, 2021.

[21] This is the same as setup as in Gehrig et al. (2011), whose focus is on firm 2's entry decision when firm 1 can price discriminate and there are consumer switching costs. Our focus is on how data sharing affects firm 2's entry decision.

[22] One may consider an alternative timeline in which firm 1 makes and announces its data sharing decision first, after which firm 2 makes an entry decision. This timeline can lead to time inconsistency whereby firm 1's announcement not to share data becomes not credible once firm 2 enters.

distribution, which facilitates clear discussions. To simplify notation, we assume away time discount. In what follows, we provide heuristic discussions only with the detailed analysis shown in the proof of Proposition 6.

To begin with, suppose firm 1 does not have any customer data. Then, upon firm 2's entry, the two firms compete as in standard behavior-based price discrimination with personalized pricing (Choe et al. 2018): they compete à la Hotelling in $t = 1$ and, in $t = 2$, customer data collected in $t = 1$ is used for price discrimination. As shown by Choe et al. (2018), the game has two pure-strategy equilibria, one being a mirror image of the other, from which we select the one that favors firm 1.[23] In that equilibrium, firm 1's total profit over two periods is $\Pi_1 = 109/169 \approx 0.645$ and firm 2's total profit is $\Pi_2 = 383/676 \approx 0.572$. Thus firm 2 enters if $E < 0.572$.

Suppose now firm 1 has data on all consumers. Without data sharing, the two firms compete in $t = 1$ as in our baseline model, leading to the marginal consumer z^N. In $t = 2$, firm 2 chooses a uniform price on $[0, z^N]$ and personalized prices on $[z^N, 1]$. Solving the two-period model gives us $z^N = 7/10$, $\Pi_1 = 17/20$, $\Pi_2 = 7/20$. Recall that, in the benchmark one-period model without data sharing, we have $z^N = 3/4$. Thus, firm 2's market share in $t = 1$ is larger in the two-period model. It is because firm 2 becomes more aggressive in $t = 1$ when it anticipates subsequent competition in $t = 2$: securing a larger market share in $t = 1$ enables firm 2 to extract more surplus through personalized pricing in $t = 2$. Indeed, without data sharing, firm 2's uniform price in $t = 1$ is $1/2$ in the one-period model, but it is $2/5$ in the two-period model. Compared to the case where firm 1 does not have any customer data, we can say that data plays a role as an entry barrier when $0.35 < E < 0.572$. But this ignores firm 1's incentives to share data once firm 2 enters. The possibility of beneficial data sharing mitigates the extent to which customer data can be an effective entry barrier. We discuss this below.

Suppose firm 2 enters and firm 1 chooses to share data on $[a, b]$. Then, in $t = 2$, firm 2 has access to data on $[a, b]$ as well as on customers it serves in $t = 1$. Solving the two-period model shows that the optimal data sharing is $[a^*, b^*] = [0.710, 0.943]$ with each firm's total profit given by $\Pi_1 = 1.201$, $\Pi_2 = 0.431$. Recall that, in the one-period model where F is a uniform distribution, we have $[a^*, b^*] = [0.71, 0.97]$. Thus, firm 1 shares less data in the two-period model, in particular, by withholding data on consumers who are most loyal to firm 2. It is because dynamic considerations make

[23] For example, firm 1 may choose its uniform price in $t = 1$ anticipating firm 2's entry.

firm 2 more aggressive in $t = 1$ as discussed above, which firm 1 counters by sharing less data. Nonetheless, firm 1's optimal data sharing is qualitatively the same as in the one-period model in that the data on firm 2's most loyal customers is withheld. Both firms benefit from the optimal data sharing and firm 2 enters if $E < 0.431$. Note that, without firm 1's optimal data sharing, firm 2 enters if $E < 0.35$. We summarize these discussions below.

Proposition 6 *Suppose firm 2 makes a decision to enter the market at a fixed entry cost E, after which firm 1 makes a decision to share data with firm 2.*

- *If firm 1 does not have customer data, then firm 2 enters if $E < 0.572$.*
- *If firm 1 is fully informed but data sharing is not possible, then firm 2 enters if $E < 0.35$. Thus, data acts as an entry barrier when $0.35 < E < 0.572$.*
- *If firm 1 is fully informed and data sharing is possible, then firm 2 enters if $E < 0.431$ and firm 1 shares data on $[0.710, 0.943]$. Thus, data acts as an entry barrier when $0.431 < E < 0.572$, and data sharing mitigates data's role as an entry barrier when $0.35 < E < 0.431$.*

Proof: See the online appendix.

4.4 When firm 1 charges a fee for sales

Many data-rich firms such as Alibaba or Amazon are digital platforms that operate in a dual mode: they run marketplaces for third-party sellers while selling their own products on the marketplaces.[24] The platform typically charges a fee for sales to third-party sellers. We extend our baseline model to consider a simple case where firm 1 charges a fixed per-unit sales fee to firm 2.[25]

Suppose firm 1 charges a per-unit fee $k \geq 0$. This induces firm 1 to price less aggressively for consumers loyal to firm 2. In the absence of sales fee, Bertrand competition for each consumer implies that firm 1 will choose zero personalized price for consumers more loyal to firm 2. Instead, firm 1 can cede these consumers to firm

[24] See Hagiu et al. (2022) for a general analysis of dual-mode platforms as well as many examples.

[25] Many platforms use a two-part fee structure that consists of fixed fee per item sold plus variable fee that is proportional to the sale amount. Examples include Amazon, eBay, Shopify, etc. Incorporating such a fee structure to our model is beyond the scope of current paper, which we leave for future work.

2 and earn k per each consumer. Given this, firm 2 will raise its price by shifting the sales fee entirely onto its customers. It is because consumers have inelastic demand insofar as firm 1 does not undercut firm 2. Since firm 1 can collect k from firm 2, it has no incentive to undercut firm 2 when firm 2 shifts entire k onto its customers. In addition, firm 2's higher uniform price allows firm 1 to increase its personalized prices. In short, the sales fee softens competition and raises all prices by k, which is then collected by firm 1. Other than that, the sales fee does not change equilibrium market shares, nor firm 1's optimal data sharing. We formalize this below.

Let us start with the case without data sharing. The marginal consumer z is given by $p_1(z) + z = p_2 + (1 - z)$. Since firm 1 should be indifferent between earning $p_1(z)$ from serving consumer z, or ceding consumer z to firm 2 and earning a fee k instead, we must have $p_1(z) = k$. From these two equalities, we obtain $p_2 = 2z - 1 + k$. Plugging this into firm 2's profit, we have

$$\pi_2 = (p_2 - k)(1 - F(z)) = (2z - 1)(1 - F(z)).$$

Observe that the above profit is exactly the same as firm 2's profit in our baseline model without data sharing. Thus, z^N remains the same, hence so does each firm's market share and firm 2's profit. The only change for firm 2 is that its uniform price increases by k to $2z^N - 1 + k$. But firm 1 is better off because it can charge higher personalized prices to all customers it serves and also earns a sales fee $k(1 - F(z^N))$. Specifically, firm 1's personalized price is $p_1(x) = p_2 + (1 - 2x) = 2z^N - 2x + k$, hence increases by k compared to the case without the sales fee. Thus, firm 1's profit is

$$\pi_1 = \int_0^{z^N} (2z^N - 1 + k)dF(x) + k(1 - F(z^N)) = \int_0^{z^N} (2z^N - 2x)dF(x) + k. \quad (12)$$

In the right hand side of the second equality in (12), the first term is firm 1's profit in the baseline model without data sharing. Thus, the sales fee has the effect of charging a price premium k to *every* consumer, which becomes additional profit for firm 1.

Suppose now firm 1 shares data on $[a, b]$. It is easy to see that beneficial data sharing should continue to have $a > 1/2$ and $b > z^N$. For consumer $x \in [a, b]$, firm 1's indifference condition is $p_1(x) = k$ and consumer x's indifference condition is $p_1(x) = p_2(x) + (1 - 2x)$. From these two, we obtain $p_2(x) = 2x - 1 + k$, and firm

2 serves all consumers on $[a, b]$.[26] For consumers outside $[a, b]$, firm 2's uniform price can be chosen to serve only $[b, 1]$ or additional consumers to the left of a. As in our baseline model, firm 1's optimal data sharing hinges on which uniform price firm 2 should be induced to choose.

First, if firm 2 chooses uniform price to serve only $[b, 1]$, then its uniform price is $p_{21} = 2b - 1 + k$. In this case, firm 1's profit, denoted by π_{11}, is

$$\pi_{11} = \int_0^a (2b - 2x + k)dF(x) + k(1 - F(a)) = \int_0^a (2b - 2x)dF(x) + k,$$

where the first term on the right hand side of the second equality is exactly the same as firm 1's profit given in (8). Second, if firm 2 chooses uniform price to serve $[z', a] \cup [b, 1]$ for some $z' \leq a$, then its uniform price is $p_{22} = 2z' - 1 + k$. Then firm 1's profit is $\pi_{12} = \int_0^{z'} (2z' - 2x)dF(x) + k$. Subtracting k from each, we have exactly the same profit for each case as in Section 3.2. In addition, one can easily check that firm 2's profit in each case is also exactly the same as that in Section 3.2. Thus, Lemma 2 applies and firm 2's (IC) is the same as before. Since the addition of constant term k to firm 1's profit does not change its optimization problem, it follows that firm 1's optimal data sharing remains the same as in our baseline model. We summarize the results in Proposition 7.

Proposition 7 *Suppose firm 1 charges firm 2 a per-unit fee for sales $k \geq 0$.*

- *In the equilibrium with or without data sharing, each firm's market share and firm 2's profit are the same as those without the sales fee, but all prices and firm 1's profit increase by k.*
- *Firm 1's optimal data sharing is the same with or without the sales fee.*

4.5 Data sharing with side-payment

We have so far focused on firm 1's incentive to voluntarily share data. We abstracted away the possibility of side-payment associated with data sharing in order to highlight the strategic benefits of data sharing. However, one may wonder how the outcome may change if firm 1 can charge for its shared data. We consider this case in this section. For simplicity, we assume that firm 1 makes a take-it-or-leave-it offer

[26] We assume consumers on $[a, b]$ choose firm 2 in case of indifference.

of a lump-sum, non-negative fee for the data, denoted by Φ. The timing is as follows: firm 1 chooses Φ and $[a, b]$; firm 2 accepts or rejects firm 1's offer; if firm 2 accepts firm 1's offer, then firm 2 pays Φ to firm 1 in exchange for the data on $[a, b]$; the two firms compete in price. The rest of the model remains the same as before, including the use of data for personalized pricing.

