MONASH University

MONASH BUSINESS SCHOOL

CENTRE FOR GLOBAL BUSINESS DISCUSSION PAPER SERIES

Unilateral Sharing of Customer Data for Strategic Purposes

Chongwoo Choe, Monash University
Jiajia Cong, Fudan University
Chengsi Wang, Monash University

Discussion Paper Number 2021-07
September, 2021

Unilateral Sharing of Customer Data for Strategic Purposes*

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

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

Chengsi Wang†
Department of Economics, Monash University
chengsiwang@gmail.com

September 24, 2021

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 consumers that are more loyal to the competitor while keeping the data on the competitor's most loyal consumers to itself, the firm can induce the competitor to raise its price for consumers it does not have data on. This makes both firms better off than without data sharing.

Keywords: customer data sharing, price discrimination

1 Introduction

In 2017, *The Economist* famously declared that data is the new oil in the digital economy.[1] Customer-generated data analyzed with powerful machine-learning 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). Customer data can also allow firms to sharpen its pricing tools and

*We thank Arthur Campbell, Zhijun Chen, Jay Pil Choi, Heiko Gerlach, Noriaki Matsushima, Nicolas Schutz, Julian Wright, and Xiaojian Zhao for helpful comments. We gratefully acknowledge financial support from the Australian Research Council (grant number DP210102015). The usual disclaimer applies.

†The corresponding author.

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

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 becomes 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.[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 can be better off than without data sharing. There are two key drivers of this result. First, the informed firm strategically selects the proportion of data to share with a competitor. Second, suitably chosen data sharing can induce the competitor to raise its price for consumers it does not have data on, which benefits the sharing firm. Below we sketch our model and the 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 so 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, 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 of the market: for the former, firm 2 chooses personalized prices and, for the latter, it chooses a uniform price. Firm 1 can choose the amount of data to share so that firm 2's uniform price is chosen to serve only those consumers who are very close to firm 2's location. Consequently, firm 2 will raise its uniform price above the level it chooses 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 a subset of consumers in firm 2's customer base. 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 the 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.

We provide some discussions below on the practical relevance of the type of data

[2]This conclusion can be easily derived from the analysis in Thisse and Vives (1998), and is formally shown in many studies. See, for example, Kim and Choi (2010) or Chen et al. (2021).

sharing considered in this paper—B2B sharing of a selected set of data for strategic purposes. Although data sharing is mandated in a few industries,[3] the B2B data sharing considered in this paper is growing in many other industries. For example, everis (2018) reports results from a survey conducted in 2017-2018 covering 129 companies across 24 countries in the European Economic Area. The survey covers various industries such as automotive, transport and logistics, agriculture, telecom, etc., and finds that a majority of companies responded to share only a small proportion of data they generate with the proportion depending on their business strategies, and that most B2B data sharing occurs within their own business sector.

More specific examples can be found in the practices of dual-model online platforms that both run a marketplace for third-party sellers and sell their own private-label products. These platforms allow third-party sellers access to some of the consumer data they collect. For example, Amazon sells its own private brands but also provides third-party sellers some data on consumer identification and consumers' transaction details, together with the aggregated data on business performance, consumer behaviour and market trends (PPMI, 2020). This enables third-party sellers to make more accurate inference about consumer preferences. Another example is from the mobile app market such as Apple Store or Google Play. App developers can directly collect consumer data by tracking consumers' devices. App stores, which also offer their own apps, play a role of gatekeeper by determining how difficult it is for app developers to track consumers' devices (PPMI, 2020).[4] Our final example is from the airline industry. Airlines typically share passenger records for the purpose of quality assurance, but they have discretion on what they share. For example, Delta did not share its frequent flyer data within the data sharing system (Jentzsch et al., 2013).