An immediate observation is that any mutually beneficial data sharing in the absence of side-payments continues to be mutually beneficial when side-payments are allowed. Firm 1 can simply share the same amount of data and use Φ to extract full additional profit from firm 2. That is, if sharing data on $[a, b]$ is mutually beneficial with corresponding profits $\pi_1(a, b) \geq \pi_1^N$ and $\pi_2(a, b) \geq \pi_2^N$, then firm 1 can continue to share data on $[a, b]$ and choose $\Phi = \pi_2(a, b) - \pi_2^N$. Thus, we expect firm 1 to be better off when side-payments are allowed.

A natural question is then whether firm 1 shares more data when side-payments are allowed. Intuition tells us it should because, given the extra tool to extract surplus from firm 2, firm 1 would be less stringent in sharing data. Since firm 1 can use Φ to extract any profit from firm 2, i.e., $\Phi = \pi_2(a, b) - \pi_2^N$, we can focus on data sharing that maximizes industry profit. That is, firm 1 chooses (a, b) to maximize $\pi_1(a, b) + \pi_2(a, b)$ subject to $\Phi = \pi_2(a, b) - \pi_2^N \geq 0$.

Suppose the optimal data sharing without side-payment is given by $[a^*, b^*]$. Suppose now firm 1 shares additional data on $[a', a^*]$ when side-payment is possible where $a' = a^* - \epsilon$ for arbitrarily small ϵ. Note that firm 2 chooses its uniform price to serve only $[b^*, 1]$ given data sharing on $[a^*, b^*]$. Since $a' < a^*$, it continues to do so when additional data on $[a', a^*]$ is shared. Thus, the only change in industry profit is from $[a', a^*]$. If it increases, then firm 1 shares more data when side-payment is allowed. In the absence of side-payment, consumer a' is served by firm 1 for profit $2b^* - 2a'$. After additional data sharing, she is served by firm 2 for profit $2a' - 1$. Thus, industry profit from $[a', a^*]$ increases if $a' > (2b^* + 1)/4$ or, given that ϵ is arbitrarily small, if $a^* > (2b^* + 1)/4$. In sum, if $a^* > (2b^* + 1)/4$ at the optimal data sharing $[a^*, b^*]$ without side-payment, then firm 1 can increase its profit by sharing more data on $[a', a^*]$ when side-payment is allowed. When $a^* \leq (2b^* + 1)/4$, additional data sharing on $[a', a^*]$ needs to be accompanied by additional data sharing on $[b^*, b']$ so that firm 2's uniform price increases. In this case, the optimal data sharing can be derived only numerically. For example, when F is uniform, the optimal data sharing with side-payment is $[a, b] = [0.66, 0.984]$. Note that, when F is uniform, the optimal data sharing without side-payment is $[a^*, b^*] = [0.71, 0.97]$. Thus, we have $a^* \leq (2b^* + 1)/4$ so that both a^* and b^* need to be adjusted in a way firm 1 shares more data when side-payment is allowed. To summarize, firm 1 shares more data and enjoys larger profit when side-payment is allowed.

4.6 Data broker

A data-poor firm may be able to obtain data from sources other than the data-rich firm, for example, from data brokers. Data brokers' optimal strategy of selling data to competing firms has been the focus of several recent studies (e.g., Montes et al., 2019; Bounie et al., 2021). Since firm 1 in our model is fully informed, the data broker can sell data only to firm 2, but this can potentially change firm 1's data-sharing strategy. We discuss this below, assuming F is a uniform distribution.

Suppose that a data broker can sell the full set of customer data to firm 2 at a fixed price P. If firm 2 purchases data from the data broker, then the two firms compete in personalized prices as in Thisse and Vives (1988), resulting in the gross profit of 1/4 for each firm. Firm 2's net profit is $1/4 - P$. Recall that, in the baseline model with the optimal data sharing, the equilibrium profits are $\pi_1^* = 0.87$ and $\pi_2^* = 0.21$. Thus, the optimal data sharing in the baseline model continues to be optimal if $P \geq 0.25 - 0.21 = 0.04$. However, if $P < 0.04$, then firm 2 has incentives to purchase data from the data broker, which decreases firm 1's profit from 0.87 to 0.25. Thus, firm 1 needs to adjust its data-sharing strategy to leave more surplus to firm 2 in order to avoid the tough competition in personalized prices. This leads to an expansion of $[a^*, b^*]$ relative to the baseline model. For each value of $P < 0.04$, the optimal data sharing can be solved numerically. For the case where F is a uniform distribution, one can check that, as P decreases from 0.04 to 0, firm 1 shares more data by decreasing a^* and increasing b^*.[27]

4.7 Partially-covered markets

In the baseline model, we assumed v is large enough so that the market is fully covered. We now discuss the case where the market is not fully covered even when it is served by both firms, which can lead to quite different outcomes when firms compete in personalized pricing (Rhodes and Zhou, 2022). Suppose $v < z^N$ where

[27]The calculation is available from the authors.

z^N is the marginal consumer in the benchmark without data sharing. Then, without data sharing, firm 2 serves $[z^N, 1]$ with uniform price $p_2^N = 2/h(z^N)$ as shown in (1), and firm 1 serves $[0, v]$ with personalized price $p_1(x) = v - x$ for all $x \in [0, v]$. Consequently, consumers on $[v, z^N]$ are not served by either firm, and the market is not fully covered without data sharing if $v < z^N$.

In this case, the uncovered segment $[v, z^N]$ separates the two firms, which makes them local monopolies. Thus, they do not compete with each other and their prices are no longer strategic complements. It then follows that firm 1 does not benefit when firm 2 increases its uniform price, implying that mutually beneficial data sharing does not exist. Nonetheless, data sharing can benefit firm 2 and increase total surplus by enabling firm 2 to serve consumers who would not be served without data sharing. For example, suppose $v = 1/2$ and F is uniform so that $z^N = 3/4$. Then, when firm 1 shares data on $[1/2, 3/4]$, firm 2 can profitably serve all consumers on $[1/2, 3/4]$ with personalized price $p_2(x) = x - 1/2 \geq 0$. This increases firm 2's profit but does not change consumer surplus, hence total surplus increases. Although firm 1's profit does not change in this case, firm 1 can also benefit from data sharing if it can charge firm 2 for data sharing.

4.8 When firm 1 cannot price discriminate

The main insight that unilateral data sharing can be mutually beneficial hinges on firm 2's ability to price discriminate after data sharing. It does not depend on firm 1's ability to price discriminate. We now show that the same insight continues to hold even when firm 1 charges a uniform price. Data sharing allows firm 2 to delink its pricing decisions on the segment with shared data and the rest of the market, the latter in competition with firm 1. As in our baseline model, suitably chosen data sharing can soften competition on the second segment, which allows firm 1 to raise its uniform price without losing its market share. For simplicity, our discussion below is based on the case where F is uniform.

Suppose firm 1 is fully informed but cannot price discriminate, for example, due to lack of data analytics necessary for price discrimination. Firm 2 does not have any data but has access to data analytics, and can exercise personalized pricing after data sharing. In the benchmark without data sharing, the symmetric Hotelling equilibrium prices are $p_1 = p_2 = 1$, firm 1 serves $[0, 1/2]$, and each firm earns profit $1/2$.

Suppose now firm 1 shares data on $[1/2, 3/5]$. We will show that this benefits both firms. Denote firm 1's uniform price by p'_1 and firm 2's uniform price p'_2. After data sharing, firm 2 will serve $[1/2, 3/5]$ with personalized prices, hence Hotelling competition is only on $[0, 1/2] \cup [3/5, 1]$. Suppose $p'_1 = 1 + \epsilon$ for sufficiently small $\epsilon > 0$. Firm 2 can undercut firm 1 slightly by choosing $p'_2 < 1 + \epsilon$ to serve $[\gamma, 1/2] \cup [3/5, 1]$ where γ is sufficiently close to 1/2 given that ϵ is sufficiently small. Or it can choose $p'_2 = p'_1 + 1/5 = 6/5 + \epsilon$ which makes consumer $x = 3/5$ indifferent, hence firm 2 serves only $[3/5, 1]$. Given that ϵ is sufficiently small, firm 2's profit from its uniform price is higher from the second choice. In addition, firm 2 can charge personalized price $p_2(x) = p'_1 + (2x - 1) > 1$ for all consumers on $[1/2, 3/5]$. Thus, firm 2 is better off after data sharing. Firm 1 is also better off because it serves the same set of consumers but charges a higher price.[28] The discussion above shows that our main insight continues to hold even when firm 1 cannot price discriminate; what matters is firm 2's ability to price discriminate after data sharing.

5 Discussions and Implications for Management

5.1 B2B data sharing in practice

Although data sharing is mandated in a few industries such as banking and energy,[29] the B2B data sharing considered in this paper is growing in many other industries. Based on a survey conducted during 2017-2018 in 31 countries in the European Economic Area, Arnaut et al. (2018a) reports examples of B2B data sharing in various industries such as automotive, transport and logistics, agriculture, telecom, etc. The survey finds that data sharing can take different forms ranging from unilateral to more collaborate approaches, and that most B2B data sharing occurs within their own business sector.[30] Particularly relevant to our model, the survey finds that a majority of participating companies choose to share a small proportion of data they

[28] In the online appendix, we provide a more formal argument for when data sharing on $[a, b]$ can be mutually beneficial by solving for equilibrium prices given $[a, b]$.

[29] In Europe, sector-specific data sharing is imposed in banking and finance, automotive, energy, electronic communications and postal services. See Feasey and de Streel (2020) for details.

[30] The type and frequency of data shared and the mechanism for data sharing vary depending on businesses and industries. For example, Green Energy Options (geo), a UK-based supplier of energy monitoring devices, shares data on energy use and efficiency on a real-time basis, by relying on FTP technologies or online data repositories. For more specific examples, see Arnaut et al. (2018b).

generate, rather than their complete datasets, depending on their business strategies. But the survey does not distinguish between non-personal data and personal data, the latter being the main type of data that can be used to facilitate price discrimination.

More specific examples of unilateral data sharing with competitors can be found in the practices of dual-mode online platforms such as Alibaba, Amazon, JD.com, Target, Walmart, etc. They run a marketplace for third-party sellers but also sell their own in-house products on the marketplace, thus competing with third-party sellers (Hagiu et al. 2022). These platforms use their advantage in collecting and analyzing consumer data to help third-party sellers to personalize their offers to customers. For example, Amazon provides third-party sellers with some raw data on consumer identification and consumers' transaction details, together with the aggregated data on business performance, consumer behaviour and market trends (PPMI, 2020). Moreover, Amazon helps third-party sellers on its platform, some of whom are competitors to Amazon's in-house brands, to target their coupons to specific groups such as Amazon Prime customers, mothers, students, customers who have viewed or purchased from certain Amazon Standard Identification Numbers, etc. As we show in this paper, this type of unilateral data sharing can benefit both Amazon and its third-party sellers, irrespective of how third-party sellers personalize their offers using the shared data and whether Amazon charges fees for their platform service.

Chinese e-commerce giants, Alibaba and JD.com recently started offering free data analytics services to sellers on their platforms through their initiatives, called the A100 program and the Zu Chongzhi platform, respectively. They profile customers by collecting and analyzing their data about gender, age, occupation, location, purchase history, keywords used in search, payment method, brand and product preferences, etc. Alibaba leverages its customer data to help its online sellers to design products and marketing strategies in order to target specific consumer segments, but goes on further to provide a promotional tool that can be used to implement price discrimination (Zhang et al., 2020). In this sense, Alibaba shares data with sellers on its platform not only directly but also indirectly through data analytics and other promotional tools. Some sellers can even improve the targetability of their marketing campaign in their brick-and-mortar shops. For instance, Bestore, a Chinese snack food chain, encourages shoppers to save their facial data when they pay with Alibaba's face-scanning payment tablets in its stores, and crafts campaigns to target specific niches of customers by utilizing data shared by Alibaba.[31] At the same time, Alibaba and JD.com compete with these sellers in the offline market by investing heavily in their own brick-and-mortar stores in recent years. Although more details about the data shared by these platforms are not available, it is conceivable that these platforms choose what to share and how in a way that is mutually beneficial so that sellers choose to remain on their platforms.

5.2 Implications for management

The main takeaway from our analysis is that strategic use of customer data can be an example of a fat-cat strategy. A data-rich firm can choose a proportion of customer data to share with a data-poor competitor, which can then soften price competition. In doing this, two general conditions need to be met. First, shared data should be on the segment of the market that is likely to be more loyal to the competitor, but the data on the competitor's most loyal segment needs to be withheld. This condition does not seem hard to satisfy given the prevalent use of various market segmentation strategies. Second, the primary use of the data is for price discrimination. Because our logic holds whether shared data is used for personalized pricing or third-degree price discrimination, the latter ubiquitous in practice, we believe the second condition is not hard to satisfy either.

While the strategic benefits of data sharing described above are clear, the question remains whether such data sharing may not infringe competition laws. For example, Article 101(1) of the Treaty on the Functioning of the European Union (TFEU) can limit data sharing among competitors if it restricts competition. An exception justifiable under Article 101(3) is when such data sharing can have beneficial effects that outweigh anticompetitive effects (Feasey and de Streel, 2020, pp. 40-41). As our analysis has shown, the welfare effects of data sharing depend on market conditions and the nature of competition. First, if the data-rich firm increases its market share further as a result of data sharing, then both consumer surplus and total surplus decrease. This is true whether competition is in personalized pricing or in third-degree price discrimination. Second, if data sharing results in the data-rich firm conceding some of its market share to the competitor, then total surplus increases thanks to the improved quality of consumer-firm matching. The effect on consumer surplus

[31] https://www.reuters.com/article/us-china-tech-retail-idUSKCN1TL09D

depends on the nature of competition in this case: consumer surplus decreases when competition is in personalized pricing, but the effect on consumer surplus is ambiguous when competition is in third-degree price discrimination, depending further on market conditions. In view of these implications, one may say that regulators do not have a *prima facie* case against the kind of data sharing studied in this paper.

A necessary condition for implementing the data sharing strategy described in this paper is the availability of data analytics. Possessing a large amount of raw, unstructured data in itself would not enable the data-rich firm to identify and select which data to share with its competitors and how. The firm also needs capabilities to process and analyze the data. Thus, the investment costs in data analytics may work as a technology barrier in implementing the optimal data sharing strategy studied in this paper. However, as rapid advancements in digital technologies and computational sciences reduce the costs significantly and bring down the technology barrier, the strategic use of data sharing identified in this paper will gain more practical relevance.

6 Conclusion

We have shown when a data-rich firm can benefit by unilaterally sharing its customer data with a competitor when the data can be used for price discrimination. Our main point is that the firm can soften competition by sharing data on consumers who are more loyal to the competitor than to the firm itself, but withholding the data on the competitor's most loyal consumers. This induces the competitor to choose a high non-discriminatory price for its most loyal consumers, which allows the sharing firm to raise its own prices. Given the growing prevalence of data sharing among firms, our analysis offers an important insight into the strategic use of data sharing, and its implications for competition policy.

We discuss below some directions for future research. First, we can examine if our insight continues to hold more generally. Although we have employed a stylized Hotelling model for expositional clarity, we expect our key insight to be valid in oligopoly models with product differentiation where one firm has sufficient data advantage over its competitors. But a more general setup will allow us to study additional issues such as the number and identity of the competitors with which the dominant firm would like to share data.[32] Second, given that unilateral data sharing

[32]For example, the most straightforward extension of our model to the case with three firms is a

is often used by dual-mode online platforms, we have considered the case where the platform charges a fixed fee for sales. It would be interesting to extend our model to study dual-mode platforms by incorporating a more general fee structure platforms use, and its impact on data sharing. Third, data does not only flow from platforms to sellers, but the other way around. For example, Amazon is reported to have used data from third-party sellers to develop its own products.[33] A key in designing a successful online marketplace is to gain a better understanding of the interaction of these data flows that go in opposite directions. Finally, the extent and effectiveness of data-sharing are constrained by consumers' privacy choice such as opt-in decisions, which can be endogenized in further research.

Appendix

Proof of Lemma 1

Suppose firm 1 chooses to share data on $[a', b'] \cup [a, b]$ with $a' < b' < a < b$. We prove that firm 1 prefers sharing data on $[a, b]$ only, or $[a', b']$ only. The same argument can be applied to any pairwise comparison and show that firm 1 prefers sharing data on n intervals to sharing data on $n+1$ intervals. Then, by transitivity, firm 1 prefers sharing data on a single interval to any other types of data sharing.

Suppose firm 1 shares data on $[a', b'] \cup [a, b]$. We start with two observations. First, for firm 1 to prefer sharing data on $[a', b'] \cup [a, b]$ to that on $[a, b]$, a necessary condition is $b' > 1/2$. If $b' \leq 1/2$, then sharing data on $[a', b']$ in addition to $[a, b]$ does not affect firm 2's uniform price because $[0, 1/2]$ is not contestable by firm 2, but it decreases firm 1's profit from $[a', b']$. Second, without data sharing, firm 2's profit is $(2z - 1)(1 - F(z))$ where z is the marginal consumer's location, which is maximized at z^N. This implies that data sharing with $b \leq z^N$ does not affect firm 2's choice of uniform price, hence cannot benefit firm 1. These observations lead us to focus on the cases where $b' > 1/2$ and $b > z^N$.

unit-length linear city where fully-informed firm 1 is located at 1/2 and uninformed firm 2 is at 0, and uninformed firm 3 is at 1. It is easy to see that our main insight continues to be valid in this case. But it is less obvious in other cases, for example, a circular city where each firm has a loyal customer base on each side of its location. In the online appendix, we offer some discussions on data sharing among three firms on a Salop circle.

[33]https://www.wsj.com/articles/amazon-scooped-up-data-from-its-own-sellers-to-launch-competing-products-11587650015

Let z^* be the marginal consumer's location that maximizes firm 2's profit, hence firm 2's uniform price is $p_2^* = 2z^* - 1$. There are four possibilities: (i) $z^* = z_1 = b$; (ii) $z^* = z_2 \in (b', a)$; (iii) $z^* = z_3 = b'$; (iv) $z^* = z_4 \leq a'$. Note that, when data is shared only on $[a, b]$, the profit-maximizing marginal consumer's location is either z_1 or z_2.

In the first two cases, i.e., $z^* = z_1$ or z_2, the same z^* is profit-maximizing for firm 2 when only $[a, b]$ is chosen for data sharing, leading to the same uniform price. Then firm 1 is better off sharing data on only $[a, b]$ because additional data sharing on $[a', b']$ reduces its profit from that segment. In the fourth case, $z^* = z_4$, we have $z_4 < \min\{z_1, z_2\}$, hence firm 2's uniform price is lower than when only $[a, b]$ is chosen for data sharing (i.e., $2z_4 - 1 < \min\{2z_1 - 1, 2z_2 - 1\}$). Once again, firm 1 is better off sharing data on only $[a, b]$. This leaves us the third case where $z^* = z_3 = b'$. Given $b' > 1/2$, there are two possibilities. First, if $b' \in (1/2, z^N)$, then, as shown previously, firm 2 will set the marginal consumer's location at z^N when data is shared only on $[a', b']$. Thus firm 1 is better off sharing data only on $[a', b']$ because firm 2's price increases from $2b' - 1$ to $2z^N - 1$. Second, if $b' \geq z^N$, then firm 2 will continue to set the marginal consumer's location at b' when data is shared only on $[a', b']$. In this case, firm 1 is indifferent between sharing data on $[a', b'] \cup [a, b]$ and only $[a', b']$ as firm 1 cannot sell to consumers on $[a, b]$ in either case. ∎

Proof of Lemma 2

It suffices to show that firm 1 cannot benefit from data sharing that induces firm 2 to choose p_{22}. For this, we will show $z' < z^N$, hence $p_{22} < p_2^N$, which proves that firm 1 is worse off with data sharing in this case.