The rest of the paper is organized as follows. After a brief review of the related studies below, we outline the model in Section 2. In Section 3, we analyze the model when firm 2 can use the shared data for personalized pricing. In Section 4, we consider the case where the shared data can be used for segmented marketing. Section 5 concludes the paper with discussions on the directions for future research and Appendix contains deferred proofs.

Related studies

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.

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;

[3]See Feasey and de Streel (2020) for details and related regulatory issues.

[4]Wen and Zhu (2019) reports that Google introduced about 200 mobile apps on Android between 2008 and 2015, many of which compete with third-party apps available on Google Play. For example, Google Play Books competes with Amazon Kindle, Barnes & Noble Nook, etc.

Armantier and Richard, 2003). More recent studies identify various conditions where mutual information sharing can increase industry profits. These conditions pertain to firms' asymmetric abilities to target customers (Chen et al., 2001), consumer switching costs (Shy and Stenbacka, 2013), or two-dimensional customer information (Jentzsch et al., 2013). In contrast, our focus in on unilateral sharing of customer information. Liu and Serfes (2006) provides a model where unilateral information sharing subject to side payments can increase industry profits. But we show that unilateral information sharing can benefit both firms even in the absence of side payments.

There is a growing body of literature on personalized pricing. Customer 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). This insight is extended to a dynamic model of behavior-based pricing by Choe et al. (2018), and with additional product personalization by Zhang (2011). Chen and Iyer (2002) studies the firm's decision to invest in customer addressability where personalized pricing can be used only for addressable customers. Ezrachi and Stucke (2016) contains discussions and many examples of personalized pricing in practice. Our work contributes to this literature by analyzing the implications of unilateral data sharing, which is not formally studied in the existing literature.

2 The model

Consumers are distributed uniformly on a Hotelling city $[0, 1]$. Firm 1 is located at point zero and has full information on consumers' preferences, i.e., their locations.[5] Firm 2 is located at point one and initially has no information. A consumer derives a gross value v from a product from either firm and incurs a transportation cost d_i, the distance between the consumer's location and firm i's location. So if a consumer buys from a 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 even when served by only one firm.

Firm 1 can costlessly share consumer data with firm 2. Data sharing takes a simple form where firm 1 chooses an interval $[a, b] \subseteq [0, 1]$ such that firm 2 can observe locations of all consumers on $[a, b]$. As we show in Lemma 1 below, firm 1 will never share data on disconnected intervals. We do not consider any monetary payment between firms, as it may create possibilities for collusion. 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 q_2. Given data sharing on $[a, b]$, firm 2 chooses personalized price $p_2(x)$ for $x \in [a, b]$ and the uniform price q_2 for the rest of the consumers. We normalize the cost of production to zero.

The timing of the game is as follows. First, firm 1 makes a data sharing decision by choosing $[a, b]$. Firm 2 will accept any offer of data sharing, as it has an option not

[5] As will be clear, our reasoning continues to apply if firm 1 has information on $[0, b]$ where $b > 3/4$.

to use the shared data for its pricing strategy. The subsequent pricing game proceeds as follows. Firm 2 moves first by choosing its uniform price when relevant. After this, the two firms move simultaneously by making private offers of personalized prices to consumers they have information on. The sequential timing in price offers is standard in the literature on personalized pricing (Thisse and Vives, 1998; Shaffer and Zhang, 2002; Choe et al., 2018; Chen et al., 2020). We also make another standard assumption that firm 2 can prevent consumers who are given private offers from choosing its uniform price. This can be done, for example, through search discrimination.

3 Analysis

3.1 Benchmarks

We start with two benchmark cases. First, consider the case without data sharing. Then the marginal consumer z satisfies $p_1(z)+z = q_2+(1-z)$. Since firm 1 can lower $p_1(z)$ down to zero, we obtain $z = (1+q_2)/2$. Then firm 2 chooses q_2 to maximize $\pi_2 = q_2(1-z)$, hence $q_2 = 1/2$ and $z = 3/4$. Given q_2, firm 1 chooses $p_1(x) = q_2 + (1 - 2x) = 3/2 - 2x$ for all $x \in [0, 3/4]$ and $p_1(x) = 0$ for all $x \in [3/4, 1]$. Thus firm 2 serves the segment $[3/4, 1]$ even without any data sharing, which we call *firm 2's customer base*. The profits and firm 2's uniform price without data sharing are as follows, where the superscript 'N' indicates no data sharing:

$$\pi_1^N = \frac{9}{16},\ \pi_2^N = \frac{1}{8},\ q_2^N = \frac{1}{2}. \tag{1}$$

Second, if firm 1 shares data on all consumers, then we have the outcome in Thisse and Vives (1998): firm 1 serves $[0, 1/2]$ with $p_1(x) = 1 - 2x$ and firm 2 serves the rest with $p_2(x) = 2x - 1$. Thus each firm earns profit equal to 1/4. Clearly, firm 1 is worse off than without data sharing: not only its market share shrinks from $[0, 3/4]$ to $[0, 1/2]$ but also its personalized price decreases from $p_1(x) = 3/2 - 2x$ to $p_1(x) = 1 - 2x$.

3.2 Optimal data sharing

We now turn to the analysis of the full game and solve for a subgame-perfect Nash equilibrium. Let us start with the following observations. First, given data sharing on $[a, b]$, firm 2's problem of choosing the uniform price can be delinked from its choice of personalized prices. This intensifies competition on $[a, b]$, hence 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 indifferent between choosing either firm. Specifically, given firm 2's uniform price q_2, firm 1 will set $p_1(x) = q_2 + (1 - 2x)$. Thus higher q_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

$q_2^N = 1/2$. From these observation, we can show that firm 1 will never share data on disconnected intervals, as claimed earlier. We formally state this below.

Lemma 1 *When firm 1 shares data, it is necessarily on a connected interval.*

Proof: See the appendix.

Given that $[3/4, 1]$ is firm 2's customer base that firm 2 can serve even without data sharing, we consider two possible configurations of $[a, b]$: $b \leq 3/4$ and $b > 3/4$.

Consider first the case with $b \leq 3/4$. Because $[3/4, 1]$ is firm 2's customer base, firm 2 will continue to set $q_2 = 1/2$ to serve all consumers on $[3/4, 1]$. Moreover, data sharing allows firm 2 to choose personalized prices on $[a, b]$. If $b \leq 1/2$, firm 1 serves all consumers on $[a, b]$ but at reduced personalized prices because $p_2(x) = 0$. That is, $p_1(x) = p_2(x) + (1 - 2x) = 1 - 2x < 3/2 - 2x$. If $b > 1/2$, then firm 1 can no longer profitably serve consumers on $[1/2, b]$. Thus data sharing hurts firm 1 in this case.

For the rest of this section, we focus on the case with $b \geq 3/4$. First, suppose $a \leq 3/4 \leq b$. Firm 2's choice of uniform price depends on which segment of consumers it intends to serve. If firm 2 serves only $[b, 1]$, then it chooses $q_2^1 = 2b - 1$ and earns profit

$$\pi_2^1 = (1 - b)(2b - 1). \tag{2}$$

If it serves additional consumers on $[z, a]$ where $z = (1 + q_2)/2$, then its profit is $\pi_2^2 = (1 - b + a - z)q_2$. The profit-maximizing uniform price in this case is $q_2^2 = 1/2 - (b - a)$ and the maximized profit is

$$\pi_2^2 = \frac{(1 + 2a - 2b)^2}{8}. \tag{3}$$

Firm 1 cannot benefit from data sharing in the second case because $q_2^2 \leq 1/2$. Thus firm 1 will choose $[a, b]$ to induce firm 2 to set q_2^1 to serve only $[b, 1]$. Then firm 1 serves $[0, a]$ with personalized price $p_1(x) = q_2^1 + (1 - 2x) = 2b - 2x$, hence its profit is $\pi_1 = \int_0^a (2b - 2x)dx = 2ab - a^2$. Thus firm 1's problem can be stated as follows:

$$\text{Choose } (a, b) \text{ to maximize } \pi_1 = 2ab - a^2 \text{ subject to } a \leq 3/4 \leq b \text{ and } \pi_2^1 \geq \pi_2^2. \tag{4}$$