From the definition of z^N, we have $(1 - 2z^N)h(z^N) + 2 = 0$. Define $g(x) := 2 + (1 - 2x)h(x)$. Then, since $h'(x) > 0$ by the monotone hazard rate condition, we have $g'(x) = -2h(x) + (1 - 2x)h'(x) < 0$ for all $x \geq 1/2$. Rewrite z' as

$$z' = \frac{1}{2} + \frac{1}{h(z')} - \frac{F(b) - F(a)}{f(z')} < \frac{1}{2} + \frac{1}{h(z')}.$$

Thus we have $g(z') = 2 + (1 - 2z')h(z') > 0$. Since $g(z^N) = 0$, $g'(x) < 0$ for all $x \geq 1/2$, and $z' > 1/2$, we have $z' < z^N$. ∎

Proof of Proposition 1

Consider data sharing on $[a, b]$ with $b > a$, which is at no loss of generality thanks to Lemma 1. By Lemma 2, firm 1 can benefit from data sharing only if firm 2 uses its uniform price to serve only consumers on $[b,1]$ where $b<1$. Given this, $\pi_1(a,b)$ is as given in (8) and $\pi_2(a,b)$ is as given in (7). We first show $\pi_1(a,b)>\pi_1^N$ and $\pi_2(a,b)>\pi_2^N$ for all $1>b>a=z^N$. Since $p_2^N=2z^N-1$, we can rewrite π_1^N as $\pi_1^N=2z^N F(z^N)-2\int_0^{z^N} xdF(x)$. Then, we have

$$\begin{aligned}\pi_1(a,b)-\pi_1^N &= 2\left(bF(a)-z^N F(z^N)\right)-2\left(\int_0^a xdF(x)-\int_0^{z^N} xdF(x)\right)\\ &= 2(b-z^N)F(z^N)>0 \text{ for all } b>a=z^N.\end{aligned}$$

For firm 2, we have

$$\begin{aligned}\pi_2(a,b)-\pi_2^N &= \int_a^b (2x-1)dF(x)+(2b-1)(1-F(b))-(2z^N-1)(1-F(z^N))\\ &= \int_{z^N}^b (2x-1)dF(x)+\int_b^1 (2b-1)dF(x)-\int_{z^N}^1 (2z^N-1)dF(x) \text{ if } a=z^N\\ &> \int_b^1 (2b-1)dF(x)+\left(\int_{z^N}^b (2z^N-1)dF(x)-\int_{z^N}^1 (2z^N-1)dF(x)\right)\\ &= \int_b^1 (2b-2z^N)dF(x)>0 \text{ for all } b>z^N.\end{aligned}$$

We now show there exists $b'\in(z^N,1)$ such that data sharing on $[z^N,b']$ satisfies firm 2's (IC). This, together with what we showed above, proves that sharing data on $[z^N,b']$ benefits both firms. Fix $a=z^N$ so that we express $\pi_{22}(z^N,b)$ as a function of b only, denoted by $\hat{\pi}_{22}(b):=\pi_{22}(z^N,b)$. It is straightforward to check $\pi_{21}(z^N)=\hat{\pi}_{22}(z^N)$ and $\pi_{21}(1)<\hat{\pi}_{22}(1)$. So mutually beneficial data sharing, if exists, must be for some $b<1$. Recall that $\pi_{21}(b)$ is maximized at $b=z^N$ by the definition of z^N, hence $\pi_{21}(z^N)>\pi_{21}(z)$ for all $z\neq z^N$. But we also have $\pi_{21}(z')=(2z'-1)(1-F(z'))\geq \hat{\pi}_{22}(b)=(2z'-1)(1-F(b)+F(z^N)-F(z'))$. Put together, we have

$$\hat{\pi}_{22}(z^N)=\pi_{21}(z^N)>\pi_{21}(z')>\hat{\pi}_{22}(b).$$

Since $\hat{\pi}_{22}(z^N)>\hat{\pi}_{22}(b)$ for any $b>z^N$, we have $\hat{\pi}'_{22}(z^N)<0$. By the definition of z^N, we have $\pi'_{21}(z^N)=0$. Since $\hat{\pi}'_{22}(z^N)<0$ and $\pi_{21}(z^N)=\hat{\pi}_{22}(z^N)$, there exists $b'>z^N$

such that $\pi_{21}(b') > \hat{\pi}_{22}(b')$., i.e., data sharing on $[z^N, b']$ satisfies firm 2's (IC). ∎

Proof of Proposition 2

Given that $z^N = 3/4$, we have $\mathcal{S} = \{(a, b) \mid 1/2 < a < b,\ 3/4 < b\}$. It is easy to derive firm 2's (IC) as follows:

$$(2b-1)(1-b) \geq \frac{(1+2a-2b)^2}{8}. \tag{13}$$

Since firm 2's profit from setting personalized prices on $[a, b]$ is $\int_a^b (2x-1)dx = a(1-a) - b(1-b)$, we can write firm 2's (IR) as

$$(2b-1)(1-b) + a(1-a) - b(1-b) \geq \pi_2^N = \frac{1}{8}. \tag{14}$$

Given the above, firm 1 serves consumers on $[0, a]$ with personalized prices $p_1(x) = (2b-1)+(1-2x) = 2(b-x)$, hence its profit is $\int_0^a 2(b-x)dx = 2ab - a^2$. Thus, firm 1's problem can be stated as follows:

$$\max_{a,b}\ 2ab - a^2,\ \text{subject to } (a, b) \in \mathcal{S},\ (13), \text{ and } (14).$$

We consider the following two cases.

Case (i): $a < 3/4$**.** In this case, firm 1's problem is

$$\max_{a,b}\ 2ab - a^2,\ \text{subject to } \frac{1}{2} < a < \frac{3}{4} < b \leq 1, (13), \text{ and } (14).$$

We first solve the relaxed problem by considering only constraint (13), after which we verify that the other constraints are satisfied at the solution to the relaxed problem. The Lagrangian is then

$$\mathcal{L} = 2ab - a^2 + \lambda\left((2b-1)(1-b) - \frac{(1+2a-2b)^2}{8}\right).$$

The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial a} = -a(\lambda+2) + b(\lambda+2) - \frac{\lambda}{2} = 0,$$
$$\frac{\partial \mathcal{L}}{\partial b} = a(\lambda+2) + \left(\frac{7}{2} - 5b\right)\lambda = 0,$$
$$\lambda\left((2b-1)(1-b) - \frac{(1+2a-2b)^2}{8}\right) = 0.$$

The above equation system yields

$$a^* \simeq 0.7077,\ b^* \simeq 0.9699,\ \lambda^* \simeq 0.2052, \tag{15}$$

and the associated profit is $\pi_1^* \simeq 0.8720$. Clearly, the above solution satisfies both constraints, $\frac{1}{2} < a^* < \frac{3}{4} < b^* \leq 1$ and (14). Therefore, a^* and b^* in (15) are indeed the optimal solution for the case where $a < 3/4$.

Case (ii): $3/4 \leq a$**.** Firm 1's problem in this case is

$$\max_{a,b}\ 2ab - a^2,\ \text{subject to } \frac{3}{4} \leq a < b \leq 1, (13), \text{ and } (14).$$

Once again, we solve the relaxed problem by considering only the constraints $3/4 \leq a$ and (13). We will verify the remaining constraints are satisfied at the solution to the relaxed problem. Then the Lagrangian is

$$\mathcal{L} = 2ab - a^2 + \lambda\left((2b-1)(1-b) - \frac{(1+2a-2b)^2}{8}\right) + \lambda_1\left(a - \frac{3}{4}\right).$$

The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial a} = -a(\lambda+2) + b(\lambda+2) - \frac{\lambda}{2} + \lambda_1,$$
$$\frac{\partial \mathcal{L}}{\partial b} = a(\lambda+2) + \left(\frac{7}{2} - 5b\right)\lambda,$$
$$\lambda\left((2b-1)(1-b) - \frac{(1+2a-2b)^2}{8}\right) = 0,$$
$$\lambda_1\left(a - \frac{3}{4}\right) = 0.$$

The above equation system leads to

$$a^* = \frac{3}{4},\ b^* = \frac{19}{20},\ \lambda^* = 3,\ \lambda_1^* = 0.5 \tag{16}$$

and the associated profit is $\pi_1^* \simeq 0.8625$. Moreover, firm 2's profit is $(2b^*-1)(1-b^*)+a^*(1-a^*)-b^*(1-b^*) = \frac{37}{200} > 1/8$, and all the remaining constraints are satisfied.

Comparing the two cases, we can conclude that firm 1 will choose $a^* \simeq 0.7077$, $b^* \simeq 0.9699$ and obtain profit $\pi_1^* \simeq 0.8720$. Firm 2's profit is then $\pi_2^* = 0.2059$. ∎

Proof of Proposition 3

It remains to show that consumer surplus decreases when $a^* < z^N$. We compare the total costs consumers incur before and after data sharing. The total costs are the sum of prices paid and transportation costs. Thus the total costs for all consumers before data sharing are

$$\pi_1^N + \pi_2^N + \int_0^{z^N} x dF(x) + \int_{z^N}^1 (1-x)dF(x). \tag{17}$$

After data sharing, the total costs are

$$\pi_1^* + \pi_2^* + \int_0^{a^*} x dF(x) + \int_{a^*}^1 (1-x)dF(x). \tag{18}$$

Subtracting (17) from (18) and noting that $a^* < z^N$, we obtain

$$\begin{aligned}&(\pi_1^* - \pi_1^N) + (\pi_2^* - \pi_2^N) + \int_{a^*}^{z^N} (1-2x)dF(x)\\ =\ &(\pi_1^* - \pi_1^N) + \int_{z^N}^{b^*}(2x-1)dF(x) + \int_{b^*}^1 (2b^*-1)dF(x) - \int_{z^N}^1 (2z^N-1)dF(x) > 0.\end{aligned}$$

In the above, the equality follows from $\pi_2^N = \int_{z^N}^1 (2z^N-1)dF(x)$ and $\pi_2^* = \int_{a^*}^{b^*}(2x-1)dF(x) + \int_{b^*}^1 (2b^*-1)dF(x)$, and the inequality holds because $\pi_1^* \geq \pi_1^N$ and $\int_{z^N}^{b^*}(2x-1)dF(x)+\int_{b^*}^1(2b^*-1)dF(x) > \int_{z^N}^1(2z^N-1)dF(x)$. This proves that consumer surplus decreases after data sharing when $a^* < z^N$. ∎

References

Armantier, O. and O. Richard (2003). Exchanges of cost information in the airline industry. *RAND Journal of Economics*, 34(3): 461-477.