It is clear that firm 1 chooses $a^* = 3/4$ because sharing extra data on $[a, 3/4]$ is costly to firm 1 without increasing q_2^1. Given $a^* = 3/4$, the constraint $\pi_2^1 \geq \pi_2^2$ is satisfied for $b \in [3/4, 19/20]$. Since π_1 increases in b, firm 1's optimal choice of b is $b^* = 19/20$, leading to its personalized price $p_1^*(x) = 2b^* - 2x = 19/10 - 2x$. Thus firm 1 earns profit $\pi_1^* = 69/80 > 9/16 = \pi_1^N$, the latter being firm 1's profit without data sharing. The larger profit is due to higher personalized prices. By sharing data on $[3/4, 19/20]$, firm 1 induces firm 2 to set $q_2^* = 9/10$, which allows firm 1 to charge $p_1^*(x) = 19/10 - 2x$ to serve consumers on $[0, 3/4]$. Without data sharing, firm 1 serves the same set of consumers, but with lower personalized price $p_1(x) = 3/2 - 2x$ because firm 2's uniform price is $q_2^N = 1/2$. In addition, firm 2 is also better off compared to the benchmark without data sharing. Firm 2 serves the same set of consumers with or without data sharing, but data sharing allows it to charge higher prices to all consumers it serves. Given data sharing, firm 2 serves consumers on $[3/4, 19/20]$ with personalized price $p_2^*(x) = 2x - 1 \geq 1/2$ and consumers on $[19/20, 1]$ with uniform price $q_2^* = 9/10$. Without data sharing, firm 2 sets $q_2^N = 1/2$ for all consumers. We can calculate firm 2's profit given data sharing as $\pi_2^* = 37/200 > 1/8 = \pi_2^N$, the latter being firm 2's profit without data sharing. In sum, the profits and firm 2's uniform price in this case are

$$\pi_1^* = \frac{69}{80}, \ \pi_2^* = \frac{37}{200}, \ q_2^* = \frac{9}{10}. \tag{5}$$

The next case is $3/4 \leq a \leq b$. But this leads to exactly the same analysis as in the case with $a \leq 3/4 \leq b$. In order to benefit from data sharing, firm 1 should necessarily induce firm 2 to choose the uniform price $q_2^1 = 2b - 1$ to serve consumers on $[b, 1]$ only. Clearly, firm 1 benefits from larger b. But increasing b needs to be accompanied by decreasing a to satisfy the constraint $\pi_2^1 \geq \pi_2^2$. As shown in the previous case, firm 1's optimal choice given $3/4 \leq a \leq b$ is $a = 3/4$ and $b = 19/20$.

We can summarize the intuition for our findings 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. But firm 1 should keep data on firm 2's most loyal customers to itself. Sharing data in this way can soften competition as firm 2 will try to extract surplus from its most loyal customers by charging a high uniform price. This allows firm 1 to raise its own personalized prices. Thus firm 1 benefits from data sharing through higher personalized prices but at no cost of reduced market share. The following proposition states our main results.