Armstrong, M. and J. Zhou (2022). Consumer information and the limits to competition. *American Economic Review*, 112(2): 534-577.

Arnaut, C., Pont, M., Scaria, E., Berghmans, A., and S. Leconte (2018a). Study on data sharing between companies in Europe. Study prepared for the European Commission DG Communications Networks, Content & Technology by everis Benelux.

Arnaut, C., Pont, M., Scaria, E., Berghmans, A., and S. Leconte (2018b). Study on data sharing between companies in Europe - case studies. Study prepared for the European Commission DG Communications Networks, Content & Technology by everis Benelux.

Bergemann, D. and S. Morris (2019). Information design: A unified perspective. *Journal of Economic Literature*, 57(1): 44-95.

Bounie, D., Dubus, A. and P. Waelbroeck (2021). Selling strategic information in digital competitive markets. *RAND Journal of Economics*, 52(2): 283-313.

Chen, Z., Choe, C. and N. Matsushima (2020). Competitive personalized pricing. *Management Science*, 66(9): 4003-4023.

Chen, Z., Choe, C., Cong, J. and N. Matsushima (2022). Data-driven mergers and personalization. *RAND Journal of Economics*, 53(1): 3-31.

Chen, Y., Narasimhan, C. and Z. J. Zhang (2001). Individual marketing with imperfect targetability. *Marketing Science*, 20(1): 23-41.

Chen, Y. and G. Iyer (2002). Consumer addressability and customized pricing. *Marketing Science*, 21(2): 197-208.

Choe, C., King. S. and N. Matsushima (2018). Pricing with cookies: behavior-based price discrimination and spatial competition. *Management Science*, 64(12): 5669-5687.

Choe, C., Matsushima, N. and M. J. Tremblay (2022). Behavior-based personalized pricing: When firms can share customer information. *International Journal of Industrial Organization*, 82, 102846.

Choudhary V., Ghose A., Mukhopadhyay T. and U. Rajan (2005). Personalized pricing and quality differentiation. *Management Science*, 51(7): 1120-1130.

Crémer J., de Montjoye, Y.-A. and H. Schweitzer (2019). Competition policy for the digital era. Retrieved from https://bit.ly/2JMr3g3.

De Cornière, A. and G. Taylor (2020). Data and competition: a simple framework, with applications to mergers and market structures. TSE Working Paper, No 1076 (revised December 2021).

Ezrachi A. and M. E. Stucke (2016). *Virtual Competition: The Promise and Perils of the Algorithm-Driven Economy*, Harvard University Press.

Feasey, R. and A. de Streel (2020). Data sharing for digital markets contestability: towards a governance framework. Centre on Regulation in Europe, available at https://ssrn.com/abstract=3855489.

Fudenberg, D. and J. Tirole (1984). The fat-cat effect, the puppy-dog ploy, and the lean and hungry look. *American Economic Review, Papers and Proceedings*, 74(2): 361-366.

Fudenberg, D. and J. Tirole (2000). Consumer poaching and brand switching. *RAND Journal of Economics*, 31(4): 634-657.

Fudenberg, D. and J. M. Villas-Boas (2012). Price discrimination in the digital economy. Chapter 10 in *Oxford Handbook of the Digital Economy*, ed. by Peitz, M. and J. Waldfogel, Oxford University Press.

Furman, J., Coyle, D., Fletcher, A., McAuley, D. and P. Marsden (2019). Unlocking digital competition: Report of the Digital Competition Expert Panel. Retrieved from https://bit.ly/33jH58S.

Gal-Or, E. (1985). Information sharing in oligopoly. *Econometrica*, 53(2): 329-343.

Gehrig, T., Shy, O. and R. Stenbacka (2011). History-based price discrimination and entry in markets with switching costs: A welfare analysis. *European Economic Review*, 55: 732-739.

Hagiu, A. and J. Wright (2020). When data creates competitive advantage. *Harvard Business Review*, January-February issue.

Hagiu, A., Tat-How, T. and J. Wright (2022). Should platforms be allowed to sell on their own marketplaces? *RAND Journal of Economics*, 53(2): 297-327.

Jentzsch, N., Sapi, G. and I. Suleymanova (2013). Targeted pricing and customer data sharing among rivals. *International Journal of Industrial Organization*, 31(2): 131-144.

Kim, B.-C. and J. P. Choi (2010). Consumer information sharing: strategic incentives and new implications. *Journal of Economics and Management Strategy*, 19(2): 403-433.

Liu, Q. and K. Serfes (2006). Consumer information sharing among rival firms. *European Economic Review*, 50(6): 1571-1600.

Montes, R., Sand-Zantman, W. and T. Valletti (2019). The value of personal information in online markets with endogenous privacy. *Management Science*, 65(3): 1342-1362.

PPMI (2000). Business user and third-party access to online platform data. Analytical paper 5, Observatory on the Online Platform Economy. Retrieved from https://platformobservatory.eu/app/uploads/2020/09/Analytical-Paper-5-Business-user-and-third-party-access-to-data_final.pdf.

Rhodes, A. and J. Zhou (2022). Personalized pricing and competition. Working paper.

Shaffer, G. and Z. J. Zhang (2002). Competitive one-to-one promotions. *Management Science*, 48(9): 1143-1160.

Shy, O. and R. Stenbacka (2013). Investment in customer recognition and information exchange. *Information Economics and Policy*, 25: 92-106.

Thisse J.-F. and X. Vives (1988). On the strategic choice of spatial price policy. *American Economic Review*, 78(1): 127-137.

Zhang, D., Dai, H., Dong, L., Qi, F., Zhang, N., Liu, X., Liu, Z. and J. Yang (2020). The long-term and spillover effects of price promotions on retailing platforms: evidence from a large randomized experiment on Alibaba. *Management Science*, 66(6): 2589-2609.

Zhang, J. (2011). The perils of behavior-based personalization. *Marketing Science*, 30(1): 170-186.

Table 1: Comparison of outcomes with and without data sharing

PDF	(α, β)	z^N	$[a^*, b^*]$	π_1^*-π_1^N	π_2^*-π_2^N	CS^*-CS^N	TS^*-TS^N
unimodal	$(2, 4)$	0.61	[0.50, 0.98]	+	+	−	+
U-shaped	$(0.9, 0.9)$	0.76	[0.72, 0.97]	+	+	−	+
decreasing	$(0.9, 2)$	0.67	[0.50, 0.99]	+	+	−	+
increasing	$(5, 1)$	0.81	[0.83, 0.96]	+	+	−	−
uniform	$(1, 1)$	0.75	[0.71, 0.97]	+	+	−	+

Note: + indicates positive; − indicates negative; CS denotes consumer surplus; TS denotes total surplus.

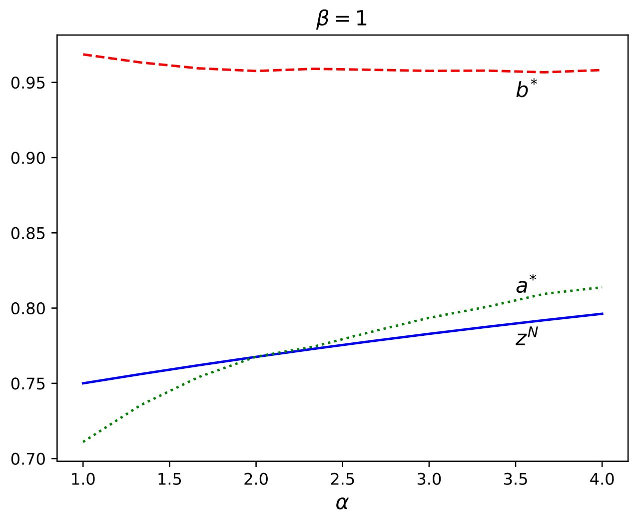

Figure 1: Optimal $[a^*, b^*]$ under beta distribution

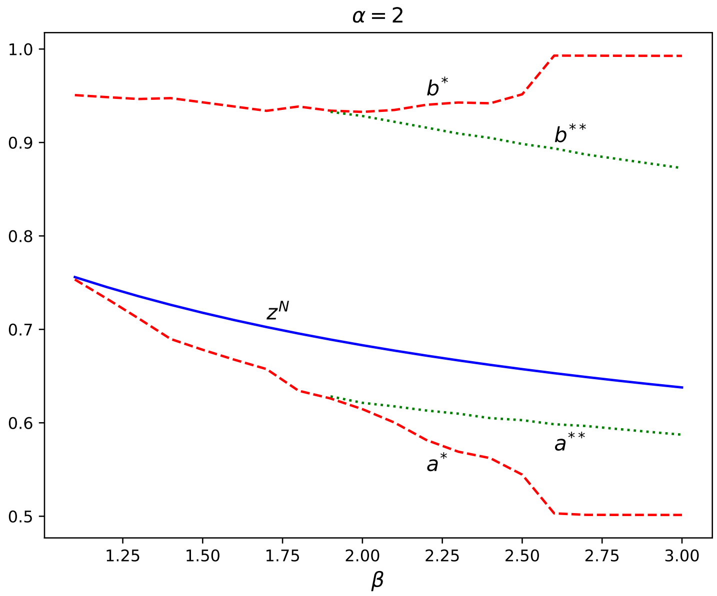

Figure 2: Optimal $[a, b]$ under beta distribution and third-degree price discrimination

Online Appendix

Proof of Proposition 4

We have already shown that firm 2's (IC) is satisfied in the absence of search discrimination. Thus, firm 2's optimal pricing strategy is to choose $p_2 = 2b - 1$ for consumers on $[b, 1]$ and $p_2(x) = 2x - 1$ for consumers on $[a, b]$. Given this, firm 1 chooses $[a, b]$ to maximize its profit $\pi_1(a, b) = \int_0^a (2b - 2x) dF(x)$ subject to firm 2's participation constraint:

$$\pi_2(a, b) = (2b - 1)(1 - F(b)) + \int_a^b (2x - 1) dF(x) \geq \pi_2^N. \tag{A1}$$

We claim that $b = 1$ in the solution to firm 1's problem. One can show that both $\pi_1(a, b)$ and $\pi_2(a, b)$ are increasing in b. This implies that firm 1 can increase profit by increasing b while satisfying (A1). Given $b = 1$, firm 1 then chooses a to satisfy (A1) with equality: $\hat{a}$ solves $\int_{\hat{a}}^1 (2x - 1) dF(x) = \pi_2^N$, hence $\hat{a} > z^N$.

Next, we show $\hat{a} > a^*$. Suppose $a^* \geq \hat{a}$. Then we have

$$\begin{aligned}
\int_{\hat{a}}^1 (2x - 1) dF(x) &= \int_{\hat{a}}^{a^*} (2x - 1) dF(x) + \int_{a^*}^{b^*} (2x - 1) dF(x) + \int_{b^*}^1 (2x - 1) dF(x) \\
&> \int_{a^*}^{b^*} (2x - 1) dF(x) + (2b^* - 1)(1 - F(b^*)) \\
&\geq (2z^N - 1)\left(1 - F(z^N)\right) = \pi_2^N.
\end{aligned}$$

This contradicts the definition of $\hat{a}$.

Finally, we show that no search discrimination reduces consumer surplus and total surplus. The effect on total surplus is clear because $\hat{a} > \max\{z^N, a^*\}$. Consider now consumer surplus compared to the case without data sharing. All consumers who stay with the same firm are worse off because prices are higher. Consumer $x \in [z^N, \hat{a}]$ switches to firm 1 after data sharing, hence incurs total cost $\hat{p}_1(x) + x \geq 2 - x$; without data sharing, her total cost is $p_2^N + (1 - x) = 2z^N - x$. Because the former is larger, she is worse off. Next is the case with data sharing and search discrimination. There are four possible cases: (i) $a^* < z^N < b^*$ and $z^N < \hat{a} < b^*$; (ii) $a^* < z^N < b^*$ and $b^* < \hat{a}$; (iii) $z^N < a^* < b^*$ and $a^* < \hat{a} < b^*$; and (iv) $z^N < a^* < b^*$ and $b^* < \hat{a}$. We prove our claim for case (i) only, as the other cases are similar. When $a^* < z^N < b^*$ and $z^N < \hat{a} < b^*$, all consumers on $[0, a^*] \cup [b^*, 1]$ stay with the same firm, and pay higher prices under no search discrimination. So they are all worse off. Consumer $x \in [a^*, \hat{a}]$ switches to firm 1 under no search discrimination. So her total cost changes from $p_2^*(x) + (1-x) = 2x - 1 + (1-x) = x$ to $\hat{p}_1(x) + x \geq 2 - x$. The latter is larger, so she is worse off. Last, consumer $x \in [\hat{a}, b^*]$ stays with firm 2 and pays the same personalized price. ∎

Proof of Lemma 3

Claim 1: Under the optimal data sharing, we have $q_{2l} = 2a - 1$, hence $z_l = a$.

Denote firm 2's profit from $[a, b]$ by $\tilde{\pi}_{21}$ when it serves the entire segment $[a, b]$, hence $\tilde{\pi}_{21} = (2a-1)(F(b) - F(a))$, and by $\tilde{\pi}_{22}$ when it serves a subset $[z_l, b] \subset [a, b]$, hence $\tilde{\pi}_{22} = (2z_l - 1)(F(b) - F(z_l))$. Given $b > z^N$, we need to consider only the following two cases: (i) $z^N \leq a < b$ and (ii) $a < z^N < b$.

(i) Suppose $z^N \leq a < b$. Define $G(x) := (2x-1)(1-F(x))$. By the definition of z^N, we have $G'(z^N) = 0$ and G is strictly concave due to the monotone hazard rate condition. Thus we have $G'(x) \leq 0$ for all $x \geq z^N$. Define $H(x) := (2x - 1)(F(b) - F(x))$, hence $\tilde{\pi}_{21} = H(a)$ and $\tilde{\pi}_{22} = H(z_l)$ where $a \leq z_l$. To show $z_l = a$, it suffices to show $H(a) \geq H(z_l)$. It is easy to see $H'(x) < G'(x)$ for all $x \in [z^N, 1)$. Since $G'(x) \leq 0$ for all $x \geq z^N$, we have $H'(x) < 0$ for all $x \in [z^N, 1)$. This proves $H(a) \geq H(z_l)$, and hence $z_l = a$.

(ii) Suppose $a < z^N < b$. We will show that it is not optimal for firm 1 to share data that leads to $z_l > a$. Suppose firm 2's profit $H(x) = (2x-1)(F(b) - F(x))$ is maximized at $z_l > a$. The analysis in case (i) above shows that $H(x)$ is decreasing for all $x \geq z^N$. Thus, $z_l > a$ implies that we must have $z_l < z^N$. Firm 1 serves consumers on $[a, z_l]$ with personalized prices $p_1(x) = 2z_l - 2x$ in this case. But firm 1 can increase its profit by sharing data on $[z_l, b]$ instead of $[a, b]$. By doing this, firm 1's profit from $[0, a]$ does not change, but its profit from $[a, z_l]$ increases because its personalized price on this segment increases from $p_1(x) = 2z_l - 2x$ to $p_1(x) = 2b - 2x$. Thus, we must have $z_l = a$.

Claim 2: Firm 1's optimal data sharing should satisfy $\pi_{21}(b) \geq \pi_{22}(a, b)$.

There are two possible cases, $z^N \leq a < b$ or $a < z^N < b$. Since the proof is very similar in both cases, we prove the claim only for the first case. Suppose $z^N \leq a < b$. Firm 2's problem on $[a, b]$ has either a boundary solution where $z_l = a$, or an interior solution where $a < z_l < b$. By Claim 1, we only need to consider the case where $z_l = a$. Suppose $z_l = a$ so that firm 2 serves all consumers on $[a, b]$ with price

$q_{2l} = 2a - 1 > p_2^N$. If $q_{2h} = 2b - 1$, then firm 1's profit is $\pi_1 = \int_0^a (2b - 2x)dF(x)$. If $q_{2h} = 2z' - 1$ for some $z' < a$, then firm 1's profit is

$$\pi_1 = \int_0^{z'} (2z' - 2x)dF(x) < \int_0^a (2b - 2x)dF(x).$$

Thus, firm 1 is better off when firm 2's chooses $q_{2h} = 2b - 1$ to serve $[b, 1]$ only, i.e., firm 2's (IC) needs to be satisfied. ∎

Proof of Proposition 6

(i) Case where firm 1 does not have customer data

We start with the case where firm 1 does not have any customer data. Suppose firm 2 enters the market. In $t = 1$, the marginal consumer is $z = 1/2 + (p_2 - p_1)/2$ where p_i $(i = 1, 2)$ is firm i's uniform price in $t = 1$. At the end of $t = 1$, firm 1 collects data on $[0, z]$, called firm 1's target segment, and firm 2 collects data on $[z, 1]$, called firm 2's target segment. In $t = 2$, firms charge personalized prices for targeted consumers and a uniform price for the rest. This is a standard model of behavior-based price discrimination analyzed in Choe et al. (2018). There are two pure-strategy equilibria, one being a mirror image of the other. Denote firm i's uniform price in period 1 by p_i^1, and its total profit over two periods by Π_i, $i = 1, 2$. By substituting $\delta_c = \delta_f = t = 1$ into Proposition 1 in Choe et al. (2018), we obtain the following in the equilibrium where firm 1's total profit over two periods is larger than firm 2's.[1]

$$p_1^1 = 8/13,\ p_2^1 = 10/13,\ z = 15/26,\ \Pi_1 = 109/169,\ \Pi_2 = 387/676. \qquad \text{(A2)}$$

Thus, firm 2 enters the market if $E < 387/676 \approx 0.572$.

(ii) Case where firm 1 is fully informed but data sharing is not possible

Suppose now firm 1 has data on all consumers. First, suppose that firm 1 does not share data with firm 2. Let z^N be the marginal consumer in $t = 1$ upon firm 2's entry. Denote firm i's personalized price for consumer x in period t by $p_i^t(x)$, whenever relevant, and firm 2's uniform price in period t by p_2^t. In $t = 1$, firm 1 serves $[0, z^N]$ with $p_1^1(x) = 2z^N - 2x$ for $x \in [0, z^N]$ and firm 2 serves $[z^N, 1]$ with $p_2^1 = 2z^N - 1$. In $t = 2$, firm 2 serves $[z^N, 1]$ with $p_2^2(x) = 2x - 1$, and serves additional consumers on

[1] In the other equilibrium, we have $p_1^1 = 10/13$, $p_2^1 = 8/13$, $z = 11/26$, $\Pi_1 = 387/676$, $\Pi_2 = 109/169$.