Proposition 1 *Suppose firm 1 can share data on $[a, b]$ with firm 2, which can be used for personalized pricing. Then firm 1's optimal choice is given by $[a^*, b^*] = [3/4, 19/20]$.*

- *Firm 1 serves consumers on $[0, 3/4]$ with personalized price $p_1^*(x) = 19/10 - 2x$ and earns profit $\pi_1^* = 69/80 > 9/16 = \pi_1^N$.*
- *Firm 2 serves consumers on $[3/4, 19/20]$ with personalized price $p_2^*(x) = 2x - 1$, consumers on $[19/20, 1]$ with uniform price $q_2^* = 9/10$, and earns profit $\pi_2^* = 37/200 > 1/8 = \pi_2^N$.*

Then how does data sharing affect consumer surplus and total surplus? Given full market coverage, total surplus can be proxied by the average distance travelled by a consumer. Since each firm's market share remains the same with or without data sharing, so does the average distance travelled by a consumer. Thus, data sharing is efficiency-neutral. But consumer surplus decreases, simply because both firms benefit from data sharing while total surplus remains the same.

4 Data sharing for segmented marketing

Even when firm 1 shares data, firm 2 may lack capabilities such as data analytics necessary to process the shared data for personalized pricing. This may allow firm 2 to employ only segmented marketing, or third-degree price discrimination, choosing one price for the segment with shared data and another for the rest of the market. But firm 1 continues to use personalized pricing for all consumers. We analyze this case below.

As discussed in the previous section, firm 1 can benefit from data sharing only when it induces firm 2 to increase uniform price above $q_2^N = 1/2$. 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 z_l satisfies $p_1(z_l) + z_l = q_{2l} + 1 - z_l$ with $p_1(z_l) = 0$. Firm 2 chooses q_{2l} to maximize profit from this segment $\pi_{2l} = q_{2l}(b - z_l)$. This leads to $q_{2l} = b - (1/2)$. Since $q_{2l} \leq 1/2$, a necessary condition for beneficial data sharing is $q_{2h} > 1/2$. Then, it is immediate that we must have $b > 3/4$; otherwise, firm 2 will choose $q_{2h} = 1/2$ to serve all consumers on $[3/4, 1]$.

As in Section 3.2, firm 1 can choose $b > 3/4$ and induce firm 2 to set $q_{2h} = 2b-1 > 1/2$ to serve consumers on $[b, 1]$ only. But now, firm 2's profit from $[a, b]$ is smaller than when it can exercise personalized pricing. Therefore, to provide necessary incentives to firm 2 to choose $q_{2h} = 2b - 1$, firm 1 needs to compensate firm 2 with more shared data; otherwise, firm 2 will choose q_{2h} to serve consumers on $[z, a] \cup [b, 1]$ for some $z \leq a$. More data sharing implies firm 1 will expand $[a, b]$ including data on some consumers outside firm 2's customer base, i.e., $a \leq 3/4$, but, at the same time, firm 1 can choose $b \geq b^*$, leading to $q_{2h} > q_2^*$.

Proposition 2 *Suppose firm 1 can share data on $[a, b]$ with firm 2, which can be used for segmented marketing. Then firm 1's optimal choice is given by $[a^{**}, b^{**}] = [(11 + \sqrt{2})/17, (27 + 4\sqrt{2})/34] \simeq [0.73, 0.96]$.*

- *Firm 1 serves all consumers on $[0, a^{**}]$ with personalized price $p_1^{**}(x) = 2b^{**} - 2x$ and earns profit $\pi_1^{**} = 7(26 + 7\sqrt{2})/289 \simeq 0.87 > \pi_1^* > \pi_1^N$, where π_1^* is firm 1's profit when firm 2 can exercise personalized pricing.*

- *Firm 2 serves all consumers on $[a^{**}, b^{**}]$ with $q_{2l}^{**} = 2a^{**} - 1$, all consumers on $[b^{**}, 1]$ with $q_{2h}^{**} = 2b^{**} - 1$, and earns profit $\pi_2^{**} = (71 + 8\sqrt{2})/578 \simeq 0.14$, hence $\pi_2^* > \pi_2^{**} > \pi_2^N$ where π_2^* is firm 2's profit when it can exercise personalized pricing.*

Proof: See the appendix.