$[z', z^N]$ with p_2^2 such that $z' = (1 + p_2^2)/2$. Maximizing firm 2's $t = 2$ profit leads to $p_2^2 = z^N - 1/2$, hence $z' = z^N/2 + 1/4 < z^N$ where the inequality is due to $z^N > 1/2$. In $t = 2$, firm 1 serves $[0, z']$ with $p_1^2(x) = 2z' - 2x = z^N + 1/2 - 2x$. Thus, each firm's total profit over two periods can be written as follows:

$$\Pi_1^N = \int_0^{z^N} (2z^N - 2x)dx + \int_0^{z^N/2+1/4} (z^N + 1/2 - 2x)dx = \frac{20(z^N)^2 + 4z^N + 1}{16},$$

$$\Pi_2^N = (2z^N - 1)(1 - z^N) + \int_{z^N}^{1} (2x - 1)dx + (z^N - 1/2)(z^N/2 - 1/4)$$
$$= \frac{-20(z^N)^2 + 28z^N - 7}{8}.$$

Maximizing Π_2^N, we find z^N and the two firms' equilibrium profits given by

$$z^N = 7/10,\ \Pi_1^N = 17/20,\ \Pi_2^N = 7/20. \tag{A3}$$

Thus, if data sharing is not possible, then firm 2 does not enter if $E > 7/20$. Compared to the case where firm 1 does not have any customer data, we can say that data plays a role as an entry barrier when $0.35 < E < 0.572$.

(iii) Case where firm 1 is fully informed and data sharing is possible

Suppose firm 1 shares customer data on $[a, b]$ with firm 2. Denote firm i's personalized price for consumer x in period t by $p_i^t(x)$, and firm 2's uniform price in period t by p_2^t. Consider $t = 1$. As in our baseline model, firm 2 can use uniform price p_2^1 to serve either $[b, 1]$ only or additional consumers to the left of a, say, $[z', a]$. We discuss each case below.

First, if firm 2 chooses $p_2^1 = 2b - 1$ to serve only $[b, 1]$ in $t = 1$, then its target segment in $t = 2$ becomes $[a, 1]$. In $t = 2$, as shown previously, firm 2 chooses $p_2^2 = a - 1/2$ to serve $[a/2 + 1/4, a]$ and charges $p_2^2(x) = 2x - 1$ to serve $[a, 1]$. Then each firm's total profit over two periods can be written as follows:

$$\Pi_1 = \int_0^a (2b - 2x)dx + \int_0^{a/2+1/4} (a + 1/2 - 2x)dx = \frac{1 - 12a^2 + 4a + 32ab}{16},$$

$$\Pi_2 = \int_a^b (2x - 1)dx + (2b - 1)(1 - b) + \int_a^1 (2x - 1)dx + (a - 1/2)[a - (a/2 + 1/4)]$$

$$= \frac{-12a^2 - 8b^2 + 12a + 16b - 7}{8}.$$

Second, if firm 2 chooses $p_2^1 = 2z' - 1$ to serve $[z', a] \cup [b, 1]$ in $t = 1$, then its target segment in $t = 1$ becomes $[z', 1]$. Analogous to the previous case, firm 2 chooses $p_2^2 = z' - 1/2$ to serve $[z'/2 + 1/4, z']$ and charges $p_2^2(x) = 2x - 1$ to serve $[z', 1]$. Firm 2's total profit in this case is

$$\Pi_2' = \int_a^b (2x-1)dx + (2z'-1)(1-b+a-z') + \int_{z'}^1 (2x-1)dx + (z'-1/2)[z'-(z'/2+1/4)].$$

Maximizing the above profit, we obtain $z' = (7 - 4(b - a))/10$, with the resulting equilibrium profit given by

$$\Pi_2' = \frac{-12a^2 + 7(2b - 1)^2 - 16ab + 28a}{20}.$$

As in our baseline model, we can show that firm 1's profitable data sharing must induce firm 2 to choose its uniform price in $t = 1$ to serve only $[b, 1]$. To see this, suppose data sharing on $[a, b]$ induces firm 2 to choose its uniform price to serve $[z', a] \cup [b, 1]$ in $t = 1$ for some $z' < a$. As shown in the proof of Lemma 2, we have $z' < z^N$ in this case. Then, it is straightforward to check that firm 1's $t = 1$ profit under data sharing is smaller than that without data sharing. In $t = 2$, firm 1's profit is $\int_0^{z'/2+1/4} (z' + 1/2 - 2x)dx$ under data sharing, and $\int_0^{z^N/2+1/4} (z^N + 1/2 - 2x)dx$. The latter is larger because $z' < z^N$.

Then, firm 1's problem is to choose $[a, b]$ to maximize Π_1 subject to $\Pi_2 \geq \Pi_2'$ and $\Pi_2 \geq \Pi_2^N = 7/20$. Solving the problem gives us $[a^*, b^*] = [0.710, 0.943]$, with each firm's total profit over two periods given by

$$\Pi_1^* = 1.201, \ \Pi_2^* = 0.431. \tag{A4}$$

Given firm 1's optimal data sharing, firm 2 enters if $E < 0.431$. Note that, without firm 1's data sharing, firm 2 enters if $E < 0.572$. Thus, the possibility of data sharing mitigates the role of customer data as an entry barrier when $0.35 < E < 0.431$. ∎

Analysis for Section 4.8 when firm 1 cannot price discriminate

We show when data sharing can be mutually beneficial even when firm 1 cannot price discriminate and chooses a uniform price. Other than that, everything else is the same as in the baseline model. That is, firm 1 is fully informed, firm 2 is uninformed, and shared data can be used by firm 2 for personalized pricing.

In the benchmark without data sharing, we have the symmetric Hotelling equilibrium where each firm chooses a uniform price equal to 1, firm 1 serves $[0, 1/2]$, and each firm earns profit $1/2$. Consider now the case where firm 1 shares data on $[a, b]$ for $b < 1$. We want to construct an equilibrium such that (i) firm 1 charges the local monopoly's price p_1 for consumers on $[0, a]$, (ii) firm 2 charges the local monopoly's price p_2 for consumers on $[b, 1]$, and (iii) firm 2 charges personalized prices for consumers on $[a, b]$. In what follows, we identify necessary and sufficient conditions for such an equilibrium.

Step 1. Let us first derive firm i's local monopoly price. Firm 1's profit is $\pi_1^m = p_1 \min\{v - p_1, a\}$. When $v - p_1 \geq a$, firm 1's profit is $\pi_1^m = p_1 a$. The optimal price is $p_1 = v - a$ and firm 1 earns $(v - a)a$. When $v - p_1 \leq a$ (i.e., $p_1 \geq v - a$), firm 1's profit is $\pi_1^m = p_1(v - p_1)$. In this case, firm 1's optimal price is $p_1 = v - a$ iff $a \leq v/2$ and is $p_1 = v/2$ iff $a \geq v/2$. In sum, firm 1's optimal local monopoly price is

$$p_1^* = \begin{cases} v - a, & \text{if } 2a \leq v \\ v/2, & \text{if } 2a \geq v. \end{cases} \tag{A5}$$

Firm 1's profit is

$$\pi_1^* = \begin{cases} (v - a)a, & \text{if } 2a \leq v \\ v^2/4, & \text{if } 2a \geq v. \end{cases}$$

Similarly, we can derive firm 2's optimal local monopoly price as

$$p_2^* = \begin{cases} v - (1 - b), & \text{if } 2(1 - b) \leq v \\ v/2, & \text{if } 2(1 - b) \geq v. \end{cases} \tag{A6}$$

Firm 2's profit from the uniform price is

$$\pi_2^* = \begin{cases} (v-(1-b))(1-b), & \text{if } 2(1-b) \le v \\ v^2/4, & \text{if } 2(1-b) \ge v. \end{cases}$$

From the above, there are four different cases for a given tuple (v, a, b), based on the relationship among v, $2a$, and $2(1-b)$: (i) $v \ge \max\{2a, 2(1-b)\}$, (ii) $2a \le v \le 2(1-b)$, (iii) $2(1-b) \le v \le 2a$, (iv) $v \le \min\{2a, 2(1-b)\}$. In all but case (i), the market is not fully covered. In what follows, we focus on case (i). Then, we have $p_1^* = v - a$ and $p_2^* = v - (1-b)$.

Step 2. We check consumers' purchase decisions are optimal. It is easy to check that consumers on $[0, a]$ purchase from firm 1 as their utility from firm 2 is negative; consumers on $[b, 1]$ purchase from firm 2 as their utility from firm 1 is negative. Consumer $x \in [a, b]$ faces personalized price

$$p_2(x) = \min\{v-(1-x), p_1^* + (2x-1)\} = v-(1-x). \tag{A7}$$

All consumers on $[a, b]$ choose firm 2 because their utility from firm 1 is negative.

Step 3. We now establish the conditions under which no firm has incentive to lower its price to poach the rival's customers. Let us start with firm 1. Under deviation price p_1^d, the marginal consumer is $\hat{x}^d = \frac{1}{2} + \frac{p_2^* - p_1^d}{2}$. A profitable deviation requires $\hat{x}^d \ge b$, implying $p_1^d \le v - b$. Firm 1 maximizes deviation profit $\pi_1^d = p_1^d(\hat{x}^d - b + a) = \frac{1}{2}p_1^d(v+2a-b-p_1^d)$ subject to $p_1^d \le v-b$. First, if $(v+2a-b)/2 \ge v-b$, then the optimal $p_1^d = v-b$, implying $\hat{x}^d = b$. But such a deviation is never profitable because firm 1 lowers its price without gaining any extra demand. Next, if $(v+2a-b)/2 \le v-b$, then the optimal $p_1^d = (v+2a-b)/2$ and firm 1's deviation profit is $\pi_1^d = (v+2a-b)^2/8$. Therefore, firm 1 has no incentive to poach firm 2's customers if and only if

$$(v+2a-b)/2 \ge v-b \quad \text{or} \quad (v+2a-b)/2 \le v-b \text{ and } (v+2a-b)^2/8 \le (v-a)a. \tag{A8}$$

Similarly, we can show that firm 2 has no incentive to poach firm 1's customers if and only if

$$v+a+2b \le 3 \quad \text{or} \quad v+a+2b \ge 3 \text{ and } (v+a-2b+1)^2/8 \le (v-(1-b))(1-b). \tag{A9}$$

Step 4: Each firm should earn higher profit under data sharing than in the benchmark without data sharing. These conditions are

$$\pi_1^* = (v-a)a \geq 1/2 = \pi_1^N, \tag{A10}$$

$$\pi_1^* = (v-(1-b))(1-b) + \int_a^b (v-(1-x))dx \geq 1/2 = \pi_2^N. \tag{A11}$$

Step 5. If we can find a tuple (v, a, b) such that $2a \leq v$, $2(1-b) \leq v$, (A8), (A9), (A10), and (A11) hold, then we can say (A5), (A6), and (A7) give us equilibrium prices and the data sharing $[a, b]$ is mutually beneficial. From the above, we can identify a set of (v, a, b) satisfying the following: (i) $1.45 \leq v \leq 2.28$ (large v incentivizes a firm to poach the rival's customers, as shown by (A8) and (A9)); (ii) $0.26 \leq a \leq 0.58$; (iii) $0.46 \leq b \leq 0.84$; (iv) $b < 1/2$ is possible when $v \in [1.52, 1.86]$; (v) b cannot be too close to 1 because, otherwise, firm 2 has incentives to poach firm 1's customers, as shown in (A9).