Compared to the case where firm 2 can use personalized pricing, firm 1 shares more data with firm 2, including the data on consumers outside firm 2's customer base, i.e., $[a^{**}, 3/4]$. This allows firm 1 to increase b, which induces firm 2 to increase its uniform price to $q_{2h}^{**} \simeq 0.92 > 9/10 = q_2^*$. The benefit from this outweighs the cost from conceding the additional segment $[a^{**}, 3/4]$ to firm 2. Thus firm 1 earns higher profit than in Section 3.2. Simply put, firm 2's reduced ability to price discriminate allows firm 1 to increase its profit by sharing more customer data. Firm 2's profit is smaller than that in Section 3.2, but still larger than that without data sharing.

Firm 1's market share in this case is $[0, (11 + \sqrt{2})/17] \simeq [0, 0.73]$. But firm 1's market share is $[0, 3/4]$ without data sharing as well as when firm 2 exercises personalized pricing. Thus the average distance a consumer travels is now shorter and, therefore, total surplus increases. Larger total surplus combined with firm 2's reduced ability to price discriminate imply that consumer surplus is larger than when firm 2 can exercise personalized pricing, although it is still smaller than that without data sharing.

5 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. This intuition holds whether the shared data is used for personalized pricing or segmented marketing. Given the growing prevalence of data sharing among firms, our analysis offers an important insight into the strategic use of data sharing.

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. Second, since unilateral data sharing is often used by dual-model online platforms, it would be interesting to extend our model to study online platforms and explicitly account for the commission revenue 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.[6] 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.

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

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's profit in this case cannot be larger than when data is shared on $[a, b]$ only, or $[a', b']$ only. This shows that firm 1 prefers sharing data on a single interval to sharing data on two intervals. 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 $[a', b']$. Thus firm 1 is better off sharing data on only $[a, b]$ if $b' \leq 1/2$. Second, without data sharing, firm 2's profit is $(1 - z)(2z - 1)$ where z is the marginal consumer's location, which is maximized at $z = 3/4$. This implies that data sharing with $b \leq 3/4$ 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 $b > 3/4$.

Let z^* be the marginal consumer's location that maximizes firm 2's profit, hence firm 2's uniform price is $q_2^* = 2z^* - 1$. There are four possibilities: (i) $z^* = z_1 = b$; (ii) $z^* = z_2 = (3 - 2(b - a))/4$ if $z_2 \in [b', a)$; (iii) $z^* = z_3 = b'$; (iv) $z^* = z_4 = \max\{1/2, (3 - 2((b - a) + (b' - a')))/4\}$ if $z_4 < a'$. 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 $z_4 < \min\{z_1, z_2\}$. Thus firm 1 is again better off sharing data on only $[a, b]$. This leaves us the third case with $z^* = z_3$. Given $b' > 1/2$, there are two possibilities. First, if $b' \in (1/2, 3/4)$, then, as shown previously, firm 2 will set the marginal consumer's location at $3/4$ when data is shared only on $[a', b']$. Thus firm 1 is better off sharing data only on $[a', b']$. Second, if $b' \geq 3/4$, 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 $[a', b']$. ∎

Proof of Proposition 2

First, suppose $a \leq 3/4 \leq b$. Then firm 2's choice of q_{2h} is the same as that in Section 3.2. That is, beneficial data sharing should necessarily induce firm 2 to choose $q_{2h} = 2b - 1$ to serve only $[b, 1]$. Then, firm 1 serves $[0, a]$ with $p_1(x) = q_{2h} + (1 - 2x) = 2(b - x)$, and serves $[a, z_l]$ with $p_1(x) = q_{2l} + (1 - 2x) = (2b + 1 - 4x)/2$ where $z_l = (1 + 2b)/4$. We need to check $z_l \in [a, b]$. Since $b \geq 3/4$ implies $b > z_l$, we only need to consider two cases: $z_l \geq a$, $z_l \leq a$.