For example, when $v = 1.8$, data sharing on $(a, b) = (0.5, 0.6)$ is mutually beneficial with equilibrium uniform prices given by $p_1^* = 1.3$ and $p_2^* = 1.4$. In Figure 3, we plot the set of mutually beneficial data sharing (a, b) when $v = 2$. An example is $(a, b) = (0.45, 0.6)$, which leads to $p_1^* = 1.55$ and $p_2^* = 1.6$. Since $\pi_1^* = (v-a)a$ which increases in a for all $a \leq v/2 = 1$ when $v = 2$, firm 1's optimal data sharing is the one with the largest value of a in the shaded region in Figure 3.

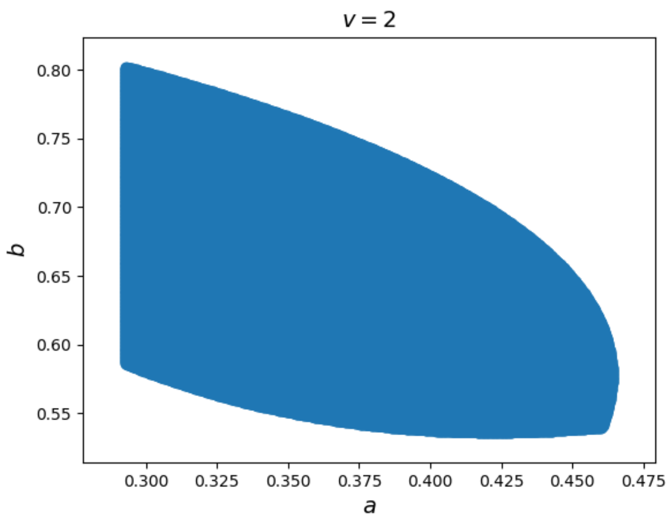

Figure 3: Mutually beneficial data sharing when $v = 2$

Data sharing on a Salop circle

Consider a unit-perimeter Salop circle where three firms are located equidistant from one another and F is a uniform distribution. We label firm 1's location 0 (as well as 1), firm 2's location 1/3, and firm 3's location 2/3. Firm 1 is fully informed and firms 2 and 3 are uninformed. The rest is the same as in our baseline model.

1. Benchmark without data sharing

Let z_1 denote the marginal consumer's location on $[0, 1/3]$ where relevant competition is between firms 1 and 2. Then we have

$$p_1(z_1) + z_1 = p_2 + (1/3 - z_1).$$

Since $p_1(z_1) = 0$, we have $z_1 = p_2/2 + 1/6$. Let $1/3 + z_2$ denote the marginal consumer's location on $[1/3, 2/3]$ where relevant competition is between firms 2 and 3, and firm 1's personalized price is zero for all consumers on this segment. Then we have

$$p_2 + z_2 = p_3 + (1/3 - z_2).$$

This gives us $z_2 = (p_3 - p_2)/2 + 1/6$. Let $2/3 + z_3$ denote the marginal consumer's location on $[2/3, 1]$ where relevant competition is between firms 1 and 3. Then we have

$$p_3 + z_3 = p_1(z_3) + (1/3 - z_3).$$

Since $p_1(z_3) = 0$, we have $z_3 = 1/6 - p_3/2$.

We now solve for the uniform prices chosen by firms 2 and 3. Firm 2's profit is

$$\pi_2(p_2, p_3) = p_2 (1/3 + z_2 - z_1) = p_2 (1/3 + p_3/2 - p_2).$$

Firm 3's profit is

$$\pi_3(p_2, p_3) = p_3 (2/3 + z_3 - (1/3 + z_2)) = p_3 (1/3 + p_2/2 - p_3).$$

Solving the two firms' profit maximization problems simultaneously, we obtain $p_2 = p_3 := p^N = 2/9$, and $z_1 = 5/18, z_2 = 1/6$, and $z_3 = 1/18$. Thus the marginal consumer's location on each segment is given by $5/18, 1/2$, and $13/18$. That is, firm 1 serves $[0, 5/18] \cup [13/18, 1]$ with personalized price $p_1(x) = p_2 + (1/3 - 2x) = 5/9 - 2x$ on $[0, 5/18]$. Firm 1's profit from $[13/18, 1]$ is the same as that from $[0, 5/18]$. Thus its equilibrium profit is

$$\pi_1^N = 2 \int_0^{5/18} (5/9 - 2x)\, dx = 25/162.$$

Firm 2 serves $[5/18, 1/2]$ with $p^N = 2/9$, and firm 3 serves the rest of the market with $p^N = 2/9$. Thus they have the same profit given by

$$\pi_2^N = \pi_3^N = (2/9)\,(1/2 - 5/18) = (2/9)\,(13/18 - 1/2) = 4/81.$$

2. Beneficial data sharing

The above analysis shows that firm 2's customer base is $[5/18, 1/2]$, of which $[5/18, 1/3]$ is the segment that lies between firms 1 and 2, and firm 3's customer base is $[1/2, 13/18]$ with $[2/3, 13/18]$ being the segment lying between firms 1 and 3. Unlike the case with linear city, firms 2 and 3 now have their customer base on each side of their location. For example, the subset $[1/3, 1/2]$ of firm 2's customer base is not contestable by firm 1. This implies that firm 1's withholding of data on the subset of $[5/18, 1/3]$ and/or $[2/3, 13/18]$ does not have the same effect as in the case with linear city, mainly because firms 2 and 3 compete for the segment $[1/3, 2/3]$. On the other hand, firm 1's problem is qualitatively the same as before. The main purpose of data sharing is to induce firms 2 and 3 to increase their uniform prices above $p^N = 2/9$ while the subsequent change in firm 1's market share, if any, is of secondary importance.

We show below that firm 1's sharing of data on entire $[5/18, 1/3]$ with firm 2 and entire $[2/3, 13/18]$ with firm 3 can benefit all three firms relative to the benchmark without data sharing. Given data sharing on $[5/18, 1/3]$, firm 2 chooses p_2 targeting only the segment $[1/3, 2/3]$. Likewise, given data sharing on $[2/3, 13/18]$, firm 3 chooses p_3 targeting only the segment $[1/3, 2/3]$. Let $1/3 + z$ be the marginal consumer on $[1/3, 2/3]$. Then we have

$$p_2 + z = p_3 + (1/3 - z)\,,$$

which gives us $z = (p_3 - p_2)/2 + 1/6$. Firm 2 chooses p_2 to maximize $p_2 z$ and firm 3 chooses p_3 to maximize $p_3(1/3 - z)$. This leads to the symmetric equilibrium uniform price $p_2^* = p_3^* := p^* = 1/3$ and $z = 1/6$. Note that firm 1 chooses personalized price equal to zero for all consumers on $[1/3, 1/2]$. Thus, given $p_2^* = p_3^* := p^* = 1/3$, consumers on $[1/3, 1/2]$ are indifferent between choosing firms 1 and 2, and consumers on $[1/2, 2/3]$ are indifferent between choosing between firms 1 and 3. From the above, it's clear that all three firms benefit from data sharing. First, firm 1 does not lose any market share but can charge higher personalized prices to consumers it serves because $p^* > p^N$. Specifically, firm 1's personalized price increases from $5/9 - 2x$ to $2/3 - 2x$, hence its profit is

$$\pi_1^* = 2\int_0^{5/18} (2/3 - 2x)\, dx = 35/162 > 25/162 = \pi_1^N.$$

Firm 2's market share also stays the same as before but firm 2 charges higher personalized prices to consumers on $[5/18, 1/3]$ and higher uniform price to consumers on $[1/3, 1/2]$. Specifically, for consumer $x \in [5/18, 1/3]$, firm 2's personalized price is given by $p_2(x) + (1/3 - x) = p_1(x) + x$ where $p_1(x) = 0$, hence $p_2(x) = 2x - 1/3 \geq 2/9 = p^N$ for all $x \in [5/18, 1/3]$. Thus firm 2's profit can be calculated as

$$\pi_2^* = \int_{5/18}^{1/3} (2x - 1/3)\, dx + p^* \,(1/2 - 1/3) = 23/324 > 4/81 = \pi_2^N.$$

By symmetry, firm 3 has the same profit as firm 2. Thus, all three firms benefit when firm 1 shares data on $[5/18, 1/3]$ with firm 2 and data on $[2/3, 13/18]$ with firm 3.

The above is one simple example of beneficial data sharing, and there could be other possibilities. But one clear observation is that sharing additional data on $[1/3, 2/3]$ would not increase uniform prices chosen by firms 2 and 3 above the uniform price $p^* = 1/3$. It is because firm 1 chooses personalized price equal to zero for all consumers on $[1/3, 1/2]$, which sets an upper bound on the uniform price that can be chosen by firms 2 and 3. For example, consumer $x \in [1/3, 2/3]$ chooses firm 2 if and only if $p_2 + (x - 1/3) \leq \min\{p_1(x) + x,\ p_3 + (2/3 - x)\} \leq p_1(x) + x = x$ given $p_1(x) = 0$. Thus $p_2 \leq 1/3$.