Suppose $z_l \geq a$ or, equivalently, $b \geq (4a - 1)/2$. Then firm 1's profit is

$$\pi_1 = \int_0^a (2b - 2x)dx + \int_a^{(1+2b)/4} \left(\frac{2b + 1 - 4x}{2}\right) dx = \frac{8a(2b - 1) + (1 + 2b)^2}{16}.$$

Firm 1 chooses (a, b) to maximize the above profit subject to constraints (i) $b \geq (4a-1)/2$, (ii) $b \geq 3/4$, (iii) $a \leq 3/4$, and (iv) firm 2 chooses q_{2h} to serve only $[b, 1]$, i.e., $\pi_2^1 \geq \pi_2^2$ as in Section 3.2. We solve the problem ignoring (iii), after which we show that the solution satisfies (iii). Since π_1 strictly increases in a, (i) must be binding. Plugging $a = (2b + 1)/4$ into π_2^2, constraints (ii) and (iv) hold simultaneously iff $b \in [3/4, (27 + 4\sqrt{2})/34]$. Since π_1 strictly increases in b, firm 1 optimally sets $b = (27+4\sqrt{2})/34 \simeq 0.96$, and hence $a = (11 + \sqrt{2})/17 \simeq 0.73$, which also satisfies (iii). Then firm 1's profit is $\pi_1 = 7(26 + 7\sqrt{2})/289 \simeq 0.87$.

Suppose $z_l < a$. In this case, firm 2 serves all consumers on $[a, b]$ by setting $q_{2l} = 2a - 1$. Since firm 1 serves all consumers on $[0, a]$, its problem is the same as in Section 3.2, but with the extra constraint $b < (4a - 1)/2$. However, the solution to firm 1's unconstrained problem, $(a, b) = (3/4, 19/20)$, satisfies the extra constraint. Thus firm 1's profit is the same as in (5): $\pi_1 = \pi_1^* = 69/80 \simeq 0.86$.

From the above, we find that firm 1's optimal data sharing obtains in the first case: $[a^{**}, b^{**}] = [(11 + \sqrt{2})/17, (27 + 4\sqrt{2})/34]$. Then firm 2 serves all consumers on $[a^{**}, b^{**}]$ with $q_{2l}^{**} = 2a^{**} - 1$, all consumers on $[b^{**}, 1]$ with $q_{2h}^{**} = 2b^{**} - 1$, and firm 1 serves all consumers on $[0, a^{**}]$ with $p_1^{**}(x) = q_{2h}^{**} + (1 - 2x) = 2b^{**} - 2x$. The profits and firm 2's uniform price relevant to firm 1's profit are then

$$\pi_1^{**} = \frac{7(26 + 7\sqrt{2})}{289} \simeq 0.87, \ \pi_2^{**} = \frac{71 + 8\sqrt{2}}{578} \simeq 0.14, \ q_{2h}^{**} = \frac{10 + 4\sqrt{2}}{17} \simeq 0.92.$$

As before, the case with $3/4 \leq a \leq b$ leads to the same analysis as the case with $a \leq 3/4 \leq b$. ∎

References

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

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 (2021). Data-driven mergers and personalization. *RAND Journal of Economics*, in press.

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.

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

everis (2018). Study on data sharing between companies in Europe. Study prepared for the European Commission DG Communications Networks, Content & Technology by everis Benelux.

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.

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

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

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

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). Customer information sharing: strategic incentives and new implications. *Journal of Economics and Management Strategy*, 19(2): 403-433.

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

PPMI (2000). Business user and third-party access to online platform data. Analytical paper 5, Observatory on the Online Platform Economy.

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.

Wen, A. and F. Zhu (2019). Threat of platform-owner entry and complementor responses: Evidence from the mobile app market. *Strategic Management Journal*, 40: 1336-1367.

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